

**Universidade Estadual de Campinas**

**EA006 - Trabalho de Fim de Curso**

**A arte e a técnica da Mixagem: I.A. é uma ferramenta ou uma ameaça?**



**UNICAMP**

**Verônica Gesteira Souza - RA 187814**

**Curso 11 - Engenharia Elétrica Integral**

**Orientador: José Augusto Mannis (IA - Unicamp)**

**Campinas / Segundo Semestre de 2021**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca da Área de Engenharia e Arquitetura  
Rose Meire da Silva - CRB 8/5974

So89a Souza, Verônica Gesteira, 1998-  
A arte e a técnica da mixagem I.A. é uma ferramenta ou uma ameaça? /  
Verônica Gesteira Souza. – Campinas, SP : [s.n.], 2021.

Orientador: José Augusto Mannis.  
Trabalho de Conclusão de Curso (graduação) – Universidade Estadual de  
Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Inteligencia artificial. 2. Mixagem (Som). 3. Produção musical. 4.  
Inteligência artificial - Aplicações musicais. I. Mannis, José Augusto, 1958-. II.  
Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de  
Computação. III. Título.

Informações adicionais, complementares

**Título em outro idioma:** The art and the technique of mixing is A.I. a tool or a threat?

**Palavras-chave em inglês:**

Artificial Intelligence

Mixing (Sound)

Music Production

Artificial intelligence - Musical applications

**Titulação:** Engenharia Elétrica

**Data de entrega do trabalho definitivo:** 17-12-2021

## **Resumo**

Mixagem, no contexto da engenharia de áudio, é o nome dado ao processo complexo que envolve a manipulação de um material sonoro de múltiplas faixas e a combinação dessas faixas em um arquivo final. A mixagem passou por diversas transformações ao longo dos anos e com a evolução dos estudos em Inteligência Artificial colocou-se em questionamento qual seria o impacto dessas novas tecnologias sobre o papel do engenheiro de mixagem. Este estudo se propõe, portanto, a investigar qual o estágio atual de automação do processo de mixagem, bem como perspectivas futuras. Para tal, foi inicialmente estabelecido um panorama geral sobre as principais técnicas, ferramentas, métodos e processos utilizados pelos engenheiros de mixagem e também foi levantado as principais linhas de estudo envolvendo Sistemas Inteligentes de Mixagem (IMS). Posteriormente, foram levantadas as ferramentas que utilizam inteligência artificial atualmente disponíveis no mercado e foram elaborados ensaios para verificação de desempenho e grau de pertinência dessas ferramentas. Foi feita então uma análise descritiva do resultado obtido com as ferramentas automáticas e feita uma pesquisa de opinião onde o resultado desses ensaios foram comparados com mixagens executadas de forma mecânica. A partir dos resultados obtidos, foram levantados questionamentos acerca do potencial e dos desafios enfrentados pelos Sistemas Inteligentes de Mixagem.

Palavras-chave: Mixagem, IMS, Inteligência Artificial, Sistemas Inteligentes de Mixagem, Engenharia de Mixagem.

## Sumário

<b>1. Introdução</b>	<b>3</b>
<b>1.1 - O que é mixagem?</b>	<b>3</b>
<b>1.2 - Mixagem através da história do áudio gravado</b>	<b>4</b>
<b>1.3 - Mixagem e a Inteligência Artificial</b>	<b>6</b>
<b>1.4 - Objetivos</b>	<b>8</b>
<b>1.4.1 - Objetivo geral</b>	<b>8</b>
<b>1.4.2 - Objetivos específicos</b>	<b>9</b>
<b>2. Definições iniciais</b>	<b>9</b>
<b>2.1 - Mixagem e seus conceitos</b>	<b>9</b>
<b>2.1.1 - A visão</b>	<b>9</b>
<b>2.1.2 - Os domínios</b>	<b>10</b>
<b>2.1.3 - Os elementos</b>	<b>16</b>
<b>2.2 - Mixagens e suas ferramentas</b>	<b>21</b>
<b>3. Materiais e métodos</b>	<b>34</b>
<b>3.1 Ensaaios</b>	<b>34</b>
<b>3.2 Pesquisa perceptual</b>	<b>40</b>
<b>4. Resultados e análise</b>	<b>41</b>
<b>4.1 Análise descritiva</b>	<b>41</b>
<b>4.2 Análise perceptual</b>	<b>49</b>
<b>5. Conclusão</b>	<b>53</b>
<b>6. Considerações finais</b>	<b>54</b>
<b>REFERÊNCIAS</b>	<b>56</b>
<b>GLOSSÁRIO</b>	<b>60</b>

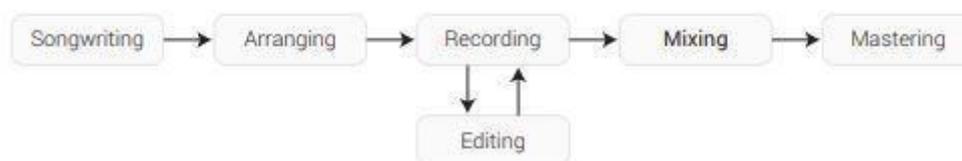
## 1.Introdução

### 1.1 O que é mixagem?

Para a engenharia de áudio, *mixagem* é o nome dado ao processo no qual um material sonoro composto de múltiplas faixas é processado e combinado em um produto final. Se tratando de uma peça musical gravada, o produto final da mixagem (*master track*) é um arquivo de áudio contendo um canal (*mono*), dois canais (*stereo*) ou mais canais dependendo da finalidade. [1]

Segundo De Man: ‘A produção de música gravada(...), da concepção até o consumo, consiste em várias etapas de processos criativos e técnicos. Composições se materializam em vibrações acústicas, que são então captadas, esculpadas, e eternizadas ou amplificadas como sinal eletrônico. Entre a performance de uma música e o comprometimento desses sinais para o meio pretendido, as diferentes fontes(..) são transformadas e fundidas em um único sinal consolidado, em um processo conhecido como mixagem.’[2]

A mixagem é, portanto, um dos estágios finais na concepção de uma música e envolve tomar decisões sonoras acerca de cada elemento musical envolvido e, então, gerar um arquivo final que será masterizado.[3] O processo de mixar é uma tarefa bastante complexa [4] que envolve a aplicação de uma série de ferramentas norteadas por conceitos tanto técnicos quanto estéticos.[5]. A etapa de produção, prévia a etapa da mixagem de uma música, envolve as etapas de composição, arranjo, gravação e edição como descrito na Figura 1.



**Figura 1: Processos presentes na produção e finalização de um produto fonográfico. [1]**

## 1.2 Mixagem através da história do áudio gravado

O início da história do áudio gravado é comumente referido ao ano de 1877 com o surgimento do *fonógrafo*, invenção de *Thomas Edison* que consistia em um cilindro giratório capaz não apenas de armazenar, como também reproduzir informações sonoras[6]. Apesar do sucesso comercial, o fonógrafo de Edison tinha certas limitações como a breve vida útil dos cilindros e a impossibilidade de duplicar um registro sonoro[6]. Como alternativa ao fonógrafo, Emile Berliner propôs um sistema de gravação e reprodução baseado em discos planos - o *gramofone*- que possuía uma série de vantagens em relação ao fonógrafo como maior espaço para armazenamento e a possibilidade de gerar cópias dos registros com facilidade, possibilitando assim o início da comercialização de gravações musicais[7]. À época, realizar uma gravação fonográfica em um estúdio exigia a presença de técnicos que realizavam tarefas de modo a garantir uma gravação inteligível, como escolher o posicionamento correto dos músicos em relação ao único captador sonoro e escolher a sensibilidade do diafragma que iria atuar no captador[6].

Outra contribuição importante para a história do áudio gravado foi dada pelo Dinamarquês Valdemar Poulsen, que inventou o primeiro sistema de gravação com o uso de magnetismo em 1898, o telegrafone, que funcionava através da atuação de um eletroímã que magnetiza um fio de aço de acordo com os sons captados, que podiam ser reproduzidos pela inversão do sistema [6]. Após a segunda guerra mundial, com a invenção do gravador de fitas magnéticas, este meio se tornou o padrão da indústria da música e permitiu a origem da gravação multicanal [8]. Até então, todos os elementos de uma música eram gravadas como um único produto final. Embora a figura do técnico de gravação pudesse se assemelhar à figura de um engenheiro de mixagem, a tarefa de mixagem como conhecemos hoje surgiu apenas no final da década de 1940, justamente com o advento da gravação multicanal.

Em 1954 a companhia estadunidense Ampex lança o primeiro gravador multicanal (com 8 canais de gravação) utilizando a tecnologia de fitas magnéticas, após solicitação do célebre guitarrista Lester William Polsfuss, conhecido como Les Paul, que tinha o desejo de gravar composições em que sua guitarra era sobreposta múltiplas vezes [8]. O grande diferencial do gravador multicanal fabricado pela Ampex era a tecnologia de gravação Sel-Sync (Selective Synchronous ou Sincronia Seletiva em português) que representou um marco na função do engenheiro de mixagem, pois até então as tecnologias não permitiam que

cada elemento fosse manipulado individualmente dentro de uma gravação nem que um instrumento pudesse ser re-gravado de forma independente. [9]

Assim se iniciou a difusão da gravação multicanal, com diferentes empresas surgindo e desenvolvendo produtos que abriram uma nova gama de possibilidades artísticas. Ao longo dos anos 1960, a gravação multicanal chamou a atenção de grandes estúdios e de grandes artistas da época como os Beach Boys e os Beatles, o que impulsionou a evolução dos dispositivos de gravação que ao longo da década de 1970 e no início da década de 1980 foram ganhando cada vez mais funcionalidades e canais de gravação [8]. No final dos anos 1980 e início dos anos 1990, os avanços tecnológicos permitiram a emergência dos primeiros consoles digitais de áudio, o que barateou os custos de produção dos consoles e também popularizou o ato de gravar e registrar sons.[9] Neste período também se iniciou a comercialização dos primeiros computadores pessoais e o surgimento dos primeiros softwares voltados para a gravação e manipulação de som, conhecidos como *Digital Audio Workstation (DAW)*.

Em 1996 a empresa alemã *Steinberg* lançou o software Cubase VST, a primeira DAW que simulou digitalmente o fluxo de trabalho de um console analógico incluindo a possibilidade de incluir efeitos virtuais (*plugins*) de empresas terceiras, o que marcou o início de uma nova era para a história da música gravada [10]. Entre o final dos anos 1990 e início dos anos 2000 novas empresas surgiram e a tecnologia das DAWs evoluiu de forma muito acelerada, de modo que em meados dos anos 2000 a utilização das DAWs se tornou o novo padrão do mercado fonográfico.[8]

Com o avanço e consolidação da tecnologia das DAWs ao longo dos anos 2000, cada vez mais fabricantes de *plugins* começaram a investir em criar e aprimorar novas ferramentas virtuais de áudio (*reverbs*, equalizadores, compressores e assim por diante) que aos poucos foram agradando aos engenheiros de mixagem e tomando o espaço das ferramentas analógicas. Ao longo do final dos anos 2000 e início da década de 2010, o processo de mixagem foi deixando de ser um processo que demandava espaçosos e caros equipamentos e passando a se tornar cada vez mais um processo feito no computador ou *in-the-box*, termo utilizado para se referir ao processo de mixar apenas com *plugins* virtuais (em português a tradução seria dentro da caixa). O processo de mixagem ao longo desta última década, portanto, se tornou mais acessível do que nunca e, embora alguns engenheiros de mixagem ainda utilizem uma mistura entre ferramentas analógicas e virtuais, a mixagem feita

exclusivamente *in-the-box* está se tornando cada vez mais absoluta na indústria fonográfica. [8]

### 1.3 Mixagem e a Inteligência Artificial

O campo de estudo conhecido como *inteligência artificial*, termo este estabelecido pela primeira vez em 1956 [11], pode ser definido de diferentes formas. Segundo Russel e Norvig: ‘(...) a inteligência preocupa-se principalmente com a ação racional. Idealmente, um agente inteligente executa a melhor ação possível em uma situação’[12]. Inteligência artificial, portanto, pode ser definida como o estudo de sistemas que agem de forma racional através de agentes inteligentes. Um agente inteligente é aquilo que irá perceber o ambiente e agir da melhor maneira possível, dado o seu propósito.[12]

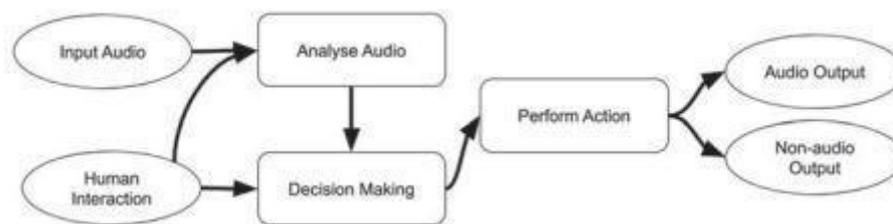
A primeira aplicação de inteligência artificial na música ocorreu em 1957, quando Lejaren Hiller e Leonard Isaacson apresentaram a primeira composição musical construída por um sistema de inteligência artificial.[5] Na mixagem, o termo ‘Mixagem Automática’ foi primeiramente usado por Dugan em 1975 e se referia ao controle de volume de apenas um canal de áudio. Apenas em 2007, no entanto, Gonzalez utilizou o termo em um trabalho sobre automação multicanal, inaugurando o campo hoje conhecido como *Intelligent Mixing Systems* (em português, Sistemas Inteligentes de Mixagem) ou IMS.[13][5][14][15]

O estudo de IMS surgiu com o propósito de investigar o problema da complexidade da mixagem. Segundo Terrell, Simpson e Sandler: ‘A mixagem é um problema quintessencial de otimização. Dado o controle de várias faixas musicais, um equilíbrio deve ser alcançado que reflita um consenso entre métodos de engenharia, objetivos artísticos e restrições de percepção auditiva.’[16] Por essa abordagem, portanto, a mixagem pode ser descrita em termos de parâmetros de controle utilizando um modelo, que pode então ser otimizado por um algoritmo inteligente. No entanto, a mixagem opera em, segundo Moffat, ‘um espaço de estados muito complexo e não-linear’[5] e o estudo de IMS se torna relevante, portanto, à medida em que estuda as possibilidades de ‘reduzir a dimensionalidade do problema da mixagem’.[5] Outra motivação para o desenvolvimento de IMS surgiu com o aumento da demanda por engenheiros de mixagem, uma vez que a acessibilidade aos meios de gravação e produção de música aumentaram drasticamente ao longo dos anos 2000 e 2010. A engenharia de mixagem, no entanto, é uma ocupação que demanda anos de formação e a contratação de

profissionais capacitados pode ter custos elevados. Um IMS poderia ser uma ferramenta auxiliar no processo de tornar a mixagem um serviço mais acessível tanto por meio da agilização do trabalho de engenheiros experientes, quanto por possibilitar que o processo seja feito por não-especialistas. [13]

Os esforços acadêmicos feitos no âmbito da mixagem automática se concentraram tanto em automatizar todo o processo de mixagem, quanto na automação de algumas ferramentas específicas utilizadas no processo de mixagem.[13] Segundo Moffat [5] um IMS possui três aspectos: processo de tomada de decisão, manipulação sonora e interação computador-humano (Figura 2).

O processo de *tomada de decisão* envolve decidir qual será o método empregado na construção do sistema inteligente em si, ou seja, na forma como regras serão modeladas e embutidas em um sistema. Os estudos envolvendo IMS, até o momento, buscaram extrair regras através do uso exclusivo de duas formas de projetar sistemas, ou da combinação de ambas: a primeira busca a extração de regras baseadas na opinião de profissionais experientes e em



**Figura 2: Fluxo de um IMS.[5]**

revisões da literatura e transfere as regras para um sistema usando ontologia ou modelando regras através de processos de *machine learning*; já a segunda, busca a extração de regras através da análise de uma base de dados composta de mixagens previamente executadas, que são aplicadas para treinar o sistema. [15]

O aspecto da *manipulação sonora* diz respeito à maneira como um IMS irá agir para alterar aspectos sonoros de um elemento da mixagem. A manipulação pode ocorrer de duas maneiras: por meio de *efeitos adaptativos* ou por meio de uma *transformação direta*. A manipulação por *efeito adaptativo* procede pela alteração de parâmetros de controle dos recursos aplicados de acordo com uma análise automatizada do sinal de áudio processado ou por uma entrada externa como, por exemplo, por um controlador operado pelo engenheiro de

mixagem ou um sensor. A *transformação direta*, por outro lado, não depende de nenhuma etapa intermediária de interação uma vez que a saída do sistema fornece o produto final programado. [5]

O conceito de *interação computador-humano* estabelece uma escala de níveis de interação humana com um IMS. Esses graus de interação são altamente dependentes do propósito do sistema inteligente e podem ser descritos como: *automático*, *independente*, *recomendativo* e *exploratório*.

No grau *automático* nenhuma interação humana é empregada e o sistema fornece o produto final processado. No grau *independente* várias tarefas são executadas pelo sistema, porém com a supervisão de um engenheiro de mixagem que valida ou não cada ação concluída pelo IMS. No grau *recomendativo* o IMS tem a habilidade de analisar e interpretar em tempo real as ações do engenheiro ao longo do processo de mixagem e fornece ao usuário recomendações de alterações a serem feitas. Finalmente, no grau *exploratório* o IMS apenas fornece informações ao engenheiro acerca do estágio atual dos processos de mixagem.[5]

O desenvolvimento tecnológico de IMS traz, portanto, claras oportunidades de avanço tanto no campo da Inteligência Artificial, à medida em que busca a solução de um problema complexo, quanto no campo da mixagem musical, uma vez que busca a sistematização de um processo muitas vezes considerado como pouco evidente e subjetivo.[5] Embora muitos estudos relevantes sobre *automação de ferramentas* [17][18][19][20][21] e sobre *automação do processos de mixagem* [14][15][16][22][23][24][25] tenham contribuído para ampliar os limites e desenvolver novas funcionalidades em IMS, ainda não há um entendimento claro sobre qual pode ser o impacto artístico e técnico efetivos desses recursos nas maneiras de mixar e no comportamento dos engenheiros de mixagem.

## **1.4 Objetivos**

### **1.4.1 - Objetivo geral**

Investigar o potencial impacto que tecnologias em Inteligência Artificial têm sobre a engenharia de mixagem.

