

UNIVERSIDADE FEDERAL DE JUIZ DE FORA
FACULDADE DE COMUNICAÇÃO
RÁDIO, TV E INTERNET

Ana Carolina de Souza Santos

Sonidos Sintéticos e falas humanas: a IA e a direção de atores na audioficção

Juiz de Fora

2025

Ana Carolina de Souza Santos

Sonidos Sintéticos e falas humanas: a IA e a direção de atores na audioficção

Memorial Descritivo apresentado como pré-requisito para a obtenção do grau de bacharel do curso de Rádio, TV e internet, da Faculdade de Comunicação da Universidade Federal de Juiz de Fora.

Orientador: Prof Dr. João Paulo Malerba
Coorientador: Profa. Dra. Márcia Cristina Vieira Falabella

Juiz de Fora

2025

Ana Carolina de Souza Santos

Sonidos Sintéticos e falas humanas: a IA e a direção de atores na audioficção

Trabalho de Conclusão de Curso apresentado ao curso de Rádio, TV e Internet, da Faculdade de Comunicação da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do grau de bacharel.

Aprovada em 11 de Agosto de 2025

BANCA EXAMINADORA

Prof. Dr. João Paulo Malerba - Orientador

Universidade Federal de Juiz de Fora

Profa. Dra. Márcia Cristina Vieira Falabella - Co Orientadora

Universidade Federal de Juiz de Fora

Prof. Dr. Talison Pires Vardiero

Universidade Federal de Juiz de Fora

Profa. Dra. Kérley Winques

Universidade Federal de Juiz de Fora



UNIVERSIDADE FEDERAL DE JUIZ DE FORA

ATA DE DEFESA DE TRABALHO DE CONCLUSÃO DE CURSO DE GRADUAÇÃO GRADUAÇÃO EM RÁDIO TV E INTERNET (RTVI)

Formato da Defesa: (X) presencial () virtual () híbrido

Ata da sessão (X) pública () privada referente à defesa do Trabalho de Conclusão de Curso intitulado "**Sonidos Sintéticos e falas humanas: a IA e a direção de atores na audioficção**", para fins de obtenção do grau de Bacharel em RTVI, pelo(a) **Ana Carolina Souza Santos** discente (matrícula_202191024_), sob orientação da Prof.(a) Dr.(a) **João Paulo Malerba** e coorientação da Prof. a) Dr(a) **Márcia Falabela**, na **Faculdade de Comunicação** da Universidade Federal de Juiz de Fora.

Ao **11** dia do mês de **Agosto** do ano de **2025**, às **14** horas, na **sala 220** da Universidade Federal de Juiz de Fora (UFJF), reuniu-se a Banca examinadora, composta pelos seguintes membros:

Titulação	Nome	Na qualidade de:
Doutor	João Paulo Malerba	Orientador
Doutora	Márcia Cristina Vieira Falabela	coorientadora
Doutora	Kerley Winques	Membro da Banca
Doutor	Talison Vardieiro	Membro da Banca

*Na qualidade de (opções a serem escolhidas):

- Orientador (a)
- Coorientador
- Membro da banca

AVALIAÇÃO DA BANCA EXAMINADORA

Tendo o(a) senhor(a) Presidente declarado aberta a sessão, mediante o prévio exame do referido trabalho por parte de cada membro da Banca, o(a) discente procedeu à apresentação de seu Trabalho de Conclusão de Curso de Graduação e foi submetido(a) à arguição pela Banca Examinadora que, em seguida, deliberou sobre o seguinte resultado:

(X) APROVADO

() REPROVADO, conforme parecer circunstanciado, registrado no campo Observações desta Ata e/ou em documento anexo, elaborado pela Banca Examinadora

Nota: _____

Observações da Banca Examinadora caso haja necessidade de anotações gerais sobre o Trabalho de Conclusão de Curso e sobre a defesa, as quais a banca julgue pertinentes

A banca recomenda que o trabalho seja indicado para premiações

Nada mais havendo a tratar, o(a) senhor(a) Presidente declarou encerrada a sessão de Defesa, sendo a presente Ata lavrada e assinada pelos(as) senhores(as) membros da Banca Examinadora e pelo(a) discente, atestando ciência do que nela consta.

INFORMAÇÕES

Para fazer jus ao título de bacharel, a versão final do Trabalho de Conclusão de curso, considerado Aprovado, devidamente conferida pela Secretaria do Curso de (colocar o nome do curso), deverá ser tramitada para o Repositório Institucional, dentro do prazo de 72 horas da realização da banca.

Juiz de Fora, 07 de agosto de 2025.