## 1.4.2 - Objetivos específicos

- A) Estabelecer um panorama histórico sobre técnicas, métodos e processos de mixagem.
- B) Apresentação técnica sobre *o funcionamento das principais ferramentas* utilizadas no *processamento de sinais* para mixagem.
- C-a) Levantar as principais *soluções em inteligência artificial* aplicadas ao âmbito da mixagem disponíveis no mercado.
- C-b) Realizar ensaios para *verificação de desempenho e grau de pertinência* dessas ferramentas.
- D) Realizar uma *pesquisa de percepção comparativa* entre mixagens feitas de forma mecânica e mixagens feitas de forma automática.
- E) Fazer uma *reflexão crítica* a partir da *análise dos resultados* para *avaliar* o potencial e a pertinência das *soluções em Inteligência Artificial* bem como os *problemas* identificados e os *desafios* para esse tipo de implementação.

## 2. Definições iniciais

### 2.2 Mixagem e seus conceitos

#### 2.1.1 - A visão

Segundo De Man, Boerum, Leonard, King, Massenburg e Reiss: ‘A busca por regras subjacentes do processo de mixagem é complicada uma vez que muitas versões de uma mesma música podem ser consideradas comercialmente viáveis’.[26] Não é possível, portanto, determinar regras para o processo de mixagem que determinem qual o resultado emergente correto ou incorreto. No entanto, o processo geral da mixagem é guiado pela visão inicialmente estabelecida pelo engenheiro responsável para a produção em realização. A elaboração da visão de mixagem é uma etapa crucial do processo como um todo, sendo este fortemente iterativo, o que norteará todas as ações do engenheiro de mixagem<sup>1</sup>. [1]

A *visão de mixagem*, também chamada de *conceito de mixagem*, é estabelecida normalmente utilizando uma combinação de práticas comumente aplicadas ao gênero ou

---

<sup>1</sup> Em inglês o *engenheiro de mixagem* é denominado pelo termo *mixer*.

estilo do produto sendo mixado e de referências observadas em outras produções similares. Alguns gêneros e estilos possuem particularidades próprias à determinadas linhas estéticas que são muito observadas no fazer artístico, bem como pelos usuários e consumidores finais. O engenheiro também pode se basear na análise de mixagens de outras músicas ou produções que eventualmente se relacionem ao projeto em curso para elaborar uma visão de mixagem.[3]

### 2.1.2 - Os *domínios* envolvidos no processo de mixagem, segundo Izhaki

Segundo Izhaki, “O processo de mixagem envolve o trabalho em cinco domínios (ou aspectos de mixagem principais):

A – *Tempo*,

B – *Frequência (Balanço Tonal)*,

C – *Intensidade (Loudness)*,

D – Espaço – *Estereofonia: Relevô sonoro*

E – Espaço – *Profundidade e Espacialidade*

Podemos considerar os domínios *estereofonia* e *profundidade* como constituintes de um domínio superior - o *espaço*”[1].

#### **(A) Domínio do tempo**

Apesar do *domínio do tempo (A)* ser direto, imediato e incontornável, ele tem sido pouco tratado por diversos autores. Já os demais domínios têm sido muito mais discutidos.

#### **(B) Domínio da frequência**

A compreensão do *domínio da frequência (B)* é essencial para o domínio das técnicas de mixagem. O ouvido humano é capaz de perceber ondas sonoras com frequências entre 20 Hz e 20 kHz. Esse domínio é normalmente dividido em partes, sobretudo registros, porém definidos com diferentes critérios. Um deles [8] considera os seguintes registros de frequências:

- subgraves (abaixo de 60 Hz),

- graves (de 60 a 250 Hz),
- médio-graves (de 250 Hz a 2 kHz),
- médio-agudas (de 2 a 4 kHz)
- agudas (acima de 4 kHz).

A manutenção de um equilíbrio adequado entre as regiões das frequências denomina-se *balanço tonal* sendo um dos pilares do processo de mixagem. Um dos principais problemas decorrentes da falta de *balanço tonal* é o de *mascamamento* que ocorre quando elementos sonoros não são percebidos devido uma sobreposição de informações em uma mesma banda de frequências. Para ser inteligível, todo *elemento de mixagem* deve poder ser percebido em um ponto ou uma região definida no espaço da imagem sonora estereofônica (localização no espaço) e ser distinguido na escuta com clareza no domínio da frequência. Nas regiões subgraves e graves o *mascamamento* tende a se tornar problemático, uma vez que, para a percepção humana<sup>2</sup>, essas bandas de frequência precisam ter muito mais energia para serem percebidas com a mesma intensidade que as bandas médias e médio-agudas [27].

No jargão profissional há uma série de nomes subjetivos atribuídos a certas situações ocorrendo nas várias bandas de frequência e por muitos serem metáforas, facilitam a comunicação e a compreensão de conceitos relacionados à sensações envolvidas em cada caso.[1]

A banda de frequências entre 4 kHz e 6 kHz é chamada de faixa de *presença*. Essa região caracteriza a sensação de proximidade e tangibilidade das vozes, dos instrumentos ou qualquer fonte sonora.

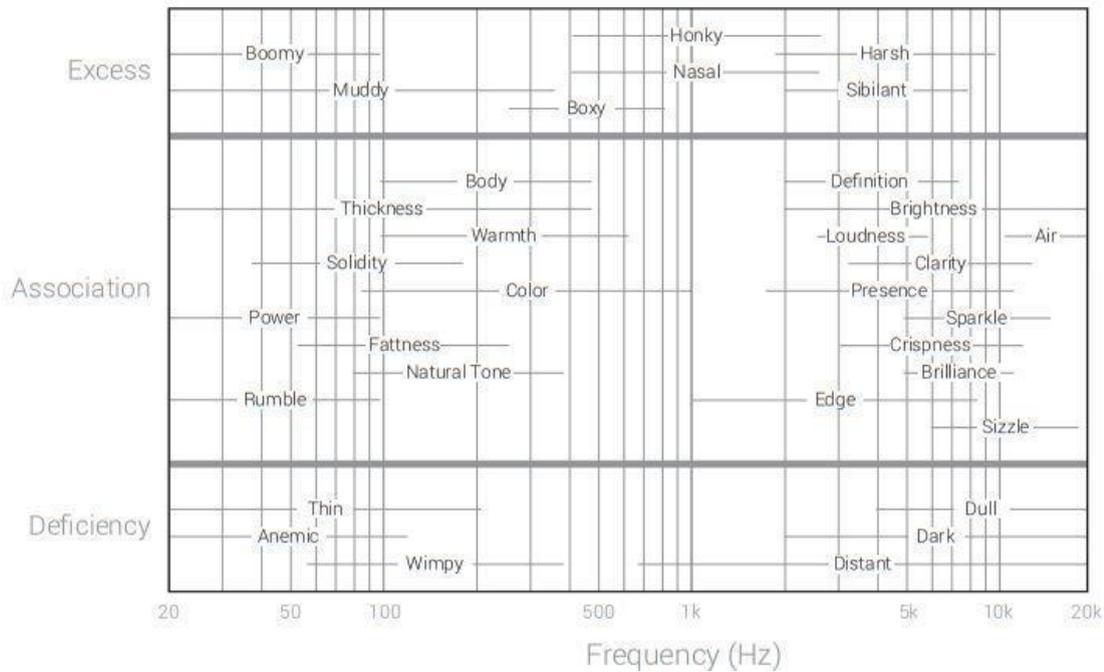
A uma saliência demasiada na banda entre 300 e 800 Hz denomina-se o efeito de *encaixotamento* (em inglês *boxy*), pois nessa região normalmente se encontram frequências excessivas nos modos de vibração de pequenas salas.

Um excesso de informação nesta região causa uma *sensação de enclausuramento*, daí o termo *boxiness*. Contudo é preciso também considerar que nessa mesma faixa de frequências estão as componentes fundamentais da maioria dos instrumentos e das vozes masculinas. [28]

---

<sup>2</sup> Curvas Isofônicas de Audibilidade de Fletcher-Munson ... [completar](#)

Na Figura 3 é possível observar alguns dos jargões utilizados pelos engenheiros de mixagem para se referir a esses efeitos.



**Figura 3: Principais jargões utilizados para se referir a efeitos subjetivos relativos a cada faixa em frequência.[1]**

### **(C) Domínio da Intensidade - Loudness**

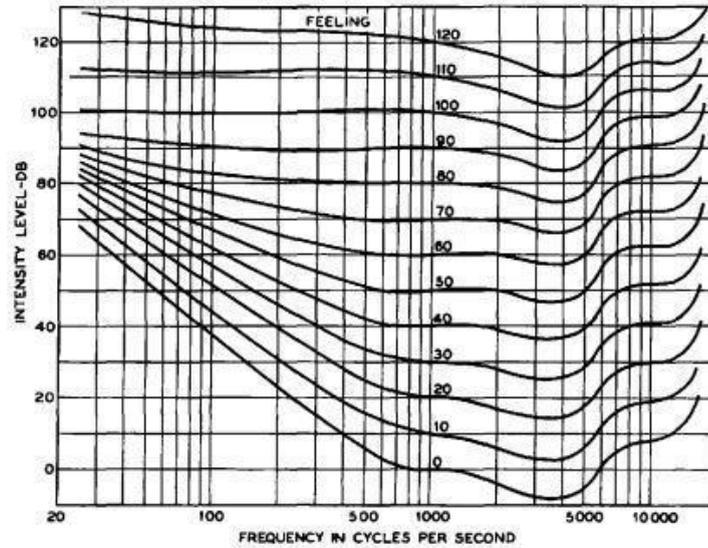
O nível de intensidade no contexto de uma mixagem é manipulado sempre de forma relativa, ou seja, deve ser definido de modo a que os elementos de mixagem sejam inteligíveis entre si e estejam presentes de maneira coesa. Normalmente a prioridade de destaque é atribuída ao elemento sonoro de maior importância em um determinado contexto de uma passagem, um trecho ou por vezes em toda uma obra em particular.[1]

É importante ter em mente dois aspectos em relação ao gerenciamento dinâmico: (1) que cada processamento aplicado a um sinal pode aumentar ou diminuir seu nível individual; e (2) que em cada sistema de processamento e mixagem existe um nível máximo em decibels a ser respeitado além do qual o sinal é distorcido e, portanto, deve-se permanecer aquém

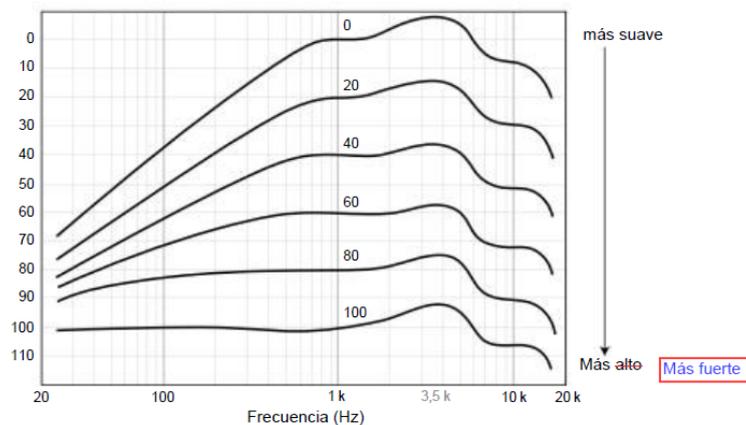
desse limite. O planejamento e definição dos níveis dos sinais sonoros ao longo da mixagem é chamado de *estrutura de ganho*. [3]

A sensação humana de intensidade sonora, denominada pelo parâmetro perceptivo *loudness*, envolve fenômenos complexos e não lineares, requerendo especial atenção e cuidado para sua gestão em todos os processos de produção fonográfica. Segundo Fletcher e Munson: '*Loudness* é um termo psicológico usado para descrever a magnitude de uma sensação auditiva' [29]. Uma outra definição é dada por Scharf: 'Loudness é a intensidade subjetiva de um som'. [30] Embora muitos estudos sobre *loudness* tenham sido conduzidos, não existem até o momento teorias suficientemente abrangentes que expliquem completamente a sensação de *loudness*. Existem, no entanto, alguns modelos psicoacústicos e fisiológicos parametrizando e medindo essa percepção que têm contribuído para um melhor entendimento sobre esta questão. [31] Um deles é o modelo de Fletcher-Munson que resulta de um experimento relatado em uma publicação em 1933, no qual um grupo de **pessoas** escutava um tom de frequência desconhecida e em seguida um outro tom de 1 kHz como referência, devendo em seguida escolher qual dos dois sinais era o mais **forte em intensidade**. [29] Esse procedimento foi repetido com o tom de referência sendo reproduzido em diversos níveis de intensidade.

Deste experimento surgiram as curvas de Fletcher-Munson (Figura 4a), das quais se conclui que quando os níveis em volume estão mais elevados a nossa percepção de *loudness* nas frequências graves e agudas é mais sensível e acentuada (Figura 4b). Com base nisso, recomenda-se que o processo de mixagem deva ser realizado em níveis de intensidade sonora não elevados de maneira a suavizar essa tendência. [1]



**Figura 4a:** Curvas de Fletcher-Munson que descrevem os resultados obtidos acerca da percepção humana de *loudness* representada por curvas isofônicas.



**Figura 4b:** Modificação da variação de intensidade presente nas curvas isofônicas da Figura 4<sup>a</sup> em função do nível de intensidade de referência e da frequência.

### (D.1) Domínio do Espaço – Estereofonia

O tratamento do *espaço* engloba os domínios da *estereofonia* e da *profundidade* e está relacionado à uma maneira tridimensional de pensar a mixagem, na qual cada um de seus elemento pode ser posicionado no espaço considerando a *largura* de sua imagem individual e

sua *profundidade* em relação à percepção do ouvinte.[8] O domínio *estereofonia*, portanto, diz respeito à qualidade da *imagem sonora* resultante, sobretudo de seu *relevo sonoro* (sensação de percepção tridimensional), considerando os aspectos de: localização espacial das fontes sonoras (*localization*); a largura de imagem sonora, tanto a individual de cada fonte sonora quanto à da imagem geral com um todo (*stereo width*); definição do contorno e da localização de cada fonte sonora (*stereo focus or sharpness*); esquema de distribuição espacial das fontes sonoras na imagem sonora geral (*stereo spread*).

Um recurso de tratamento importante para a distribuição dos elementos de mixagem é o **panorama**, através do qual cada material é situado em uma posição definida dentro da cena sonora. Quando se trabalha em dois canais (L-R) essa distribuição se estende de um canal a outro, criando como que uma tela sonora diante do ouvinte. Neste caso, o procedimento mais comum é o do *panorama por intensidade* dosando a porcentagem de energia do sinal que é direcionada a cada um dos canais, conferindo ao elemento sonoro uma posição dentro desse plano.[1]

## **(D.2) Domínio do Espaço – Profundidade e Espacialidade**

O *domínio da profundidade* diz respeito à distribuição panorâmica através de planos sonoros mais distantes (ao fundo) ou mais próximos do ouvinte. A sensação de *profundidade* pode ser obtida através do controle de: (1) nível de *intensidade* sonora (quanto mais forte mais próximo, quanto mais franco mais distante); (2) *equalização* (quando distante, há diminuição das componentes espectrais acima de 1 kHz; quando se aproximam, há aumento das bandas médio-aguda e aguda, com ênfase de 4 a 6 kHz (*presença*)); (3) *ambiência*, através do uso de processamentos de *delay* (atraso) e *reverb* (reverberação), que serão apresentados mais adiante.[3]

O *domínio da espacialidade* diz respeito à sensação de envolvimento do ouvinte pela massa sonora da mixagem e é em grande parte caracterizado pela descorrelação entre os sinais que chegam às orelhas esquerda e direita do ouvinte. Iazzetta et al. (2004) [32] definem, a partir de Barron (1971) [33] e Schroeder (1979) [34], a *Impressão Espacial* (*Spatial Impression*) como o efeito psicoacústico causado pelas reflexões sonoras que atingem o ouvinte fazendo-o criar mentalmente uma sensação acústica espacial do ambiente, sendo influenciada pela difusão sonora e pela dissimilaridade dos sons percebidos por ambas

as orelhas. A grandeza para quantificar essa sensação é a de *Correlação Cruzada Interaural* (*Interaural Cross-Correlation Family*)  $IACF$ ,  $IACC_A$ ,  $IACC_E$  e  $IACC_L$  :

$$IACF_t(\tau) = \frac{\int_{t_1}^{t_2} p_L(t)p_R(t+\tau)dt}{\sqrt{\int_{t_1}^{t_2} p_L^2 dt \int_{t_1}^{t_2} p_R^2 dt}}$$

$$IACF_t(\tau) = \frac{\int_{t_1}^{t=\infty} p_L(t)p_R(t+\tau)dt}{\sqrt{\int_{t=0}^{t=\infty} p_L^2 dt + \int_{t=0}^{t=\infty} p_R^2 dt}}$$

Onde  $L$  e  $R$  designam os sinais recebidos pelos ouvidos Esquerdo (Left) e Direito (Right), respectivamente. O valor máximo da equação acima é a unidade (1), sendo que  $t_1$  corresponde ao instante inicial (0 s – zero segundos) no qual o som direto chega ao ouvinte. A integral de  $t_1$  a  $t_2$  corresponde à energia do som direto, das primeiras reflexões bem como da reverberação até o instante  $t_2$  (ms).

O Coeficiente de Correlação Cruzada Interaural (*Interaural Cross-Correlation Coeficient*) ( $IACC$ ) é dado por:

$$IACC_t = |IACF_t(\tau)|_{\max} \quad \text{sendo } -1 < \tau < +1$$

### 2.1.3 - Os elementos no processo de mixagem

Diferentemente de Izhaki, Owsinski classifica as instâncias funcionais do processo de mixagem em *elementos*, ao invés de *domínios*.

Segundo Owsinski: “Todo gênero musical moderno (...) possui seis elementos principais(...). Eles são:

- A - **Balanço** - a relação entre os níveis de intensidade sonora dos elementos musicais,
- B - **Âmbito de frequência** - ter todas as frequências propriamente representadas,
- C - **Panorama** - posicionar um elemento musical *na imagem sonora*
- D- **Dimensão** - adicionar uma ambiência acústica a um elemento musical localizado

E - **Dinâmica** - controlar os envelopes dinâmicos (de nível de intensidade sonora) de **uma pista** ou instrumento e

F- **Interesse** – tornar a mixagem especial[8].

### **(A) Balanço**

O *balanço* entre os níveis *de intensidade sonora*, embora presente em todo o processo, é o elemento que costuma ser o ponto de partida de uma mixagem, pois está intimamente ligado ao princípio do *estabelecimento da visão* adotado para cada mixagem. O *estabelecimento da visão* e o *balanço de intensidade sonora* inicial dos elementos ocorrem a partir de uma análise preliminar do material a ser mixado, por exemplo, um determinado arranjo de uma canção. A partir dessa análise deve-se encontrar: (a) quais são os principais elementos envolvidos, distribuindo-os por uma escala de prioridade; (b) quais elementos podem estar conflitando entre si no domínio da frequência; e (c) quais as sessões (verso, refrão etc.) de maior importância para a forma da obra em mixagem.

### **(B) Âmbito de frequência**

O *âmbito de frequência*, brevemente explorado no item anterior, está relacionado a *equilibrar as características sônicas de cada um dos elementos de mixagem através da sua manipulação espectral individual*. A principal ferramenta utilizada para esse objetivo é o *equalizador*, que futuramente será aprofundado.[8] Podemos dividir *os aspectos referentes à composição espectral* dos elementos de mixagem (instrumentos, vozes, outros materiais sonoros) em quatro partes: *componente fundamental, parciais harmônicos, parciais inarmônicos e formantes*.

A frequência da componente fundamental de uma onda ou de um sinal sonoro em geral determina a sua altura, sendo esta associada às notas musicais (dó, ré, mi...). As frequências dos parciais harmônicos são grandezas múltiplas inteiras da frequência da fundamental, e as frequências dos parciais inarmônicos são grandezas múltiplas não inteiras da fundamental. Cada um dos *formantes* corresponde a frequências que marcam as zonas nas quais se concentra mais energia em um espectro.[1]

Tendo em mente que o espectro de frequências de cada elemento de mixagem (voz ou instrumento) ocupa um lugar na extensão do *Âmbito de Frequências* de toda a mixagem, ou seja, cada elemento tem uma localização espectral na mixagem como um todo, é possível considerar que a ação de equilibrar a distribuição espectral dos elementos na mixagem seja objetivamente uma ação de *balanço tonal*.

Para um *balanço tonal* adequado é preciso manter sempre a atenção em relação a alguns aspectos próprios a algumas das bandas de frequências.

As frequências sub-graves (até 60 Hz) estão mais associadas à percepção de energia do que de alturas. Nelas normalmente estão localizadas as componentes-fundamentais de sons do registro grave de instrumentos como o *bumbo sinfônico (Gran Cassa)* ou o contrabaixo.

Na faixa de frequências graves (de 120 até 250 Hz) se encontra a componente fundamental das notas graves das vozes masculinas e da maioria dos instrumentos musicais, enquanto que na faixa de médio-graves (de 250 até 2 kHz) aparecem os primeiros harmônicos das vozes e da maioria dos demais instrumentos

A faixa de médio-agudos (até 4 kHz) está muito associada à percepção da voz humana enquanto a faixa aguda (acima de 4 kHz) é onde se concentram muitos dos parciais harmônicos e inarmônicos bem como frequências de vários formantes, responsáveis por dar definição tanto à voz quanto aos instrumentos.[8]

### ***(C) Panorama***

O elemento *panorama*, previamente citado no *domínio da estereofonia*, é um recurso através do qual se pode trabalhar o movimento e a clareza da distribuição espacial dos instrumentos, das vozes e de efeitos na imagem sonora final da mixagem, estando também relacionado à etapa preliminar de *elaboração da visão* adotada para a mixagem. Estão associados ao *panorama* alguns *a priori* estéticos/técnicos de mixagem:

(a) *o espaço central* da imagem sonora é geralmente atribuído aos elementos principais, enquanto que as laterais (*L-R*) normalmente são reservadas ou para elementos aos quais se quer dar movimento; ou que estejam em conflito com alguns dos elementos principais[3];

(b) contudo, considere-se ainda que a distribuição espacial e a movimentação das fontes sonoras também podem ser trabalhadas de maneira plástica, ou seja, usando a totalidade do espaço da imagem sonora como um espaço de livre evolução.

Um problema que pode decorrer da panoramização é o excesso de informação (fontes sonoras) por vezes acumulada em uma determinada área, o que pode causar problemas de inteligibilidade, além de um *desbalanceamento* da imagem estereofônica, quando há mais energia de sinal concentrada em um lado da *imagem sonora* do que no outro[8], a não ser que essa dissimetria esteja sendo propositalmente trabalhada e incorporada no princípio da mixagem como um efeito particular e justificado de espacialização.

#### **(D) Dimensão**

O elemento *dimensão* é aquele pelo qual se ambienta *um instrumento, uma voz ou um outro elemento sonoro da mixagem* sendo trabalhado através do uso de recursos como por exemplo efeitos de reverberação (*reverb*) e de atraso (*delay*), que posteriormente serão introduzidos.

Através desses efeitos se pode moldar o elemento *dimensão* atribuído a um instrumento, voz ou material sonoro. A operação de cada um deles envolve com quatro macro objetivos:

- (a) *recriar um espaço sonoro* no qual o elemento se insere;
- (b) adicionar mais *interesse* à escuta;
- (c) contribuir para que o elemento em mixagem tenha um caráter *espacial tridimensional*;
- (d) implementar *planos e camadas espaciais* mais próximas ou mais distantes do ouvinte[8].

#### **(E) Dinâmica**

A *dinâmica* resulta da gestão e manuseio do *âmbito dinâmico geral* (utiliza-se normalmente o termo em inglês *dynamic range*), que pode ser definido como a diferença entre o nível mais *fraco* e o mais *forte* de um sinal.[28] Segundo Izhaki o elemento *dinâmica*

pode ainda ser dividido ‘entre macrodinâmica e microdinâmica. (...) *Macrodinâmica* está relacionada a variações em nível de intensidade sonora para eventos mais amplos que uma nota musical (...), enquanto *microdinâmica* está relacionada à variação em nível de intensidade sonora ocorrendo no detalhe, ou seja, no acontecimento de cada nota musical tocada(...).

Associamos *microdinâmica* ao envelope dinâmico de cada som individualmente (...)’[1] (*volume x time*) normalmente dividido em quatro partes[35]:

- i) ataque (*attack*),
- ii) decaimento (*decay*),
- iii) sustentação (*sustain*) e
- iv) liberação (ou extinção/desinência) (*release*).

A maneira pela qual se procede à manipulação da *microdinâmica* e da *macrodinâmica* em uma mixagem, pode depender de alguns fatores de técnica de escritura ou de síntese sonora, mas é principalmente vinculada ao *gênero* e ao *estilo* da música a ser mixada. Por exemplo, enquanto o uso de manipulação dinâmica é de suma importância em no repertório *pop* ou *hip-hop*, ele é mínimo, se não raro no *jazz* ou a na música clássica.[8]

### **(F) Interesse**

O elemento ***interesse*** é um conceito definido por Owsinski como: ‘Além de estar tecnicamente correta, uma mixagem deve (...) construir um clímax enquanto mantém pontos de tensão e alívio para manter o ouvinte subconscientemente envolvido’[8]. A tarefa de mixar, portanto, inclui a constante preocupação com a experiência do ouvinte ao longo da música. Quando são necessárias manipulações e operações mais complexas isso pode ser alcançado aplicando recursos de automação de parâmetros. No contexto da mixagem, a automação permite controlar a configuração de parâmetros ao longo da mixagem, podendo ser vinculada a múltiplos parâmetros como o *nível de intensidade sonora* da pista (*track*); o *panorâmico*; ou o envio, à qualidade e quantidade de modificação no processamento por um efeito durante um determinado trecho ou então durante toda a mixagem [3].

## 2.3 Mixagem e suas ferramentas

### Preparação

Antes do processo de mixagem em si, a preparação para a realização da mixagem é crucial. Após terminada a etapa de produção de uma música, cada elemento é individualmente exportado e estes arquivos são então enviados ao responsável pela mixagem. A preparação então se inicia pela organização e checagem dos *stems* (nome dado aos arquivos individuais de cada elemento da música) e um pré-processamento a ser aplicado se necessário, como por exemplo, eliminação de ruído, afinação, ou algum outro ajuste.[8]

Ainda na preparação, algumas escolhas devem ser tomadas. A principal delas é definir como será feito o roteamento dos sinais. É muito comum que se opte pelo agrupamento das faixas em *buses*, que permitem, além do processamento de cada *track* individual, processamentos gerais em todo um grupo de *tracks*. Apesar de haver diversos recursos de agrupamento, normalmente os *buses* são mais frequentemente agrupados em *instrument buses* ou no *mix bus*. Um grupo de instrumentos constitui uma espécie de “canal”, no qual todas as *tracks* (faixas) de um determinado instrumento estão incluídas nele, e são nele processadas de maneira coerente e uniforme. Um exemplo tradicional é o agrupamento da bateria no *drum bus* onde bumbo, caixas, pratos e afins são processados em conjunto. Já o *mix bus* é um canal por onde todos os elementos são enviados antes de serem enviados ao master no qual o engenheiro de mixagem pode agir, geralmente com pequenas alterações, no produto final como um todo.[1]

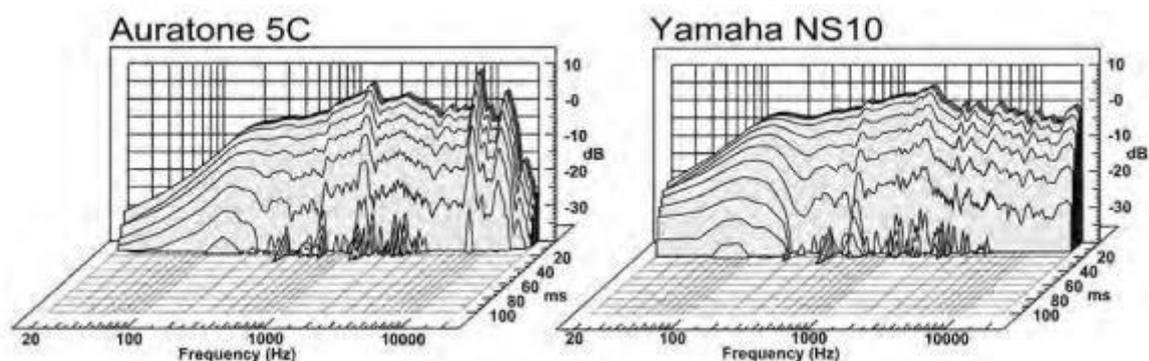
### Monitoramento

Monitoramento é o nome dado ao sistema de escuta através do qual a mixagem é controlada pela escuta, podendo ser um fone de ouvido ou um conjunto de caixas de som (também chamadas de monitores). Como cada circuito de amplificação possui uma resposta em frequência característica, o sistema de monitoramento é um ponto de fundamental importância no processo de mixagem, uma vez que influenciará toda tomada de decisão do engenheiro de mixagem.[1] No caso do uso de caixas de som, outro fator pivotal na

percepção do som é a característica acústica da sala na qual os monitores estão posicionados.[27]

No começo da mixagem moderna os engenheiros acreditavam que a mixagem deveria ser feita no melhor sistema de monitoramento possível, ou seja, no sistema que mais fielmente representasse todas as frequências, no entanto logo os engenheiros perceberam que esses sistemas não refletiam os meios nos quais os consumidores realmente escutavam música, o que significava que os consumidores estavam ouvindo um resultado final muito distante do que os engenheiros de mixagem ouviam no estúdio.[1] Houve então uma mudança de percepção e os estúdios passaram a utilizar caixas de som menores e de mais baixo custo que pudessem simular os sistemas que os consumidores utilizavam. Os modelos de monitores que mais se consagraram por conseguirem com sucesso simular o sistema *low-fi* dos consumidores e ao mesmo tempo ter definição o suficiente para servir de referência para os engenheiros foram a Auratone 5C e a Yamaha NS10.[27]

Desde 1979, com o lançamento do Sony Walkman, o uso de fones de ouvido se tornou um meio cada vez mais frequente no qual os consumidores ouvem música, o que fez com que as mixagens começassem a passar por um processo de checagem, no qual também se avalia como a mixagem será executada através de fones de ouvido. Enquanto alguns engenheiros de mixagem utilizam falantes como meio principal e fones de ouvido apenas para checagem, muitos engenheiros utilizam fones como único sistema de monitoramento. Isso se deve ao fato de que a mixagem em fones de ouvido é mais acessível, já que não requer uma sala acusticamente tratada.[1]



**Figura 5: Resposta em frequência das caixas Auratone 5C e Yamaha NS10.[27]**

## Medidores de volume

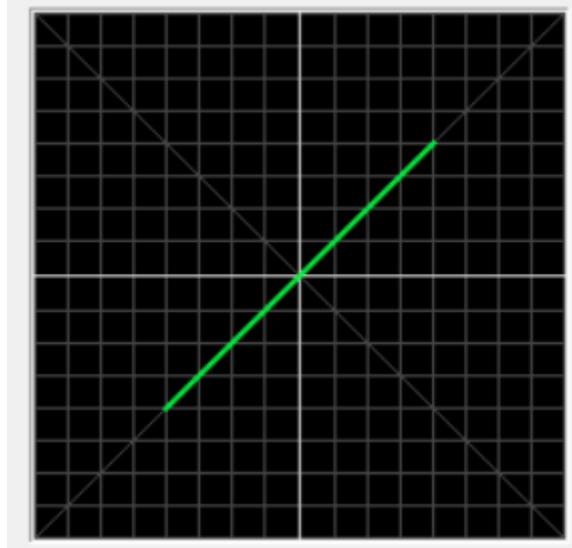
Controlar o volume dos elementos ao longo da mixagem é uma tarefa importante e o uso de medidores é de grande utilidade na medida em que oferecem um controle visual para o engenheiro. Os medidores mais utilizados são os medidores de *pico* e os de *valor médio*. Os medidores de pico mostram o valor de pico do sinal naquele instante (valor instantâneo) e tem como valor máximo em sistemas digitais o 0 dB, já os medidores de valor médio mostram uma média do nível de intensidade sonora do sinal em uma janela de tempo pré-definida. A medição do valor médio tem uma relevância especial, já que os ouvidos humanos têm uma percepção de *loudness* que depende do valor médio do sinal.[1]

## Medidores de fase

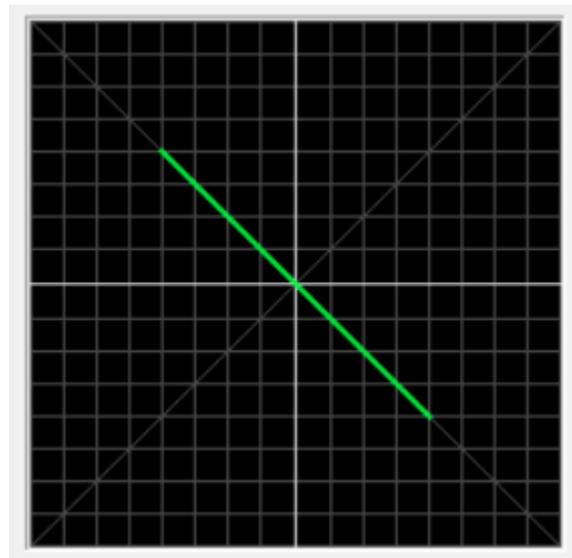
Em relação a um sinal estéreo de dois canais, é preciso ter uma atenção especial ao comportamento do sinal em fase para que haja *compatibilidade mono*. Este é o termo utilizado para se referir à adaptabilidade de uma mixagem estéreo à sua reprodução em um sistema mono (de apenas um canal) com perda mínima de informação do conteúdo, uma vez que os sistemas mono reproduzem um arquivo estéreo através da soma de ambos os canais (esquerdo (L) e direito (R)). Para auxiliar esse monitoramento, os medidores de *fase* e de *correlação* são importantes ferramentas visuais.

O medidor de *correlação* indica a compatibilidade mono de um sinal mostrando o grau de defasagem entre canal direito e canal esquerdo, no âmbito de -1 (180° de defasagem) a +1 (0° de defasagem, ou seja, quando os canais direito e esquerdo forem idênticos).

Já o medidor de *fase* mostra o grau de *descorrelação* ou *correlação* entre os sinais dos canais *esquerdo* e *direito*, através de uma modulação ortogonal como para formar figuras de Lissajous.[36] No display do medidor de fase o nível de intensidade sonora do sinal é descrito pelo tamanho do feixe de modulação e a correlação ou a descorrelação entre os sinais L e R é apontada pela inclinação do feixe de modulação (45° à direita correlação total; 45° à esquerda descorrelação total).



**Figura 6a: Medidor de fase de um sinal totalmente em fase.**



**Figura 6b: Medidor de fase de um sinal em oposição de fase (180°).**

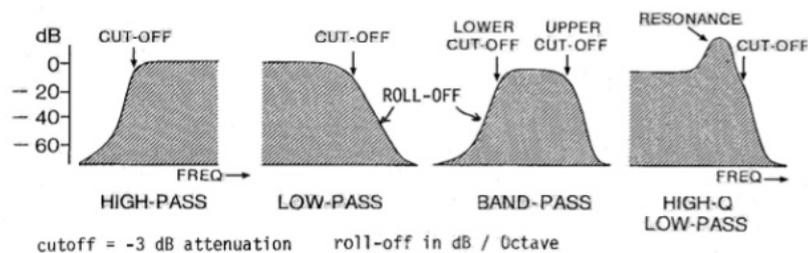
## Equalizadores

Segundo Izhaki: "Distinguir frequências e dominar a manipulação delas é talvez o maior desafio a ser enfrentado em uma mixagem"[1]. O equalizador, mesmo não sendo a única forma de manipular frequências em uma mixagem, é a principal e mais importante ferramenta com esta finalidade.

O equalizador surgiu na década de 1870 no contexto do início da engenharia telefônica com a finalidade de recuperar bandas de frequência em um sinal que foram perdidas ao longo de uma transmissão[37]. Os primeiros equalizadores, assim como os demais equipamentos de áudio naquela época, eram analógicos e, portanto, tinham certas características particulares devido aos componentes utilizados. No final da década de 1970 e início da década de 1980, com o advento do processamento digital de sinais, os primeiros equalizadores digitais surgiram como uma tentativa de discretizar equalizadores de áudio analógicos já consagrados na indústria fonográfica. Com o passar do tempo, equalizadores cada vez mais poderosos em termos de funcionalidades foram surgindo graças aos novos recursos do áudio digital[1].