Assinatura digital dos membros da Banca Examinadora



Documento assinado eletronicamente por **João Paulo Carrera Malerba, Professor(a)**, em 12/08/2025, às 08:08, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Marcia Cristina Vieira Falabella, Professor(a)**, em 12/08/2025, às 12:17, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Talison Pires Vardiero, Professor(a)**, em 12/08/2025, às 14:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Kérley Winkes, Professor(a)**, em 12/08/2025, às 14:32, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Ana Carolina de Souza Santos, Usuário Externo**, em 20/08/2025, às 10:42, conforme horário oficial de Brasília, com fundamento no § 3º do art. 4º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no Portal do SEI-Ufjf (www2.ufjf.br/SEI) através do ícone Conferência de Documentos, informando o código verificador **2539174** e o código CRC **8873821C**.

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Santos, Ana Carolina.

Sonidos Sintéticos e falas humanas: : a IA e a direção de atores na audioficção / Ana Carolina Santos. -- 2025.
88 f.

Orientador: João Paulo Malerba

Coorientadora: Márcia Cristina Vieira Falabella

Trabalho de Conclusão de Curso (graduação) - Universidade Federal de Juiz de Fora, Faculdade de Comunicação Social, 2025.

1. Audioficção. 2. Direção de atores. 3. Inteligência artificial. 4. Clonagem de voz. 5. Comunicação e tecnologia. I. Malerba, João Paulo, orient. II. Vieira Falabella, Márcia Cristina, coorient. III. Título.

AGRADECIMENTOS

Primeiramente, agradeço ao meu pai, João, que agora, espiritualmente e em meu coração, me guia e me dá forças todos os dias. Me ensinou que viver requer coragem, então eu deveria viver para ser feliz e não viver em vão. Que me ensinou a nunca desistir, a lutar pelo que quero e, acima de tudo, o verdadeiro significado do amor. À minha mãe, Jacqueline, que foi forte mesmo passando por tantas dificuldades e nunca me deixou faltar nada. Depois da perda do meu pai, precisou ser ao mesmo tempo mãe e pai e o fez com amor e coragem. Você é uma das minhas maiores inspirações. À minha irmã, que se inspira em mim e, sem saber, me fez ser forte tantas vezes para dar um exemplo para ela. Aos meus avós, aos que estão aqui e aos que já partiram, que são exemplos de ancestralidade. Agradeço também aos demais familiares.

Esse TCC não seria o mesmo sem os meus amigos Omar e Manu, então gostaria de agradecer especialmente a eles. Também agradeço aos amigos que marcaram diferentes fases da minha vida. Larissa, Miguel e Larah, vocês fizeram meu ensino médio inesquecível e sou muito grata por nossa amizade ter ultrapassado os portões da escola. À Maysa, minha amiga de vida, que sempre esteve presente em todas as fases da minha trajetória. À Ariane, pelos puxões de orelha e conselhos sempre tão pontuais. Você é uma pessoa incrível por dentro e por fora. À Isadora, você se tornou muito mais que uma amiga, virou uma irmã de outra mãe. Nossa conexão é coisa de outro mundo, e eu me inspiro muito em você. Obrigada por deixar a faculdade muito mais divertida e por estar comigo em tantos momentos. Obrigada também à minha grande cunhada Poli por sempre me acolher tão bem, você é garra e genialidade.

Não poderia deixar de dedicar também este TCC à minha garota, minha princesa, meu amor, Jeniffer. Obrigada por não ter me deixado desistir, por acreditar em mim e no meu potencial, por todos os momentos compartilhados e por me mostrar que a vida é muito mais bonita quando se ama alguém. Você me ensinou que o amor verdadeiro é leve e aconchegante. Você é luz, você é casa.

Aos professores, especialmente ao Talison, que virou também um amigo e que me ensinou muita coisa, principalmente a acreditar que eu consigo e no meu potencial, e me deu um grande exemplo do profissional que quero ser no futuro. Ao Cris, Paulo Roberto, Flávio, Janaina, Nilson e Teresa Medeiros, cujas aulas foram essenciais para minha formação. Aos técnicos da 115, que são peças fundamentais para a faculdade inteira. Ao pessoal do Fórum e da Produtora e aos amigos da faculdade, que tornaram essa jornada mais rica e especial. Às

minhas entrevistadas Talita Magnolo, que agora é orientadora de pesquisa e vem me fazendo ter um outro olhar sobre a IA. A incrível Renata Lana, que me concedeu uma conversa-chave para minha pesquisa e me fez olhar a IA com um olhar mais mercadológico e essencial. Ao professor e orientador João Paulo Malerba, por embarcar nessa loucura comigo, por sua paciência e por me ajudar a cumprir prazos mesmo quando pareciam impossíveis. À professora, coorientadora e amiga Márcia Falabella, que fez meu amor pelo teatro crescer ainda mais, despertando em mim um olhar especial para a direção de atores. Obrigada por todos os conselhos, por sua generosidade e por, em muitos momentos, ser mais que uma professora, uma amiga que admiro profundamente.