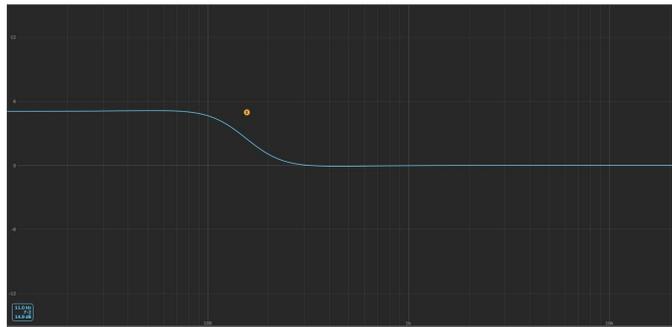
O princípio de *correção de frequência* está interligado à ação de diminuir ou acrescentar energia em determinadas bandas de frequência de um sinal, o que pode ser realizado através de *filtros* e equalizadores. Os *corretores de frequência* podem ser classificados em três grupos: passa-banda, *shelving* (prateleira em português) e paramétricos (também chamado de bell filters, sino em inglês).[1]

Os filtros passa-banda são divididos em dois: filtros passa-alta (FPA) e filtros passa-baixa (FPB). O seu funcionamento é bastante simples: a partir da inserção de uma frequência de referência (frequência de cut-off) as frequências acima serão cortadas (FPB) ou mantidas (FPA) e as frequências abaixo serão mantidas (FPB) ou cortadas (FPA). A taxa na qual essas frequências serão atenuadas dependerá de um outro parâmetro chamado de slope (taxa de inclinação em português), que é medida em dB/oitava e determina quão brusca será a atenuação a partir da frequência de cut-off. Uma slope de 6dB/oitava, por exemplo, significa que a partir da frequência de cut-off cada oitava sofrerá redução de 6 dB.[1] Um FPA e um FPB podem ser utilizados individualmente ou combinados de modo a formarem um filtro passa-banda (FPB), onde serão escolhidas duas frequências de cut-off.

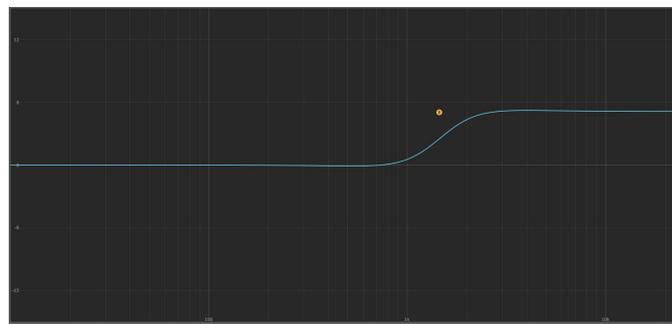


**Figura 7: Representação das curvas dos filtros: passa alta, passa baixa e passa banda.**

Os filtros *shelving* podem não apenas atenuar como também dar ganho a bandas de frequências e possuem três parâmetros: (1) *frequência* de cut-off, (2) *slope* e (3) *ganho*. Existem dois tipos de filtros prateleira: os *low-shelf* e os *high-shelf*. Os *low-shelf* atenuam ou aumentam (dependendo do valor do ganho) as frequências menores que a frequência de cut-off e não alteram as frequências acima. Já o *high-shelf* funciona de maneira inversa. O parâmetro *slope* funciona de forma similar aos filtros passa banda, mas no caso determinará a inclinação (em dB/oitava) com a qual o filtro irá atingir o ganho desejado.[1]



**Figura 8a: Filtro Low-Shelf**

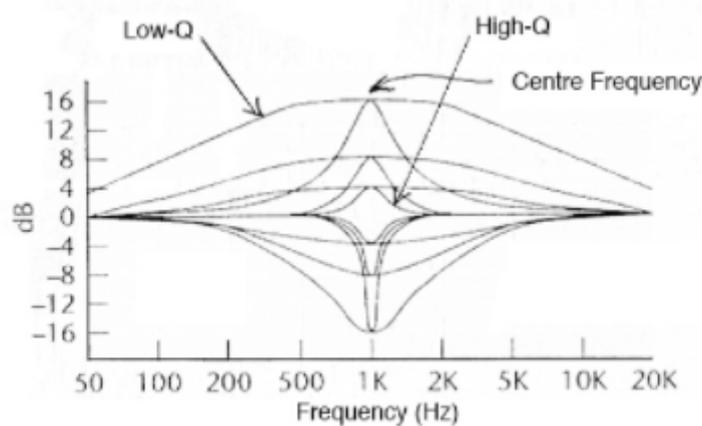


**Figura 8b: Filtro High-Shelf**

As *curvas paramétricas*, são utilizadas para agir em bandas mais precisas de frequências, podendo atenuá-las ou aumentá-las. Os parâmetros utilizados são três: frequência central, ganho e Q (fator qualidade – largura da banda de frequência). Essa curva também é chamada de *bell*, pois **seu** formato se assemelha a um sino. Seu funcionamento é dado da seguinte forma: a frequência central determinada será a frequência de referência e o ganho será a quantidade em dB que esta frequência será atenuada ou levantada. O fator Q diz respeito a quão específico será **sua ação**, ou seja, quanto maior o valor de Q mais estreito, em

termos de banda, será a ação da curva e quanto menor o valor de Q, maior será a banda de frequência que sofrerá alteração.[1]

Existem alguns tipos de equalizadores, os principais sendo dois: os equalizadores gráficos e os equalizadores paramétricos. O equalizador gráfico consiste em uma série de paramétricos com frequência central fixada cobrindo todo o espectro de frequência e são mais utilizados em som ao vivo, não sendo tão usuais no processo de mixagem. O equalizador paramétrico permite a seleção completa dos parâmetros de atuação e essa flexibilidade o torna o modelo mais utilizado na mixagem.[8]



**Figura 9: Curva paramétrica de equalização com diferentes valores de Q.**

A equalização possui uma série de aplicações ao longo do processo de mixagem, as principais sendo:

- *Manipulação tonal*: lapidar as características sonoras e o timbre de cada instrumento.
- *Separação e definição*: diferenciar informações sonoras estejam ocupando uma mesma posição no espectro de frequências de maneira a não se mascararem entre si, permitindo assim que o ouvinte possa distinguir os vários elementos sobrepostos.
- *Ajuste de nível*: caso seja necessário atenuar ou aumentar o volume de uma região específica de frequências.
- *Remoção de conteúdo indesejado*: atenuação de frequências ressonantes indesejadas, ruído ou sons capturados acidentalmente.[1]

## Compressores

O compressor é um dos dispositivos utilizados no controle **do âmbito** de variação dinâmica e sua origem foi motivada pela necessidade de controlar automaticamente a intensidade de sinal para diferentes aplicações.[1] Embora sua função original seja a de **controle dinâmico**, o compressor também é utilizado como uma ferramenta para alterar características timbrísticas do som [38] quando aplicado a somente algumas bandas de frequência. Ao longo do tempo surgiram vários modelos de compressores com diferentes parâmetros. De maneira geral a operação de um compressor envolve os seguintes parâmetros: *threshold* (limiar), *ratio* (razão), *attack* (tempo de ataque), *release* (tempo de liberação/desinência), *gain* (ganho) e *knee* (joelho). [39]

O parâmetro de *threshold* é o nível limiar de intensidade do sinal a partir do qual o compressor passa a atuar: a cada vez que o sinal de entrada ultrapassa o nível de intensidade escolhido como *threshold* o compressor inicia sua atuação para reduzir o volume. Essa redução é determinada pelo parâmetro *ratio*, que é a razão entre a intensidade do o sinal de entrada e a intensidade do sinal de saída.[39] Em outras palavras, um *ratio* de 2:1 implica que para cada 2 dB que o sinal de entrada ultrapasse do valor de *threshold*, o sinal de saída aumentará apenas 1 dB acima do valor do *threshold*. [8] A *transição* entre o sinal não comprimido e o início da compressão pode ser mais ou menos brusca, sendo determinada pelo parâmetro *knee*. Um *knee* suave irá partir do *ratio* 1:1 (não compressão) para o *ratio* selecionado de forma gradual, enquanto um *knee* agressivo irá fazer essa transição mais rapidamente. [1]

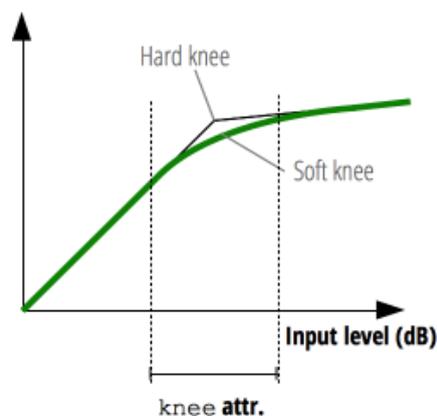


Figura 10: Representação da ação de um *soft knee* e de um *hard knee*.

O *knee*, no entanto, não diz respeito à velocidade de atuação do compressor (com a qual o compressor inicia e termina sua atuação), sendo esta dada respectivamente pela escolha do *attack* (tempo de ataque) e do *release* (tempo de liberação). Por atuarem no domínio do tempo, os parâmetros de *ataque* e *liberação* influenciam diretamente o envelope sonoro do sinal que está sendo processado. Por fim, o parâmetro *gain* (ganho) oferece a possibilidade de aumentar ou reduzir o nível de intensidade sonora do sinal no estágio de saída do processamento de compressão.

Com o advento do áudio digital e o surgimento de compressores digitais, os parâmetros se tornaram muito mais customizáveis, uma vez que os modelos analógicos tinham limitações de recursos e valores pré-estabelecidos para cada parâmetro. Apesar de possuírem menos opções de controle que os compressores digitais, alguns modelos de compressores analógicos são até hoje cultuados pelo seu desempenho e por suas características sonoras.[38] Existem quatro tipos de compressores analógicos: *valvulados*, *FET*, *ópticos* e *VCA*. [8]

Os primeiros modelos de compressores criados foram os *valvulados*, que atenuavam o sinal de saída através do uso de válvulas sensíveis ao sinal de entrada. O emprego de válvulas gera distorções harmônicas, mudando aspectos timbrísticos do som, o que faz dos compressores valvulados compressores ‘coloridos’, termo utilizado para se referir a dispositivos que adicionam intensidade em bandas de frequências do sinal de forma não linear durante o processamento[38]. Esse tipo de compressor não possui controle de *ratio* e tem valores de *ataque* e *liberação* maiores, sendo, portanto, mais lentos.[1]

O circuito dos compressores *ópticos* é formado por uma fonte luminosa atrelada ao sinal de entrada e uma foto-resistência. [8] A atenuação do sinal se dá por meio da interação entre a fonte luminosa e a foto resistência, ou seja, quanto maior o sinal de entrada no circuito, maior a intensidade da luz emitida pela fonte luminosa e, assim, maior a resistência do fotoreistor, reduzindo a intensidade do sinal.[38] Assim como os compressores valvulados, os compressores ópticos possuem tempo de resposta lento e produzem coloração ao sinal processado.[3]

Os compressores do tipo FET empregam transistores de efeito de campo que reduzem o sinal por meio do controle da sua resistência através da alteração da voltagem aplicada ao terminal *gate* [40]. Os compressores FET apresentam tempos de ataque e de liberação mais rápidos do que os valvulados e os ópticos e permitem o controle de *ratio*. [38]

Os compressores VCA são os mais recentes entre os compressores analógicos, tendo surgido na década de 1980.[8] Seu funcionamento ocorre através do controle de voltagem de um circuito de amplificação integrado. Se caracteriza por ter tempos de resposta mais rápidos, e por ser muito mais ‘transparente’, termo utilizado para ferramentas que não alteram as características em frequência do sinal (ou seja, não ‘colorido’).[38]

Além dos quatro tipos de compressores analógicos, existem, como mencionado, os compressores digitais que funcionam através de operações matemáticas, sendo altamente customizáveis e precisos[1]. Atualmente, além da vasta gama de compressores digitais disponíveis no mercado, os principais modelos de compressores analógicos foram emulados em instrumentos virtuais, trazendo toda a praticidade de uma ferramenta digital, mas mantendo as características sonoras que por tantos anos foram apreciadas.[8]

Com tantos tipos de ferramentas para compressão, cabe ao engenheiro de mixagem escolher a mais adequada de acordo com: (a) as características sonoras do elemento a ser processado; (b) a finalidade do tratamento de compressão, e (c) o contexto no qual o trabalho ou a atividade se insere. O uso dos compressores é bastante vasto e dentre as suas várias utilizações, podemos destacar as seguintes:

- *Acentuar detalhes sonoros*: comprimir o sinal para que algumas de suas nuances que porventura sejam pouco perceptíveis possam ficar mais evidentes (mais destacadas em intensidade).
- *Balancear volume*: diminuir a variação da amplitude dinâmica dos sinais de maneira a ter um maior controle do volume dos elementos durante a mixagem.
- Percepção de *loudness*: comprimir e expandir o ganho de um som com ataque forte e *sustain* fraco permite manter o nível de pico do sinal e também aumentar o nível médio durante o *sustain* o que proporciona uma sensação de aumento de *loudness*
- *Alterar o envelope dinâmico*: ao mudar o nível do sinal ao longo do tempo altera-se o envelope dinâmico de cada som, podendo enfatizar ou atenuar seu transiente de ataque, bem como o *sustain* (*sustentação*), o *decay* (decaimento) ou o *release* (*liberação/extinção*).
- *Adicionar frequências*: Como mencionado, alguns modelos analógicos de compressores possuem como característica a alteração do timbre do sinal processado o que pode ser desejado em alguns contextos.[1]

## Limiters

Uma outra ferramenta que altera a amplitude dinâmica de um sinal é o Limiter (limitador). Enquanto os compressores funcionam com uma taxa de compressão, permitindo níveis de saída superiores ao *threshold*, o *limiter* garante que o sinal não excederá o *threshold*. Em outras palavras podemos considerar o *limiter* como um compressor de *ratio*  $\infty:1$ . [3]

## Gates

O Gate é mais uma ferramenta que altera a dinâmica do sinal. Inicialmente inventado para diminuir ruído, o *gate* funciona através do estabelecimento de um nível mínimo de *threshold*, permitindo a saída apenas de um sinal que tenha intensidade superior ao nível limiar escolhido. [8]

## Expanders

O funcionamento do *expander* é oposto ao funcionamento do compressor. Ao invés de comprimir a amplitude dinâmica, ele aumenta. O *expander* também pode ser utilizado como *gate*, porém com mais parâmetros de customização.[1]

## Delay (*atraso*)

Segundo Savage, o *delay* pode ser definido como: ‘repetições individuais distintas de um som original devidas a atrasos decorrentes de reflexões da onda sonora produzidas pelo ambiente natural’.[3] Os dois principais parâmetros em um *delay* são o *tempo de atraso*, que determina o tempo entre o sinal original e a primeira repetição, e o *feedback*, que determina a quantidade de repetições subsequentes em decrescente em intensidade.[27]

De modo geral os *delays* podem ser classificados de acordo com o tempo de atraso: os *delays curtos*, com tempo de atraso entre 40-150 ms; os *delays médios*, entre 150-400 ms e os *delays longos*, acima de 400 ms. Outra forma de classificação diz respeito a como o *delay* se

comporta, sobretudo entre os canais esquerdo (L) e direito (R) em um sinal em formato estéreo com dois canais. O *delay* pode ser fixo e retornar em ambos os canais e também pode atuar de forma móvel na imagem sonora estéreo.[8]

As possibilidades e situações de aplicação são vastas quando se trata do uso de *delays*, dependendo do estilo musical e do gosto pessoal do *mixer* (engenheiro de mixagem).[27] De modo geral a funcionalidade do delay gira em torno da criação de uma ambientação acústica, adicionando sensação de espaço e profundidade; do posicionamento espacial um elemento de mixagem dentro da imagem sonora total; bem como de explorar ou preencher eventuais vazios, brechas ou silêncios.[1]

### *Reverb* (Reverberação)

Segundo Pujahari : ‘A reverberação é uma propriedade inerente a locais fechados, causada pela propagação e pela decorrente multiplicidade de reflexões (de modo difuso) dentro de um espaço estanque.’[20] Os primeiros usos de reverberação artificial para fins musicais tinham por objetivo adicionar realismo às gravações feitas em estúdio. Porém, ao longo do tempo, o *reverb* se tornou também um recurso criativo capaz de simular respostas de salas existentes no mundo físico e também de salas imaginárias.[27]

Existem cinco principais tipos de simulação de reverberação: *hall*, *sala*, *camara*, *placa* (*plate*) e não-linear. Os reverbs de hall, sala e câmara são os tipos naturais de reverb, que simulam salas acústicas. Já os reverbs dos tipos *plate* e *não-linear* são artificiais. O tipo *plate* é obtido através da vibração de uma placa metálica suspensa tendo nas extremidades transdutores elétricos de sinal sonoro, induzindo na placa o sinal direto e captando o retorno difuso. Já o *reverb* não-linear é um produto da era digital, permitindo uma customização completa dos parâmetros e a simulação de respostas acústicas de salas que talvez jamais poderiam ser construídas.[8]

Dada a quantidade numerosa de *plug-ins* disponíveis no mercado, os parâmetros de controle de reverbs podem ser os mais variados possíveis. Os mais comuns, no entanto, são: *quantidade de som direto* em relação à *quantidade de som reverberante*, pré-delay, tempo de decaimento (*tempo de reverberação* –  $T_{60}$ ), tamanho da sala (*size*), densidade, difusão e equalização da reverberação.