Por último, não poderia deixar de agradecer a mim mesma. Mesmo passando por tantas coisas ao longo da vida, sempre confiei que eu conseguiria tudo o que sonhasse. Bastava não desistir e o resto eu dava um jeito. E eu consegui. Eu vou me formar!!

Dedico este trabalho ao meu pai, que me ensinou a olhar para as estrelas e a sonhar. E agora, brilha entre elas, iluminando meu caminho para sempre.

RESUMO

Este memorial descritivo apresenta uma pesquisa experimental e prática que investiga a relação entre direção de atores, inteligência artificial sonora e clonagem de voz em produções ficcionais em áudio. A pesquisa parte de quatro eixos principais. O primeiro analisa a evolução dos formatos de áudio, retomando o percurso histórico que vai das radionovelas ao podcast, com foco nas mudanças tecnológicas, estéticas e de escuta. O segundo discute a direção de atores e suas técnicas, explorando métodos de construção da performance vocal. O terceiro aborda a expressividade da voz, os processos de escuta e os Viewpoints vocais como ferramentas de criação e presença cênica na atuação sonora. O quarto eixo investiga os usos da inteligência artificial e da clonagem de voz, refletindo não apenas sobre seus limites expressivos e impactos profissionais, mas também sobre questões éticas, como o uso indevido de vozes, os riscos de substituição de profissionais e os vieses presentes nos algoritmos de geração vocal. A parte prática consiste na criação de dois episódios a partir do mesmo roteiro: um com atuação humana dirigida e outro com vozes clonadas por IA. A comparação entre ambos permite observar os efeitos da direção sobre a performance. Um terceiro episódio reúne os resultados da pesquisa e inclui entrevistas com especialistas em voz e tecnologia, além de ouvintes não especializados, ampliando a reflexão sobre as diferenças entre a atuação humana e a simulação artificial.

Palavras-chave: Audioficção; Direção de atores; Inteligência artificial; Clonagem de voz; Comunicação e tecnologia.

ABSTRACT

This descriptive report presents an experimental and practical study investigating the relationship between actor direction, sound-based artificial intelligence, and voice cloning in fictional audio productions. The research is structured around four main axes. The first analyzes the evolution of audio formats, tracing the historical trajectory from radio dramas to podcasts, with a focus on technological, aesthetic, and listening changes. The second discusses actor direction and its techniques, exploring methods for constructing vocal performance. The third addresses vocal expressivity, listening processes, and vocal Viewpoints as tools for creation and stage presence in audio acting. The fourth axis examines the uses of artificial intelligence and voice cloning, reflecting not only on their expressive limits and professional impacts but also on ethical issues such as voice misuse, risks of professional replacement, and biases embedded in vocal generation algorithms. The practical component consists of creating two episodes from the same script: one directed with human performance and another using AI-cloned voices. The comparison between these allows observation of the effects of direction on performance. A third episode compiles the research results, including interviews with voice and technology specialists as well as non-expert listeners, broadening the reflection on differences between human acting and artificial simulation.

Keywords: Podcast; Audio fiction; Actor direction; Artificial intelligence; Voice cloning; Communication and technology.

LISTA DE APÊNDICES

APÊNDICE A - Roteiro do Episódio 01 e 02.....	61
APÊNDICE B - Perfil dos Personagens.....	67
APÊNDICE C - Moodboard dos Personagens.....	69
APÊNDICE D - Ordem do Dia – Gravação Ep 01 (17/01).....	70
APÊNDICE E - Cronograma de Produção.....	71
APÊNDICE F - Roteiro base episódio 3.....	72

LISTA DE FIGURAS:

Figura 1: Gráfico - Consumo de áudio no Brasil.....	13
Figura 2: Gráfico - Aumento de ouvintes de Podcast no Brasil.....	15
Figura 3: Gráfico - Aumento de ouvintes no mundo por gênero de podcast de 2019 a 2023..	17
Figura 4: Imagem - Tela de edição do episódio um no Adobe Premiere (2025).....	45
Figura 5: Imagem - Página da clonagem de vozes inicial e final no Eleven Labs (2025).....	46
Figura 6: Aba de ajuste de voz no Eleven Labs (2025).....	47
Figura 7: Imagem autoral, aba de ajuste de voz, texto e edição no Eleven Labs (2025).....	48
Figura 8: Página do Eleven Labs com o texto do episódio 2.....	49
Figura 9: Página do Eleven Labs com o texto do episódio 2.....	50
Figura 10: Sessão de linha do tempo e player do Eleven Labs 2.....	50
Figura 11: Tela de edição do episódio dois no Adobe Premiere.....	51