O *som direto* é o próprio sinal original, de entrada, antes de ser reverberado. O *pre-delay* se refere ao tempo entre o sinal direto e a primeira reflexão. O tempo de decaimento é o tempo de reverberação, que é o tempo necessário para que as reflexões diminuam de 60 dB de intensidade em relação ao som direto. O *tamanho da sala* (size) diz respeito justamente ao espaço físico que está sendo simulado (ao volume da sala – em m<sup>3</sup>). Os parâmetros de *densidade* e *difusão* dizem respeito ao comportamento das reflexões em campo direto e em campo reverberante, e a equalização da reverberação uma simulação da absorção em função da frequência das ondas sonoras por parte dos materiais de revestimento da sala simulada.[20]

A funcionalidade do *reverb* em uma mixagem pode ter várias nuances, desde adequar o som à estética de um gênero musical específico, à criação de um espaço sonoro, até resolver problemas técnicos que podem surgir durante uma mixagem. Dentre as principais aplicações do *reverb* podemos citar: simular profundidade espacial, dar mais realismo à escuta de uma performance musical, adequar o espaço de escuta para uma determinada música ou produção sonora; mudar o timbre de um elemento; criar um efeito de dramatização; salientar ou destacar um elemento; posicionar um elemento na imagem sonora estereofônica; ou adicionar coesão à um elemento dentro do contexto da mixagem.[1]

## Distorção

Vista muitas vezes como indesejável no processo de gravação, a distorção é uma ferramenta de muito valor na mixagem, permitindo a adição de conteúdo **em timbre** a um som enquanto mantém as características do elemento sonoro sendo mixado.[3] A distorção era uma característica muito presente na era analógica do áudio, apesar de tecnicamente seu controle em produção ter sempre sido uma questão problemática. Porém quando aplicada de maneira coerente e dosada na medida certa, a *distorção* imprimia características timbrísticas expressivas e interessantes, sendo inclusive classificada em vários tipos (*overdrive*, *valve saturation*, *fuzz* e *bit crushing*), o que fez com que seu uso fosse assimilado e mantido ulteriormente na era digital.[27]

## Exciters

Excitador é um dispositivo que usa manipulação de fase e distorção harmônica para produzir mudanças no timbre em um sinal [8]. O primeiro *excitador* foi produzido pela companhia Aphex em 1975 por acidente. Desde então, muitos outros *exciters* foram lançados. Apesar do funcionamento de cada *exciter* ser mantido em segredo pelos fabricantes, os princípios giram em torno da adição de harmônicos, da manipulação da dinâmica e da manipulação de fase, proporcionando além das alterações do conteúdo sonoro uma sensação de maior *loudness* para os ouvintes.[1]

## Transient Shaper

Modeladores de transiente são ferramentas que atuam apenas nos transientes do sinal podendo contê-los ou acentuá-los.[1] Alguns modelos de *designers* de transiente requerem um valor de *threshold*, assumindo que os transientes costumam ser mais intensos (fortes) do que o resto do sinal. No entanto há modelos de *transient shapers* que não utilizam um controle de nível de *threshold* mas possuem recursos capazes de reconhecer mudanças abruptas de nível sonoro, o que é uma característica própria aos transientes e, através disso, podem acionar seu funcionamento, como se estivessem se adaptando dinamicamente a diferentes níveis de *threshold*. [27]

## 3. Materiais e métodos

### 3.1 Ensaio

De modo a explorar as possibilidades e a pertinência das ferramentas de mixagem que utilizam inteligência artificial, o presente trabalho realizou uma série de cinco ensaios onde a mesma peça musical foi processada tanto por ferramentas automáticas quanto por meios mecânicos.

Inicialmente foi realizado um levantamento de ferramentas que utilizam inteligência artificial disponíveis no mercado. Foram encontradas soluções fabricadas por quatro

fabricantes diferentes que estão listadas na Tabela 1 juntamente com uma breve descrição de seu objetivo principal e do seu funcionamento. [41][42][43][44][45][46]

**Tabela 1: Ferramentas de mixagem automáticas**

Nome	Objetivo	Fabricante	Funcionamento
Soothe 2	Equalizador/ Compressor	Oeksound	Identifica de forma automática e em tempo real frequências ressonantes e componentes espectrais indesejados. Permite ao usuário que controle os parâmetros de supressão.
Gullfoss	Equalizador	Soundtheory	Processa de forma automática e em tempo real o sinal de entrada e aplica duas curvas de equalização: uma que adiciona frequências e outra que retira frequências. Permite ao usuário que regule a intensidade de cada curva.
smart:EQ3	Equalizador	Sonible	Analisa o sinal de entrada por um tempo e em seguida o equaliza de forma automática e em tempo real. Também pode ser usado como um equalizador comum, sendo a automação opcional.
smart:comp	Compressor	Sonible	Analisa o sinal de entrada por um tempo e em seguida o comprime de forma automática e em tempo real. Também pode ser usado como um compressor comum, sendo a automação opcional.
smart:reverb	Reverb	Sonible	Analisa o sinal de entrada por um tempo e em seguida aplica nele reverb de forma

			automática e em tempo real. Também pode ser usado como um reverb comum, sendo a automação opcional.
Neutron Sculptor	Equalizador	IZotope	Após a seleção do instrumento a ser processado e da finalidade, equaliza o sinal. Permite ajuste de parâmetros após a análise.
Neutron Relay	Fader digital	IZotope	Analisa cada elemento presente na música e automaticamente ajusta o nível de intensidade sonora para cada canal.
Neutron Mix Assistant	Equalizador, compressor, exciter, gate e transient shaper.	IZotope	Analisa cada elemento presente na música e realiza um processamento completo de cada uma delas com suas ferramentas. Cada ferramenta também pode ser adicionada e utilizada de forma mecânica.

Foram então elaborados cinco ensaios para explorar todas as funcionalidades encontradas nas ferramentas automáticas. Visando que tanto na execução mecânica quanto na coleta automática as mesmas funcionalidades estariam à disposição, na mixagem mecânica foram utilizadas as mesmas ferramentas quando estas dispunham de um modo mecânico de operação. Nos casos cuja ferramenta não permite manipulação mecânica uma outra ferramenta sem automação foi empregada. Na Tabela 2 é possível observar uma descrição de cada ensaio, bem como as ferramentas empregadas e na Tabela 3 estão expostas as visões formuladas para a execução de cada mixagem mecânica.

**Tabela 2: Descrição dos ensaios realizados**

Ensaio	Objetivo	Ferramentas empregadas na mixagem automática	Ferramentas empregadas na mixagem mecânica
1	Detecção e atenuação de frequências ressonantes indesejadas em uma faixa de guitarra.	Soothe 2	Equalizador
2	Equalizar e balancear o volume de uma composição de cinco elementos: guitarra, baixo, bumbo, caixa e prato.	Neutron Sculptor Neutron Relay	Equalizador Fader de volume
3	Equalizar, comprimir, adição de reverb e balanço de volumes de uma composição de cinco elementos: guitarra, baixo, bumbo, caixa e prato.	smart:EQ3 smart:comp smart:reverb Neutron Relay	smart:EQ3 (no modo mecânico) smart:comp (no modo mecânico) smart:reverb (no modo mecânico) Fader de volume
4	Equalização do mix bus realizada em cima do resultado da coleta 3.	Gullfoss	Equalizador
5	Realização de uma mixagem completamente automática, incluindo a seleção de ferramentas.	Neutron Mix Assistant Neutron Relay	Neutron Mix Assistant (no modo mecânico) Fader de volume

**Tabela 3: Visões formuladas para a execução da etapa mecânica de cada ensaio**

Ensaio	Visão
1	Atenuar de forma sutil as ressonâncias indesejadas presentes em uma faixa de guitarra, preservando a timbragem original da guitarra.
2	Destacar a guitarra em termos de <i>loudness</i> em relação aos outros elementos. Aumentar a percepção do prato e da caixa. Desenhar as curvas de equalização do baixo e do bumbo de modo a evitar mascaramento.
3	Destacar a guitarra em termos de <i>loudness</i> em relação aos outros elementos. Desenhar as curvas de equalização do baixo e do bumbo de modo a evitar mascaramento. Aumentar a percepção do prato e da caixa. Ambientar a guitarra, o prato e a caixa. Reter sutilmente os transientes da guitarra e do baixo. Enfatizar os transientes da caixa e do bumbo.
4	Atenuar frequências graves e subgraves de modo a conter <i>muddiness</i> e <i>boominess</i> (ver Figura 3). Acrescentar frequências acima de 10 kHz de modo a acrescentar <i>airness</i> (ver Figura 3).
5	Destacar a guitarra em termos de <i>loudness</i> em relação aos outros elementos. . Desenhar as curvas de equalização do baixo e do bumbo de modo a evitar mascaramento. Aumentar a percepção do prato e da caixa. Balancear os transientes dos elementos.

O primeiro ensaio foi realizado tendo como perspectiva a solução de um problema muitas vezes encontrado ao longo de uma mixagem que é o de conter frequências ressonantes indesejadas. Segundo Owsinski: ‘frequência ressonante é uma frequência particular ou faixa de frequências que são acentuadas, geralmente devido a algum fator acústico, eletrônico ou mecânico estranho.’[8]. A etapa mecânica foi executada através do uso de um equalizador paramétrico comum e o método empregado foi o de realizar uma varredura buscando atenuar

as frequências indesejadas. Já a etapa automática consistiu na utilização de uma configuração pré estabelecida pelo Soothe 2 com o objetivo de remover ressonâncias em guitarras.

O segundo ensaio buscou verificar duas ferramentas essenciais ao processo de mixagem: o *fader digital* e o equalizador. A etapa mecânica iniciou-se pela aplicação de equalização em todas as faixas seguindo o que havia sido previamente estabelecido na visão, iniciando a equalização pela guitarra e pelo baixo e depois seguindo para a bateria. Após a equalização, os volumes foram estabelecidos nos faders. A etapa automática baseou-se inicialmente na inserção do plug in *Neutron Sculptor* em cada canal. Em seguida foi informada a cada unidade do plug in o instrumento correspondente que seria equalizado e foi aplicada uma curva de equalização de forma automática. Após a equalização foi utilizado o plugin *Neutron Relay* para definição automática dos volumes. O *Neutron Relay* permite a escolha de instrumentos a serem priorizados em volume, neste caso a guitarra foi selecionada como prioridade.

O terceiro ensaio orientou-se para o estudo de ferramentas que possibilitam a opção de serem operadas tanto de forma totalmente automática, quanto de forma totalmente mecânica. Neste ensaio as mesmas ferramentas foram utilizadas na mesma ordem de processamento para cada uma das etapas. Na etapa mecânica foi inserido primeiramente o plug in *sonible:smartEQ3* em todas as faixas, em seguida foi inserido o plugin *sonible:smartcomp* nas faixas de guitarra, baixo, bumbo e caixa. Depois foi inserido o plugin *sonible:smartReverb* na guitarra, na caixa e no prato, e por fim os níveis em volume foram definidos. A etapa automática foi realizada com a definição de parâmetros de forma automática e com os níveis em volume sendo definido pelo plugin *Neutron Relay*.

O quarto ensaio foi realizado para investigar o uso de **equalizadores** no *mix bus*, ou seja, um processamento efetuado na faixa musical como um todo. Para este ensaio foram utilizadas as mixagens feitas **no ensaio três**. Na etapa mecânica foi utilizado um equalizador paramétrico seguindo os objetivos traçados pela visão. Na etapa automática foi empregado o plug in *Gullfoss*, e seus parâmetros para as curvas aditiva e subtrativa em frequência foram definidos na metade de sua capacidade cada.

O quinto ensaio foi desenhado para avaliar um cenário mais próximo de uma **automação completa** do processo de mixagem, pois a ferramenta *Neutron Mix Assistant*, embora permita o posterior ajuste mecânico, possui um modo que seleciona tanto as ferramentas, quanto a ordem de processamento e cada parâmetro de forma automática. Na

etapa mecânica foram selecionadas dentre as ferramentas disponíveis no Neutron as que foram julgadas como pertinentes para o alcance da visão. As ferramentas utilizadas no processamento da guitarra foram, em ordem de processamento: primeiro *compressor*, *equalizador*, *segundo compressor* e *exciter*. As ferramentas utilizadas no processamento do baixo foram, em ordem de processamento: *equalizador*, *compressor*, *exciter* e *transient shaper*. As ferramentas utilizadas no processamento do bumbo foram, em ordem de processamento: *equalizador*, *exciter*, *compressor* e *transient shaper*. As ferramentas utilizadas no processamento da caixa foram, em ordem de processamento: *equalizador*, *compressor* e *exciter*. As ferramentas utilizadas no processamento do prato foram, em ordem de processamento: *exciter* e *equalizador*. Na etapa automática o *Neutron Mix Assistant* realizou a mixagem de forma automática e o *Neutron Relay* definiu os volumes também de forma automática. As ferramentas selecionadas pelo algoritmo estão descritas na seção de resultados e análise.

### **3.2 Análise perceptual**

Após obtidos os resultados *automático* e *mecânico* de cada ensaio, foi conduzida uma pesquisa comparando os resultados obtidos no segundo, terceiro e quinto ensaios. As faixas foram previamente equiparadas em termos de *loudness* e a pesquisa foi elaborada em quatro partes. A primeira parte dispunha dos resultados da mixagem mecânica e da mixagem automática oriundos do segundo ensaio. As faixas não possuíam nenhum tipo de identificação, apenas estavam disponíveis para audição quantas vezes o respondente desejasse. Em seguida à audição, foi solicitado ao respondente que escolhesse qual das duas versões da música foi sua preferida e caso não houvesse preferência, havia possibilidade de escolher a alternativa ‘sem preferência’. O mesmo procedimento foi repetido para os resultados emergentes do terceiro e do quinto ensaio na segunda e terceira parte. Na quarta e última parte da pesquisa, foi solicitado ao respondente que selecionasse seu grau de experiência com produção musical e/ou engenharia de som, sendo que havia três graus de expertise: (a) nenhuma experiência, (b) alguma experiência e (c) profissional da área.

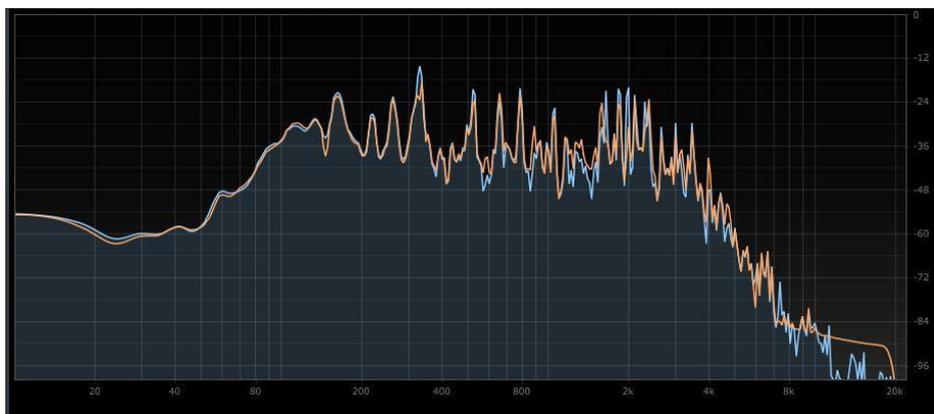
## 4. Resultados e análise

### 4.1 Análise descritiva

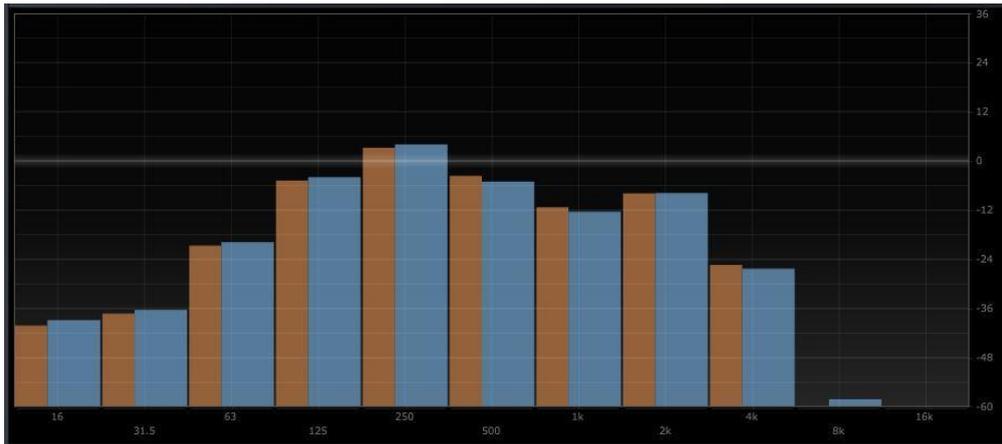
Após a realização dos ensaios, cada produto emergente foi exportado em um arquivo estéreo no formato .wav e então, as mixagens *automática* e *mecânica* para cada experimento foram equiparadas em *loudness* utilizando um medidor de *loudness*.

#### Primeiro Ensaio

No primeiro ensaio, foi feita uma análise em frequência, uma vez que o objetivo era a contenção de ressonâncias. Na faixa dos médio-graves (250 Hz - 2 kHz) tanto o processamento mecânico quanto o processamento automático obtiveram respostas similares, o produto final da mixagem automática no entanto obteve maior sucesso em atenuar uma ressonância indesejável em torno de 1,8 kHz. Na faixa dos médio-agudos (2 kHz - 4 kHz) ambos processamentos obtiveram resultados parecidos, com uma maior atenuação por parte do processamento automático em 2,8 kHz. Na faixa dos agudos (acima de 4 kHz) houve uma diferença significativa de processamento na faixa acima de 10 kHz, sendo que o produto final da mixagem automática apresentou uma atenuação de forma contundente de todas as frequências nesta faixa enquanto o processamento mecânico manteve essa faixa de frequência.



**Figura 11: Comparação entre os componentes espectrais da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

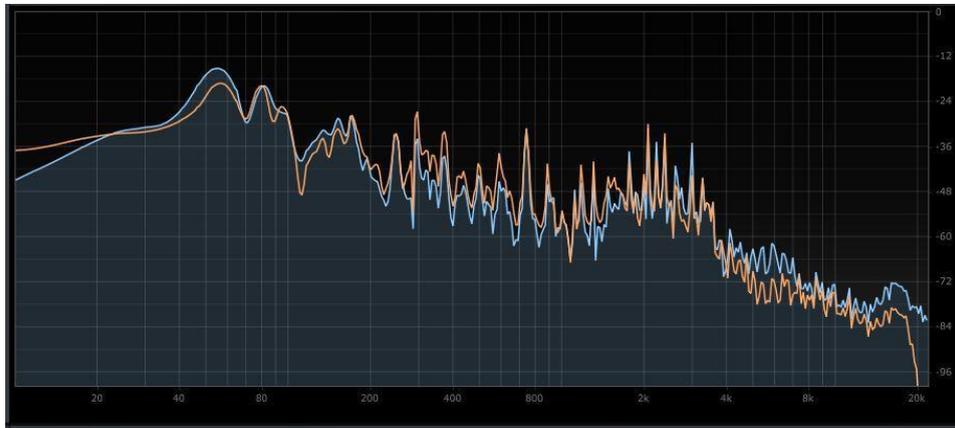


**Figura 12: Comparação entre os componentes espectrais, por oitava, da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

### Segundo Ensaio

No segundo ensaio, foi feita uma análise em frequência e em volume, já que as ferramentas empregadas eram apenas um *equalizador* e um *fader* de volume. Quanto aos volumes relativos entre os componentes, após equiparação de *loudness*, o volume escolhido para o baixo foi similar em ambas as mixagens, enquanto os demais volumes relativos divergiram. A mixagem automática apresentou maior nível de intensidade sonora para a guitarra, enquanto os elementos da bateria (bumbo, prato e caixa) ficaram mais intensos na mixagem mecânica. Em termos de frequência, é possível observar que houve discrepâncias. Na faixa do sub-grave é possível observar que a fundamental do bumbo está mais intensa na mixagem mecânica. Na faixa dos graves, onde está localizada a frequência fundamental do baixo, existe uma similaridade de energia em ambas as mixagens. Na faixa das médias-graves, região muito importante para a sonoridade de elementos como guitarra e caixa e onde está presente algumas harmônicas do baixo, houve discrepâncias. Na mixagem automática houve dois problemas nesta região de frequência: houve o mascaramento de elementos, já que a guitarra está encobrindo tanto o baixo quanto a caixa, e houve aumento de frequências ressonantes na guitarra. Na mixagem mecânica não houve mascaramento, uma vez que os elementos estão distinguíveis entre si, e também houve maior controle de ressonâncias, porém isso ocorreu ao custo de uma redução na energia da guitarra nesta região. Na região das médias-agudas ficou evidente que os componentes da bateria tiveram maior

energia na mixagem mecânica enquanto na mixagem automática houve maior ênfase na guitarra. Na região das frequências agudas as componentes do prato e da caixa ficaram em maior evidência na mixagem mecânica em comparação com a automática.

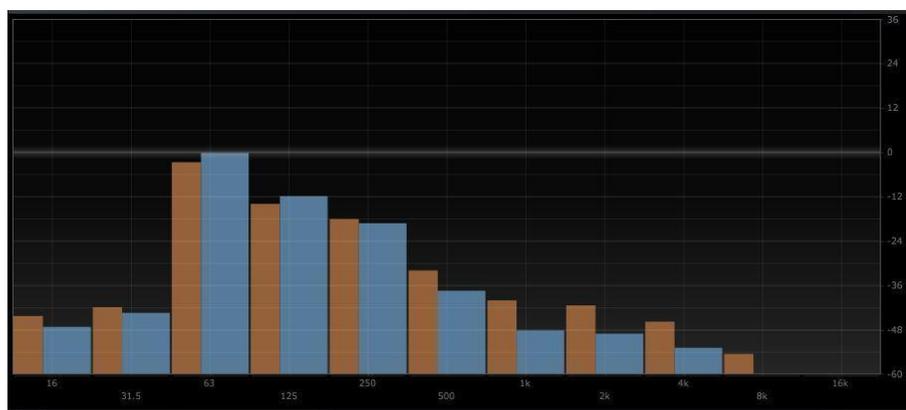


**Figura 13: Comparação entre os componentes espectrais da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

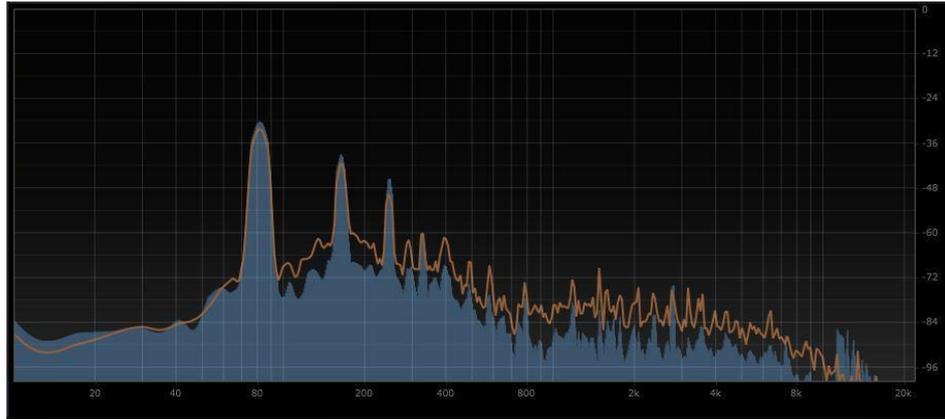
### Terceiro Ensaio

Sobre os resultados obtidos no terceiro ensaio, foi feita uma análise em termos de volume, frequência, dinâmica, campo estéreo e compatibilidade mono. Sobre os níveis de intensidade relativos, após equiparação de *loudness*, foi perceptível que a guitarra ficou mais **intensa (forte)** na mixagem automática, enquanto bumbo e baixo ficaram mais **intensos (fortes)** na mixagem mecânica. Prato e caixa obtiveram níveis de intensidade sonora similares. Na faixa dos subgraves, a mixagem mecânica obteve maior energia na frequência fundamental do bumbo em relação à mixagem automática. Na faixa das frequências graves, a fundamental do baixo possuiu maior energia na mixagem mecânica em relação à mixagem automática. Na mixagem automática houve uma maior percepção de outros elementos como a guitarra e a caixa nesta faixa de frequência, o que resultou em um leve mascaramento de componentes importantes do bumbo. Na faixa das médias-agudas houve maior energia na mixagem automática em relação à mixagem mecânica. Essa maior energia realçou algumas frequências ressonantes da guitarra. Nas faixas das médias-agudas e das agudas houve novamente uma ênfase maior na guitarra, tornando-a mais presente na mixagem automática em relação à mixagem mecânica. Quanto à dinâmica, a mixagem automática obteve, de

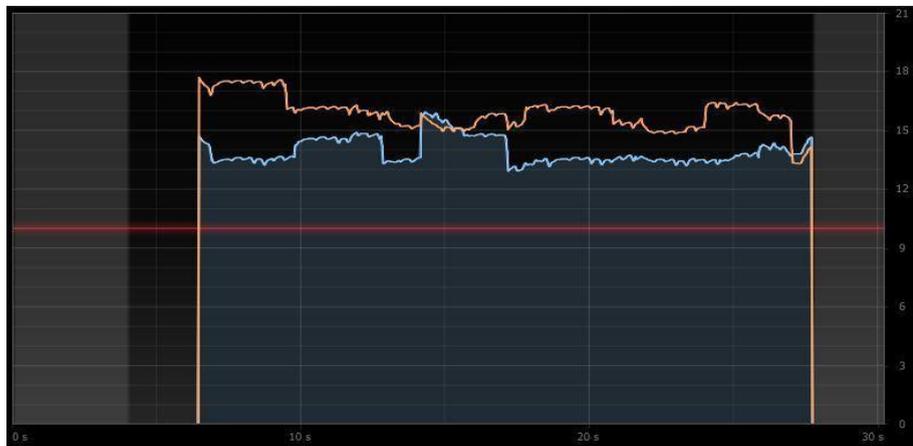
maneira geral, um maior âmbito dinâmico em comparação com a mixagem mecânica, o que significa que a mixagem mecânica foi comprimida com parâmetros mais contundentes. Neste ensaio o uso do *reverb* se concentrou em três elementos: guitarra, prato e caixa. O *reverb* na guitarra foi estabelecido com parâmetros distintos em cada mixagem, enquanto na mixagem automática o tempo de decaimento foi de 1,5 s e o tempo de *pré-delay* foi de 0 s, na mixagem mecânica o tempo de decaimento foi de 0,77 s e o tempo de *pré delay* foi de 19ms. Isso resultou numa diferença sonora significativa da guitarra em ambas as mixagens. No *reverb* do prato houve convergência no tempo de *pré-delay*, que foi de 0 s em ambas as mixagens, porém houve divergência no tempo de decaimento que na mixagem automática foi de 1,50 s e na mixagem mecânica foi de 0,46 s. No *reverb* da caixa houve convergência no tempo de *pré-delay*, que foi de 0 s em ambas as mixagens, porém houve divergência no tempo de decaimento que na mixagem automática foi de 1,50 s e na mixagem mecânica foi de 0,52 s. Essa diferença nos tempos de decaimento do prato e da caixa alteraram de forma significativa o resultado de cada mixagem, uma vez que na mixagem mecânica esses elementos soaram como se estivessem em um ambiente menor em tamanho enquanto na mixagem automática os elementos soaram como se estivessem em um ambiente maior. Ultimamente, **ouvindo ambos os resultados em mono**, é possível observar que a mixagem automática conseguiu traduzir o som da guitarra de forma bastante satisfatória em mono, porém apresentou **problemas de fase**, uma vez que o som do bumbo ficou comprometido em mono. Na mixagem mecânica, por outro lado, todos os elementos conseguiram traduzir de forma satisfatória quando somados os canais direito e esquerdo.



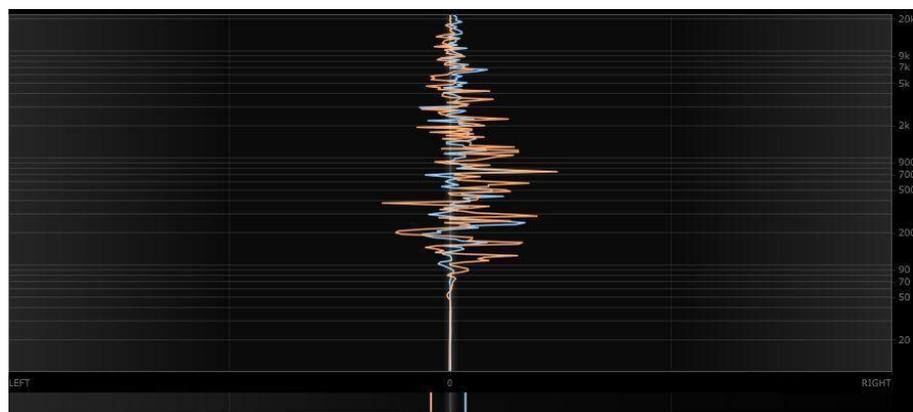
**Figura 14: Comparação entre os componentes espectrais, por oitava, da mixagem automática (em laranja) e da mixagem mecânica (em azul).**



**Figura 15: Comparação entre os componentes espectrais da mixagem automática (em laranja) e da mixagem mecânica (em azul).**



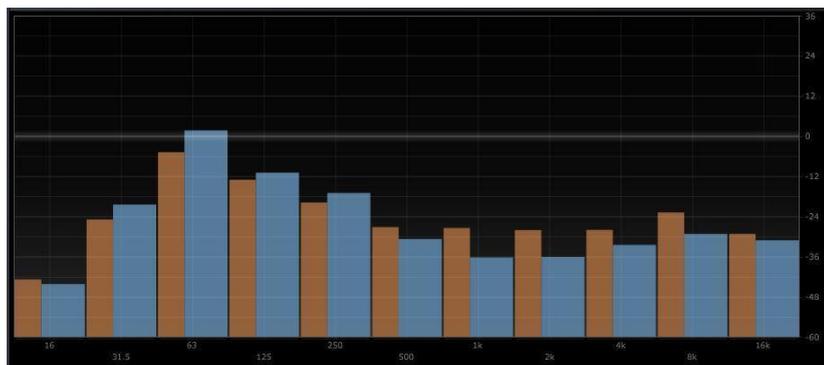
**Figura 16: Comparação entre o alcance dinâmico ao longo do tempo da mixagem automática (em laranja) e da mixagem mecânica (em azul).**



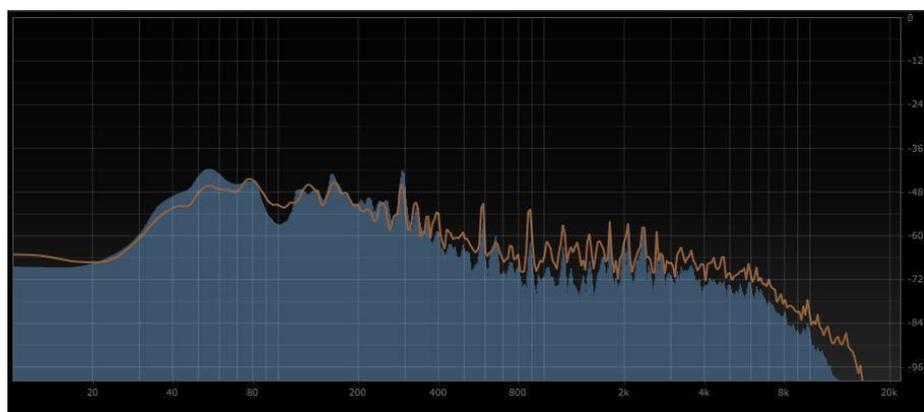
**Figura 17: Comparação entre o campo estéreo da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

## Quarto Ensaio

O ensaio de número quatro teve o propósito de investigar a aplicação da equalização no *mix bus*, utilizando os resultados obtidos no terceiro ensaio, portanto a análise dos resultados foi feita em frequência apenas. A aplicação da equalização automática no *mix bus* foi capaz de resolver um problema aparente no resultado do terceiro ensaio que foi a emergência de frequências ressonantes, porém causou um novo problema ao aumentar de forma demasiada as frequências altas, o que comprometeu o *balanço tonal* da música. A aplicação de equalização de forma mecânica foi capaz de aumentar o detalhamento dos elementos e trazer maior balanço tonal para a mixagem, no entanto a guitarra, que é o elemento principal nesta composição, não conseguiu manter a mesma energia ao longo da mixagem, sendo ofuscada em alguns momentos por outros elementos como o baixo.



**Figura 18: Comparação entre os componentes espectrais, por oitava, da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

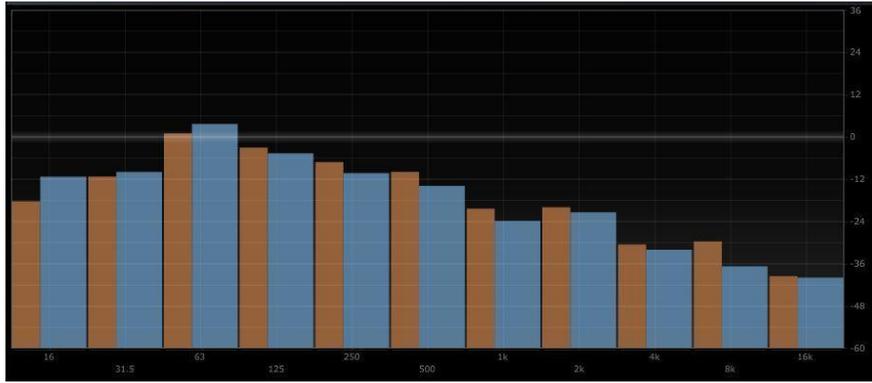


**Figura 19: Comparação entre os componentes espectrais da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

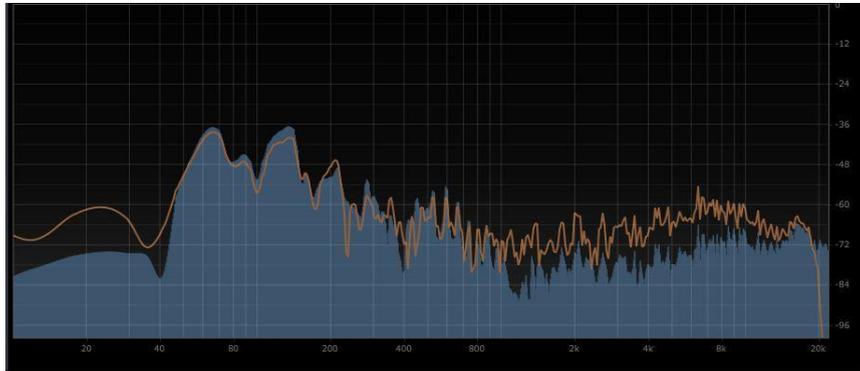
## Quinto Ensaio

No quinto ensaio os resultados foram analisados em termos de dinâmica, frequência e nível de intensidade entre elementos. Em termos de *nível de intensidade relativo*, o baixo e o bumbo tiveram níveis semelhantes, enquanto a guitarra, o prato e a caixa ficaram mais fortes na mixagem automática. Nas faixas dos subgraves e graves, a energia em ambas as mixagens foi equivalente, com a fundamental do bumbo ligeiramente mais alta na mixagem mecânica. Nas faixas dos médio agudos e dos médio graves, houve uma maior ênfase no baixo na mixagem mecânica,  **mascarando componentes da guitarra**, enquanto na mixagem automática houve maior ênfase na guitarra,  **mascarando componentes do baixo**. Nas frequências agudas houve maior energia na mixagem automática em relação à mixagem mecânica. Em termos de dinâmica, a mixagem mecânica apresentou menor âmbito dinâmico em relação à mixagem automática.

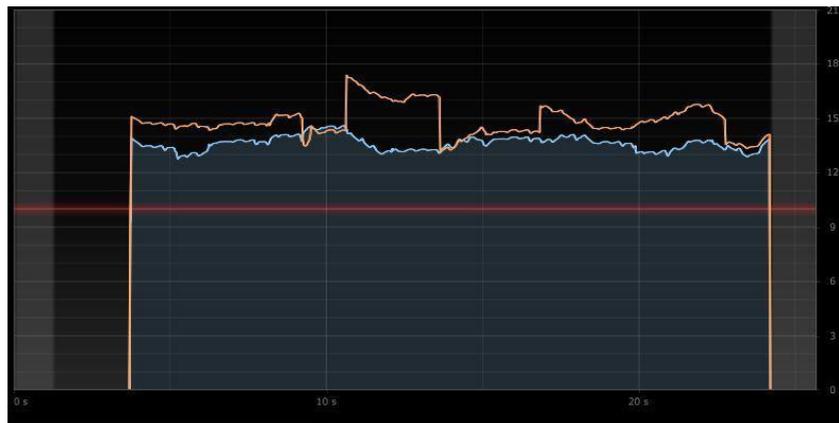
Diferentemente dos demais ensaios, no quinto ensaio a seleção de quais as ferramentas utilizadas na mixagem automática, bem como a ordem em que foram empregadas, foram definidas de forma automática pelo algoritmo do *plugin* empregado (Neutron Mix Assistant). Houve divergência entre as ferramentas selecionadas de forma mecânica e as ferramentas selecionadas de forma automática. As ferramentas utilizadas no processamento automático da guitarra foram, em ordem de processamento: equalizador, compressor e exciter. As ferramentas utilizadas no processamento automático do baixo foram, em ordem de processamento: equalizador, exciter e compressor. As ferramentas utilizadas no processamento automático do bumbo foram, em ordem de processamento: *equalizador, exciter e compressor*. As ferramentas utilizadas no processamento automático da caixa foram, em ordem de processamento: *equalizador, exciter e compressor*. As ferramentas utilizadas no processamento automático do prato foram, em ordem de processamento: primeiro *equalizador*, segundo *equalizador, exciter e compressor*.



**Figura 20: Comparação entre os componentes espectrais, por oitava, da mixagem automática (em laranja) e da mixagem mecânica (em azul).**



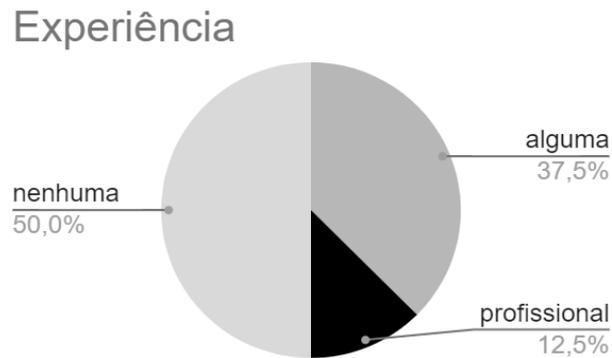
**Figura 21: Comparação entre os componentes espectrais da mixagem automática (em laranja) e da mixagem mecânica (em azul).**



**Figura 22: Comparação entre o alcance dinâmico ao longo do tempo da mixagem automática (em laranja) e da mixagem mecânica (em azul).**

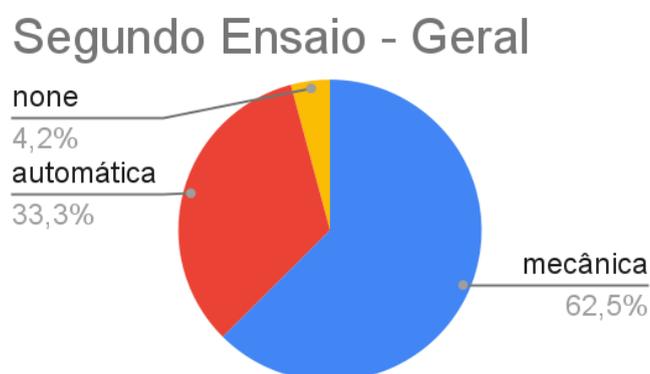
## 4.2 Análise perceptiva

Foram obtidas 24 respostas para o questionário de análise perceptual sendo que 50% dos respondentes não tinham experiência prévia com produção ou mixagem, 37,5% possuíam alguma experiência e 12,5% eram profissionais da área.



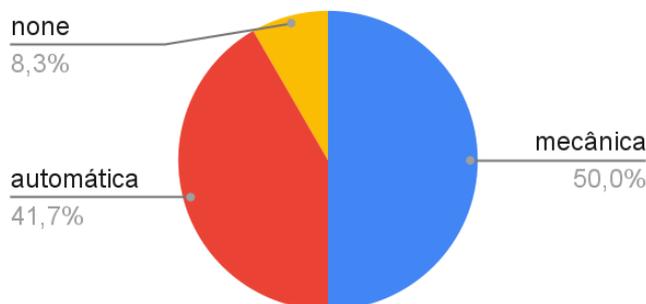
**Gráfico 1: Grau de experiência dos respondentes.**

Os resultados extraídos a partir da análise das mixagens obtidas no **segundo ensaio** estão expressados no Gráfico 2. Também estão descritas as preferências divididas entre os respondentes com nenhuma experiência e os respondentes com alguma experiência ou profissionais respectivamente nos Gráficos 3 e 4. É possível observar maior preferência geral pela mixagem mecânica, sendo que entre os indivíduos com alguma experiência ou profissionais, 3 em cada 4 escolheram a mixagem mecânica.



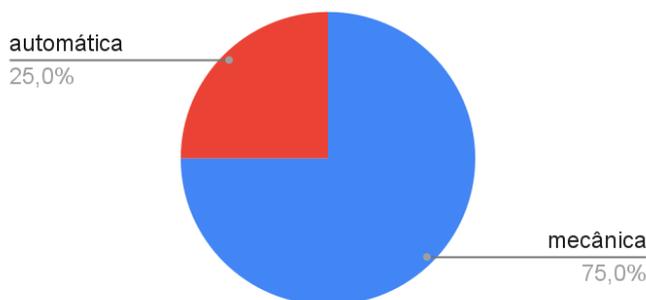
**Gráfico 2: Distribuição das preferências dos respondentes para o segundo ensaio.**

### Segundo Ensaio - S/Experiência



**Gráfico 3: Distribuição das preferências dos respondentes sem experiência na área para o segundo ensaio.**

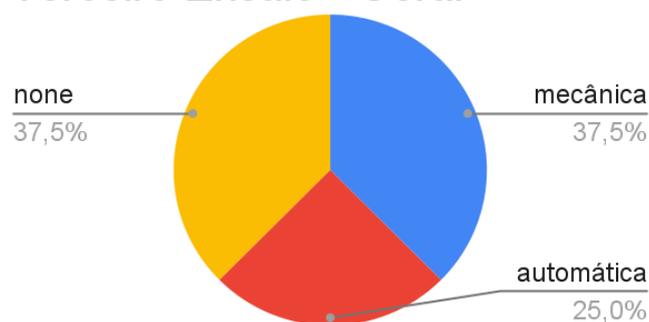
### Segundo Ensaio - C/Experiência



**Gráfico 4: Distribuição das preferências dos respondentes com alguma experiência e profissionais na área para o segundo ensaio.**

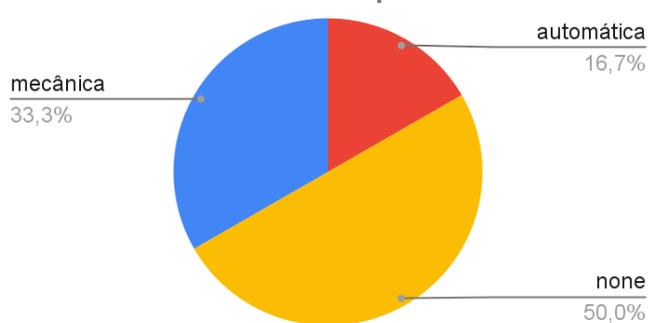
As respostas obtidas sobre as mixagens do **terceiro ensaio** estão expressas no Gráfico 5. Também estão descritas as preferências divididas entre os respondentes com nenhuma experiência e os respondentes com alguma experiência ou profissionais respectivamente nos Gráficos 6 e 7. É possível observar que de maneira geral um quarto dos respondentes preferiu a mixagem feita de forma automática, enquanto três oitavos preferiram a mixagem mecânica e outros três oitavos selecionaram a opção 'sem preferência'. Entre os indivíduos sem experiência houve uma predominância pela indefinição, enquanto entre os indivíduos com alguma experiência ou profissionais a indefinição foi menor e a preferência pela mixagem mecânica foi maior.

### Terceiro Ensaio - Geral



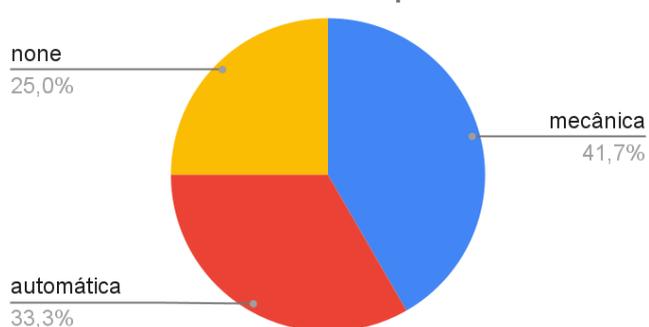
**Gráfico 5: Distribuição das preferências dos respondentes para o terceiro ensaio.**

### Terceiro Ensaio - S/Experiência



**Gráfico 6: Distribuição das preferências dos respondentes sem experiência na área para o terceiro ensaio.**

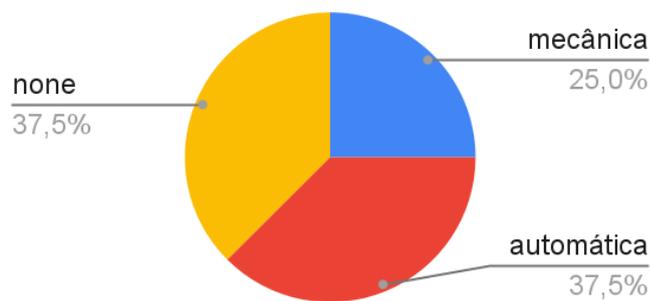
### Terceiro Ensaio - C/Experiência



**Gráfico 7: Distribuição das preferências dos respondentes com alguma experiência e profissionais na área para o terceiro ensaio.**

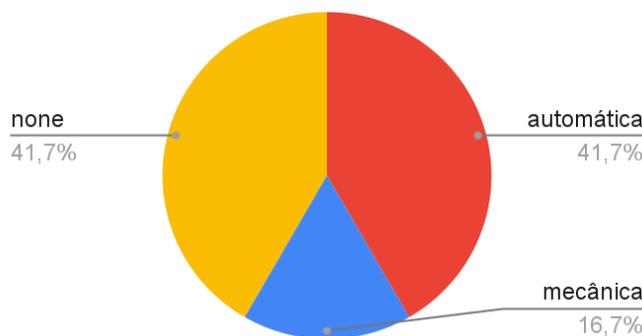
As respostas obtidas sobre as mixagens do **quinto ensaio** estão expressas no Gráfico 8. Também estão descritas as preferências divididas entre os respondentes com nenhuma experiência e os respondentes com alguma experiência ou profissionais respectivamente nos Gráficos 9 e 10. É possível observar que de maneira geral um quarto dos respondentes preferiu a mixagem feita de forma mecânica, enquanto três oitavos preferiram a mixagem automática e outros três oitavos selecionaram a opção ‘sem preferência’. Entre os indivíduos com alguma experiência ou profissionais as três opções foram igualmente distribuídas.

Quinto Ensaio - Geral



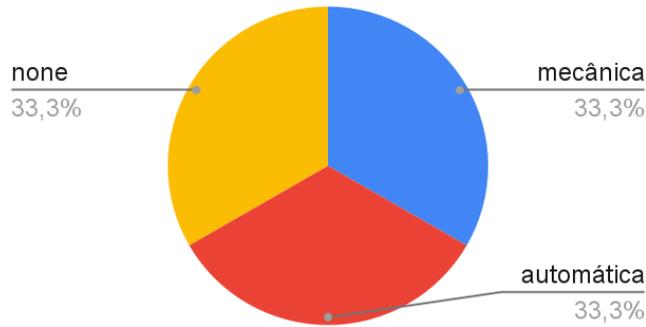
**Gráfico 8: Distribuição das preferências dos respondentes para o quinto ensaio.**

Quinto Ensaio - S/Experiência



**Gráfico 9: Distribuição das preferências dos respondentes sem experiência na área para o quinto ensaio.**

### Quinto Ensaio - C/Experiência



**Gráfico 10: Distribuição das preferências dos respondentes com alguma experiência e profissionais na área para o quinto ensaio.**

## 5. Conclusão

O desenvolvimento do presente trabalho permitiu uma maior compreensão dos aspectos históricos, técnicos e artísticos da engenharia de mixagem, bem como possibilitou uma investigação sobre as possibilidades que os avanços em Inteligência Artificial trouxeram para esta área. Isso foi possível através do estabelecimento inicial de um panorama histórico e de uma revisão dos principais estudos conduzidos acerca de Sistemas Inteligentes de Mixagem (IMS), seguido de uma definição dos principais conceitos e ferramentas aplicadas no âmbito da mixagem de som. Posteriormente foram levantadas as principais ferramentas disponíveis comercialmente que contam com formas automáticas de realizar o processo de mixagem e então foram elaborados cinco ensaios de modo a testar tais ferramentas. Nestes ensaios foram realizadas tanto mixagens automáticas, quanto mixagens mecânicas utilizando uma quantidade limitada de processamento em ambas as formas, de modo a garantir as mesmas funcionalidades. Então foi realizada uma análise descritiva para os resultados obtidos por meios artificiais e humanos em cada ensaio, onde foi possível observar as principais diferenças entre cada resultado emergente.

Finalmente foi conduzida uma pesquisa de percepção com uma amostra de respondentes variando entre pessoas sem nenhuma experiência prévia com mixagem e áreas correlatas até profissionais da área. Nos resultados da pesquisa **não foi constatada uma preferência predominante** por uma forma de mixagem em detrimento de outra, o que foi **observado foi uma alternância entre as preferências** de acordo com os resultados dos

ensaios. Dos três ensaios que foram levados para a pesquisa de percepção em um deles houve uma preferência maior pela mixagem mecânica, em outro houve um empate entre os respondentes que preferiram a mixagem mecânica e os que não tiveram preferência e em outro houve um empate entre os respondentes que preferiram a mixagem automática e os que não tiveram preferência. Entre os respondentes sem experiência e os respondentes com alguma experiência ou profissionais houve algumas diferenças perceptuais, sendo que as pessoas sem experiência escolheram com mais frequência a opção ‘sem preferência’ e as pessoas com alguma experiência na área ou profissionais preferiram com maior frequência a mixagem mecânica.

A distribuição das preferências ao longo da pesquisa indicam que, apesar das mixagens obtidas por meios automáticos terem apresentado resultados satisfatórios principalmente entre pessoas sem experiência em engenharia de mixagem e áreas correlatas, **os profissionais não reconhecem um desempenho suficiente para equiparar esses recursos automáticos com soluções mecânicas tradicionais.** Embora as tecnologias dedicadas à mixagem automática tenham progredido consideravelmente e sigam em uma tendência de desenvolvimento, no estágio atual de evolução é recomendada a verificação por parte de um especialista dos resultados obtidos por esses recursos.

Para estudos futuros, outros ensaios com músicas de diferentes gêneros poderiam ser conduzidos de modo a **testar as ferramentas em outros contextos** e de modo a conduzir **pesquisas de percepção com mais amostras sonoras.** Outras pesquisas de percepção também poderiam ser conduzidas com uma **maior amostra de respondentes.** Ressalta-se a importância de uma diversidade no nível de experiência dos respondentes, bem como uma separação entre o profissional que realizou a mixagem do profissional que realiza a análise perceptual, uma vez que estudos mostraram que existe uma tendência ao profissional preferir a sua própria mixagem quando exposto à diversas mixagens de uma mesma música.[4][26]

## **6.Considerações finais**

Embora os Sistemas Inteligentes de Mixagem atualmente disponíveis ainda possuam certas limitações [5], esses sistemas estão se desenvolvendo de forma cada vez mais rápida e obtendo resultados cada vez mais pertinentes e sofisticados.[15]

Sendo a profissão do engenheiro de mixagem uma profissão que exige o desenvolvimento de uma série de competências, dada a complexidade empregada na tarefa de mixar, o uso de IMS pode beneficiar os profissionais da área à medida em que trazem a possibilidade de melhorar o fluxo de trabalho através da automatização de tarefas repetitivas. Um ponto a ser considerado é que no atual estágio de desenvolvimento, o engenheiro precisa **checar o trabalho feito por um IMS**, então a interação humano-computador só será benéfica se as ferramentas inteligentes tiverem um desempenho satisfatório e uma compreensão acerca dos objetivos e do contexto da peça sonora sendo mixada.[23]

Segundo De Man e Reiss: 'Ao longo da história da tecnologia, a inovação tradicionalmente encontrou resistência e ceticismo, em particular de usuários profissionais que temem ver suas funções interrompidas ou tornadas obsoletas.(...) No entanto, a evolução da música está intrinsecamente ligada ao desenvolvimento de novos instrumentos e ferramentas, e essencialmente a invenções utilitárias (..) **Esses avanços mudaram a natureza da profissão de engenheiro de som, de basicamente técnica para cada vez mais expressiva**. Geralmente, há mérito econômico, tecnológico e artístico em explorar o imenso poder de computação e flexibilidade que a tecnologia digital de hoje oferece, para se aventurar longe da estrutura rígida do conjunto de ferramentas de produção musical tradicional.'[13] Embora a substituição absoluta de um engenheiro de mixagem por um sistema inteligente ainda não esteja no horizonte [23] é importante levar em consideração as possibilidades benéficas para a profissão que o desenvolvimento de ferramentas mais acessíveis e com cada vez mais funcionalidades traz, uma vez que a finalidade da mixagem é fazer com que a peça sonora se torne uma versão potencializada dela mesma, seja por meios automáticos ou por meios mecânicos.

## REFERÊNCIAS

- [1] Izhaki, Roey. 'Mixing audio : concepts, practices, and tools', 3a ed., Nova Iorque, NY : Routledge, 2017.
- [2] De Man, Brecht. 'Towards a better understanding of mix engineering'. Reino Unido. Tese de doutorado - School of Electronic Engineering and Computer Science, Queen Mary University of London; 2017.
- [3] Savage, Steve. 'Mixing and mastering in the box : the guide to making great mixes and final masters on your computer', Nova Iorque, NY: Oxford University Press, 2014.
- [4] Dobrowohl, Felix A., Andrew J. Milne, and Roger T. Dean 2019. "Timbre Preferences in the Context of Mixing Music" Applied Sciences 9, no. 8: 1695. Disponível em <https://doi.org/10.3390/app9081695>
- [5] Moffat D., 'AI Music Mixing Systems'. Em: Miranda E.R. (eds) 'Handbook of Artificial Intelligence for Music'. Suíça: Springer, Cham, 2021. Disponível em: [https://doi.org/10.1007/978-3-030-72116-9\\_13](https://doi.org/10.1007/978-3-030-72116-9_13)
- [6] Araújo, Danilo Vieira Granato. 'Uma Breve História da Mixagem : Origem, Técnicas, Percepção e Futuros Avanços'. Campinas, SP. Dissertação de mestrado - Universidade Estadual de Campinas, Instituto de Artes; 2015.
- [7] Biblioteca do Congresso Americano[site]. 'The Gramophone'[acesso em 18 out 2021]. Disponível em: <https://www.loc.gov/collections/emile-berliner/articles-and-essays/gramophone/>
- [8] Owsinski, Bobby. 'The Mixing Engineer's Handbook', Boston, MA: Course Technology PTR, 2014.
- [9] Phillips, M., 'Exploring Potential of the Mix'. Em: 'Mixing Music', Nova Iorque, NY: Routledge, 2017.
- [10] Yamaha [site]. 'The History of the DAW' [acesso em 18 out 2021]. Disponível em: <https://hub.yamaha.com/proaudio/pa-history/the-history-of-the-daw/>
- [11] Harvard University[site]. "The History of Artificial Intelligence"[acesso em 18 out 2021]. Disponível em: <https://sitn.hms.harvard.edu/flash/2017/history-artificial-intelligence/>
- [12] Russell, Stuart J., 'Artificial intelligence : a modern approach', Englewood Cliffs, NJ: Prentice Hall, 1995.
- [13] De Man, B., D. Reiss, J., Stables, R., 'Ten Years of Automatic Mixing', 2017

- [14] Lefford, M. Nyssim, Gary Bromham, and David Moffat. "Mixing with Intelligent Mixing Systems: Evolving Practices and Lessons from Computer Assisted Design." 148th Convention of the Audio Engineering Society, 2-5 June, 2020, Online. Audio Engineering Society, Inc., 2020.
- [15] M. N. Lefford, G. Bromham, G. Fazekas, and D. Moffat, "Context-Aware Intelligent Mixing Systems" *J. Audio Eng. Soc.*, vol. 69, no. 3, pp. 128–141, (2021). Disponível em: <https://doi.org/10.17743/jaes.2020.004>
- [16] Terrell, Michael, Andrew Simpson, and Mark Sandler. "The mathematics of mixing." *Journal of the audio engineering society* 62.1/2 (2014): 4-13.
- [17] Hawley, Scott H., Benjamin Colburn, and Stylianos I. Mimitakis. "SignalTrain: Profiling audio compressors with deep neural networks." arXiv preprint arXiv:1905.11928 (2019).
- [18] Perez Gonzalez, Enrique. *Advanced automatic mixing tools for music*. Diss. Queen Mary University of London, 2010.
- [19] Ma, Zheng, Joshua D. Reiss, and Dawn AA Black. "Implementation of an intelligent equalization tool using Yule-Walker for music mixing and mastering." *Audio Engineering Society Convention 134*. Audio Engineering Society, 2013.
- [20] Pujahari, Abhinav. "Towards Automatic Reverb Addition for Production Oriented Multi-Track Audio Mixing." (2017).
- [21] Perez\_Gonzalez, Enrique, and JoshuaD Reiss. "A real-time semiautonomous audio panning system for music mixing." *EURASIP Journal on Advances in Signal Processing* 2010 (2010): 1-10.
- [22] Ramirez, Marco A. Martinez, and Joshua D. Reiss. "Deep learning and intelligent audio mixing." *acoustic guitar* 55 (2017): 24.
- [23] Reiss, Joshua D. "Intelligent systems for mixing multichannel audio." 2011 17th International Conference on Digital Signal Processing (DSP). IEEE, 2011.
- [24] Martinez Ramirez, M., Daniel Stoller, and David Moffat. "A Deep Learning Approach to Intelligent Drum Mixing with the Wave-U-Net." *Audio Engineering Society*, 2021.
- [25] De Man, Brecht, and Joshua D. Reiss. "A semantic approach to autonomous mixing." *Journal on the Art of Record Production (JARP)* (2013).
- [26] De Man, Brecht, et al. "Perceptual evaluation of music mixing practices." *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.
- [27] Senior, Mike. 'Mixing secrets for the small studio'. Nova Iorque, NY: Routledge, 2018.

- [28] Katz, Bob. 'Mastering Audio: the art and the science', 3a ed, Burlington, MA : Focal Press, Taylor & Francis Group, 2015.
- [29] Fletcher, Harvey, and Wilden A. Munson. "Loudness, its definition, measurement and calculation." *Bell System Technical Journal* 12.4 (1933): 377-430.
- [30] Scharf, Bertram. 'Loudness'. Em 'Handbook of Perception - Volume IV', Nova Iorque, NY: Academic Press, 1978.
- [31] Florentine, Mary, Arthur N. Popper, and Richard R. Fay. "Loudness." Nova Iorque, NY: Springer, 2011.
- [32] Iazzetta, F.; Figueiredo, F. L.; Masiero, B. S. Parâmetros Subjetivos em Salas Destinadas à Prática Musical. In: CONGRESSO IBERO-AMERICANO DE ACÚSTICA, 4, 2004a, Guimarães, Portugal. **Anais...** p. 1-8. CD-ROM: id83.
- [33] Barron, M. The subjective effects of the first reflections in concert halls – The need for lateral reflections. **J. Sound Vib.** v. 15, n. 4, p. 475-494, 22April, 1971.
- [34] Schroeder, Manfred R. Binaural dissimilarity and optimum ceilings for concert halls : more lateral sound diffusion. **J. Acoust. Soc. Am.**, v. 65, n. 4, p. 958-963, Apr. 1979.
- [35] Burred, Juan José, Axel Robel, and Thomas Sikora. "Dynamic spectral envelope modeling for timbre analysis of musical instrument sounds." *IEEE Transactions on Audio, Speech, and Language Processing* 18.3 (2009): 663-674.
- [36] Sweetwater[site]. 'How phase meters can help your mixes' [acesso em 18 out 2021]. Disponível em: <https://www.sweetwater.com/insync/phase-meters-can-help-mixes/>
- [37] Välimäki, Vesa, and Joshua D. Reiss. "All about audio equalization: Solutions and frontiers." *Applied Sciences* 6.5 (2016): 129.
- [38] Moore, Austin, Rupert Till, and Jonathan P. Wakefield. "An investigation into the sonic signature of three classic dynamic range compressors." (2016).
- [39] Eliasson, Simon. "Comparing Compressor Interface Designs: How do visual displays on digital compressors impact how audio engineers navigate an interface and the choices they make?." (2019).
- [40] Moore, Austin. 'An Investigation into Non-Linear Sonic Signatures with a Focus on Dynamic Range Compression and the 1176 Fet Compressor'. Diss. University of Huddersfield, 2017.
- [41] User Manual - Soothe 2. Disponível em: [https://storage.googleapis.com/oeksound-downloads/soothe2/soothe2\\_ManualFAQ.pdf](https://storage.googleapis.com/oeksound-downloads/soothe2/soothe2_ManualFAQ.pdf)

- [42] User Manual - Neutron 3. Disponible em:  
<https://s3.amazonaws.com/izotopedownloads/docs/neutron300/en/index.html>
- [43] User Manual - Gullfoss. Disponible em:  
<https://www.soundtheory.com/static/Gullfoss%20Manual.pdf>
- [44] User Manual - smart:EQ3. Disponible em:  
<https://www.sonible.com/wp-content/uploads/2021/05/manual-smartEQ3-1.pdf>
- [45] User Manual - smart:reverb. Disponible em:  
<https://www.sonible.com/wp-content/uploads/2020/12/manual-smartreverb.pdf>
- [46] User Manual - smart:comp. Disponible em:  
[https://www.sonible.com/wp-content/uploads/2020/03/manual-smartComp\\_download.pdf](https://www.sonible.com/wp-content/uploads/2020/03/manual-smartComp_download.pdf)

## GLOSSÁRIO

**agudas** - registro de frequências acima de 4000 Hz

**balanço tonal** - A manutenção de um equilíbrio adequado entre as regiões das frequência

**buses** - forma de rotear o sinal de um ou mais faixas de som de modo a processá-las de forma conjunta

**compatibilidade mono** - termo utilizado para se referir à adaptabilidade de uma mixagem estéreo à sua reprodução em um sistema mono (de apenas um canal) com perda mínima de informação do conteúdo, uma vez que os sistemas mono reproduzem um arquivo estéreo através da soma de ambos os canais (esquerdo (L) e direito (R))

**componente fundamental** - frequência que determina a altura de um som, sendo esta associada às notas musicais (dó, ré, mi...)

**compressores digitais** - compressores confeccionados por meios digitais. Funcionam através de operações matemáticas, sendo altamente customizáveis e precisos.

**compressores FET** - compressores analógicos que empregam transistores de efeito de campo que atuam no sinal de entrada por meio do controle da resistência dos transistores através da alteração da voltagem aplicada ao terminal gate

**compressores ópticos** - modelo de compressor analógico formado por uma fonte luminosa atrelada ao sinal de entrada e uma foto-resistência. A atenuação do sinal se dá por meio da interação entre a fonte luminosa e a foto resistência.

**compressores valvulados** - primeiros modelos de compressores analógicos criados, funcionam através do uso de válvulas sensíveis ao sinal de entrada

**compressores VCA** - modelo mais recente entre os compressores analógicos. Seu funcionamento ocorre através do controle de voltagem de um circuito de amplificação integrado

**Correlação Cruzada Interaural** - grandeza utilizada para quantificar a Impressão Espacial

**curvas de Fletcher-Munson** - curvas obtidas de forma experimental que buscam descrever a percepção de volume do ouvido humano ao longo do espectro de frequência

**decibéis** - unidade de medida de intensidade de som

**desbalanceamento da imagem estereofônica** - quando há mais energia de sinal concentrada em um lado da imagem sonora do que no outro

**Digital Audio Workstation (DAW)** - em português Estação de trabalho de áudio digital, nome dado aos softwares utilizados para gravar, editar, produzir e processar áudio.

**domínio da espacialidade** - diz respeito à sensação de envolvimento do ouvinte pela massa sonora da mixagem e é em grande parte caracterizado pela descorrelação entre os sinais que chegam às orelhas esquerda e direita do ouvinte

**domínio da estereofonia** - diz respeito à qualidade da imagem sonora resultante, sobretudo de seu relevo sonoro (sensação de percepção tridimensional), considerando os aspectos de: localização espacial das fontes sonoras (localization); a largura de imagem sonora, tanto a individual de cada fonte sonora quanto a da imagem geral com um todo (stereo width); definição do contorno e da localização de cada fonte sonora (stereo focus **or sharpness**); esquema de distribuição espacial das fontes sonoras na imagem sonora geral (stereo spread)

**domínio da frequência** - um dos domínios da mixagem que compreende as ondas sonoras com frequências que o ouvido humano é capaz de perceber, ou seja, frequências entre 20 Hz e 20 kHz. Esse domínio é normalmente dividido em registros.

**domínio da intensidade** - diz respeito à intensidade de nível sonoro. No contexto de uma mixagem os níveis de intensidade sonora são manipulados sempre de forma relativa, ou seja, deve ser definido de modo a que os elementos de mixagem sejam inteligíveis entre si e estejam presentes de maneira coesa.

**domínio da profundidade** - diz respeito à distribuição panorâmica através de planos sonoros mais distantes (ao fundo) ou mais próximos do ouvinte

**domínio do tempo** - diz respeito ao tempo no qual a mixagem está inserida de forma direta, imediata e incontornável.

**elemento âmbito de frequência** - está relacionado a equilibrar as características sônicas de cada um dos elementos de mixagem através da sua manipulação espectral individual.

**elemento balanço** - equilíbrio entre os níveis de intensidade sonora em uma mixagem

**elemento dimensão** - aquele pelo qual se ambienta um instrumento, uma voz ou um outro elemento sonoro da mixagem sendo trabalhado através do uso de recursos como por exemplo efeitos de reverberação (reverb) e de atraso (delay)

**elemento dinâmica (dinâmica)** - resultante da gestão e manuseio do âmbito dinâmico geral (utiliza-se normalmente o termo em inglês dynamic range), que pode ser definido como a diferença entre o nível mais fraco e o mais forte de um sinal.

**elemento interesse** - está relacionado à manipulações realizadas durante a mixagem visando a experiência do ouvinte ao longo da música

**elemento panorama (panorama)** - recurso através do qual se pode trabalhar o movimento e a clareza da distribuição espacial dos instrumentos, das vozes e de efeitos na imagem sonora final da mixagem

**efeitos adaptativos** - dentro do contexto de IMS, é uma forma de manipulação sonora baseada na alteração de parâmetros de controle dos recursos aplicados de acordo com uma análise automatizada do sinal de áudio processado ou por uma entrada externa como, por exemplo, por um controlador operado pelo engenheiro de mixagem ou um sensor

**estrutura de ganho** - planejamento e definição dos níveis dos sinais sonoros ao longo da mixagem

**fader digital** - ferramenta utilizada para alterar o nível de intensidade sonora de um canal de áudio.

**faixa de presença** - banda de frequências entre 4 kHz e 6 kHz caracteriza a sensação de proximidade e tangibilidade das vozes, dos instrumentos ou qualquer fonte sonora.

**filtros passa-banda** - corretor de frequência que funciona a partir da inserção de uma frequência de referência (frequência de cut-off) as frequências acima serão cortadas (filtros passa-baixa) ou mantidas (filtros passa-alta) e as frequências abaixo serão mantidas (filtros passa-baixa) ou cortadas (filtros passa-alta).

**filtros shelving** - corretor de frequência que funcionam atenuando ou aumentando (dependendo do valor do ganho) as frequências menores ou maiores (a depender do tipo de filtro) que a frequência de cutoff.

**fonógrafo** - invenção de Thomas Edison que consistia em um cilindro giratório capaz não apenas de armazenar, como também reproduzir informações sonoras

**formante** - frequência que marca a zona na qual se concentra mais energia em um espectro

**frequências ressonantes** - frequência particular ou faixa de frequências que são acentuadas, geralmente devido a algum fator acústico, eletrônico ou mecânico estranho

**gênero** - rótulos utilizados para categorizar peças musicais que possuam características em comum

**gramofone** - invenção de Emile Berliner que consistia em um sistema de gravação e reprodução baseado em discos planos, que inaugurou a possibilidade de gerar cópias dos registros sonoros com facilidade, possibilitando assim o início da comercialização de gravações musicais

**graves** - registro de frequências entre 60 e 250 Hz

**grau automático** - dentro do contexto de IMS, é um grau de interação computador-humano no qual nenhuma interação humana é empregada e o sistema fornece o produto final processado

**grau exploratório** - dentro do contexto de IMS, é um grau de interação computador-humano no qual o IMS apenas fornece informações ao engenheiro acerca do estágio atual dos processos de mixagem.

**grau independente** - dentro do contexto de IMS, é um grau de interação computador-humano no qual várias tarefas são executadas pelo sistema, porém com a supervisão de um engenheiro de mixagem que valida ou não cada ação concluída pelo IMS

**grau recomendativo** - dentro do contexto de IMS, é um grau de interação computador-humano no qual o IMS tem a habilidade de analisar e interpretar em tempo real as ações do engenheiro ao longo do processo de mixagem e fornece ao usuário recomendações de alterações a serem feitas.

**interação computador-humano** - nível de interferência humana que um IMS prevê, em outras palavras, o nível de automação do IMS.

**Impressão Espacial (Spatial Impression)** - efeito psicoacústico causado pelas reflexões sonoras que atingem o ouvinte fazendo-o criar mentalmente uma sensação acústica espacial

do ambiente, sendo influenciada pela difusão sonora e pela dissimilaridade dos sons percebidos por ambas as orelhas.

**Intelligent Mixing Systems(IMS)** - em português Sistemas Inteligentes de Mixagem, nome dado ao campo de estudo que abarca formas de automatizar o processo da mixagem por meio de inteligência artificial

**in-the-box** - nome dado ao processamento de áudio digital realizado no computador, através de ferramentas virtuais

**knee** - parâmetro de um compressor que determina ao quão brusco será a transição entre o sinal não comprimido e o início da compressão

**loudness** - pode ser definido como o termo psicológico usado para descrever a magnitude de uma sensação auditiva ou como a intensidade subjetiva de um som

**macrodinâmica** - está relacionada a variações em nível de intensidade sonora para eventos mais amplos que uma nota musical

**manipulação sonora** - maneira como um IMS irá agir de modo a alterar aspectos sonoros de um elemento musical. Pode ocorrer de duas maneiras: por meio de efeitos adaptativos ou por meio de uma transformação direta.

**mascamamento** - elementos sonoros não são percebidos devido uma sobreposição de informações em uma mesma banda de frequências, Nas regiões subgraves e graves o mascaramento tende a se tornar problemático, uma vez que, para a percepção humana, essas bandas de frequência precisam ter muito mais energia para serem percebidas com a mesma intensidade que as bandas médias e médio-agudas

**master** - produto final da mixagem que consiste em um arquivo de um ou mais canais. Também é o nome dado ao canal no qual todas os canais de um sistema multicanal são somados

**medidor de correlação** - indica a compatibilidade mono de um sinal mostrando o grau de defasagem entre canal direito e canal esquerdo

**medidor de fase** - mostra o grau de descorrelação ou correlação entre os sinais dos canais esquerdo e direito, através de uma modulação ortogonal como para formar figuras de Lissajous.

**medidores de pico** - mostram o valor de pico do sinal naquele instante (valor instantâneo)

**medidores de valor médio** - mostram uma média do nível de intensidade sonora do sinal em uma janela de tempo pré-definida

**médio-agudas** - registro de frequências entre 2000 e 4000 Hz

**médio-graves** - registro de frequências entre 250 e 2000 Hz

**microdinâmica** - está relacionada à variação em nível de intensidade sonora ocorrendo no detalhe, ou seja, no acontecimento de cada nota musical tocada

**mixagem** - nome dado ao processo no qual um material sonoro composto de múltiplas faixas é processado e combinado em um produto final

**mix bus** - é um canal por onde todos os elementos são enviados antes de serem enviados ao master no qual o engenheiro de mixagem pode agir, geralmente com pequenas alterações, no produto final como um todo

**mixer** - termo utilizado para se referir ao engenheiro de mixagem

**monitoramento** - nome dado ao equipamento no qual a mixagem será ouvida, podendo ser um fone de ouvido ou um conjunto de caixas de som

**mono** - formato de gravação e reprodução de som que emprega apenas um canal

**parciais harmônicos** - grandezas múltiplas inteiras da frequência da fundamental

**parciais inarmônicos** - grandezas múltiplas não inteiras da fundamental

**plugins** - softwares utilizados para processamento de áudio ou como instrumentos virtuais que são integrados às DAWs

**pré-delay** - se refere ao tempo entre o sinal original e a primeira reflexão

**processo de tomada de decisão** - decidir qual será o método empregado na construção do sistema inteligente em si, ou seja, na forma como regras serão modeladas e embutidas em um sistema

**stems** - nome dado aos arquivos individuais de cada elemento da música

**stereo** - formato de gravação e reprodução de som que emprega dois canais (direito e esquerdo)

**subgraves** - registro de frequências abaixo de 60 Hz

**telegrafone** - invenção de Valdemar Poulsen, foi o primeiro sistema de gravação com o uso de magnetismo. Seu funcionamento ocorria através da atuação de um eletroímã que magnetiza um fio de aço de acordo com os sons captados, que podiam ser reproduzidos pela inversão do sistema

**tempo de decaimento** - tempo necessário para que as reflexões diminuam de 60 dB de intensidade em relação ao som direto

**timbre** - qualidade sonora de um som que permite sua distinção

**transformação direta** - dentro do contexto de IMS, é uma forma de manipulação sonora que não depende de nenhuma etapa intermediária de interação uma vez que a saída do sistema fornece o produto final programado

**transiente** - sinal de grande amplitude e curta duração em tempo que marca o início de um som

**visão** - também chamada de conceito, é o nome dado para o resultado emergente desejado do processo da mixagem