SUMÁRIO

SUMÁRIO.....	
1. Introdução.....	8
2. Evolução da produção sonora: do rádio tradicional ao podcast.....	10
2.1. O rádio e sua influência.....	11
2.2. O surgimento do podcast.....	14
2.3. O crescimento dos podcasts ficcionais.....	16
2.4. Radionovelas no brasil:a voz como protagonista da atuação.....	18
3. O papel da direção de atores na construção da performance.....	19
3.1 - O sistema Stanislavski: fundamentos e principais técnicas.....	20
3.2 - A técnica Chubbuck: substituição e objetivo geral.....	22
3.3 - Harold Guskin: Como parar de atuar.....	23
4. Voz, som e viewpoints vocais.....	26
4.1. A voz como identidade.....	27
4.2. Escuta, som e ruídos: para além da fala.....	27
4.3. Viewpoints vocais: um método de escuta.....	29
5. IA: origem, conceito e contexto atual.....	31
5.1. Bancos de dados, viés algorítmico e os desafios éticos da clonagem.....	32
5.2. Clonagem de voz: experiência do ouvinte e transformações no mercado sonoro.....	34
6. Memorial Descritivo: Por dentro do Jogo.....	37
6.1. Decisão de jogar: a seleção do tema.....	37
6.2. Definição da estória: escolhendo o tabuleiro.....	38
6.3. A formação do elenco: as peças do jogo.....	39
6.4. Ensaios: arrumando as peças no tabuleiro.....	40
6.5 - Dia da gravação: play - que o jogo comece.....	42
6.6 - Pós-jogo: hora de rever a partida.....	43
6.7 - Análise: Revendo as partidas e comparando os resultados.....	51
7 - Considerações finais.....	56
REFERÊNCIAS.....	58

1. Introdução

Este trabalho apresenta uma pesquisa prática e teórica que investiga a relação entre a direção de atores, a inteligência artificial sonora e a clonagem de voz em produções ficcionais em áudio, com ênfase no formato de podcast denominado como audioficções. A proposta parte da hipótese de que a direção de atores, aliada a técnicas de atuação bem estruturadas, ainda desempenha um papel insubstituível na criação de performances vocais expressivas, algo que as inteligências artificiais, mesmo com seus avanços, ainda não conseguem replicar com autenticidade.

A pesquisa organiza-se em quatro frentes principais. A primeira trata da trajetória da produção sonora, desde o rádio e as radionovelas até os atuais podcasts, evidenciando transformações estéticas, técnicas e culturais que moldaram o modo de escutar e contar histórias em áudio. Autores como Brandão e Fernandes (2014) e Cunha e Haussen (2003) ajudam a construir esse panorama histórico, enquanto, com o crescimento do consumo de podcasts, especialmente os ficcionais, observa-se uma reconfiguração estética e tecnológica do formato. Como discute Vicente (2018), esse cenário abre espaço para novas formas de experimentação narrativa e sonora.

A segunda frente se debruça sobre a direção de atores, explorando técnicas clássicas e contemporâneas como as propostas por Stanislavski (2003), Chubbuck (2014) e Guskin (2010). O objetivo é entender como o trabalho de direção pode potencializar a expressividade vocal e garantir uma entrega mais autêntica e conectada emocionalmente, mesmo em formatos que se sustentam apenas pelo som. Essas técnicas foram fundamentais para a preparação dos atores nos episódios produzidos nesta pesquisa.

Já a terceira frente investigada diz respeito à escuta, à presença cênica e à voz como ferramenta dramatúrgica. Neste ponto, os *Viewpoints Vocais*, inspirados nas proposições de Bogart e Landau (2010), são utilizados como base para refletir sobre a construção da expressividade em ambientes sonoros. A pesquisa entende a voz como corpo em ação, capaz de ocupar o espaço acústico com ritmo, intenção e relação com o outro, mesmo em formatos puramente auditivos.

Por fim, a quarta e última linha de investigação da pesquisa aborda o uso da inteligência artificial, trazendo um panorama sobre o que é a IA e sua trajetória histórica, com foco na clonagem de voz. Gomes (2010) analisa tanto seu funcionamento quanto os impactos éticos, criativos e profissionais que envolvem seu uso. Embora essas tecnologias ofereçam

novas possibilidades criativas, também levantam preocupações quanto à substituição de profissionais da voz e ao uso indevido de identidades vocais (Rosa, 2023). Além disso, estudos como os de Silva (2020) revelam que os algoritmos podem reproduzir preconceitos estruturais, afetando especialmente grupos minorizados. No campo sonoro, isso se manifesta na valorização de vozes brancas, masculinas e neutras, em detrimento da diversidade vocal. Para Winkes e Longhi (2022; 2023), esses sistemas funcionam como mediações sociotécnicas que podem reforçar desigualdades culturais. Seus estudos mostram que os modelos de síntese vocal incorporam escolhas técnicas excludentes e padronizadoras. Assim, a pesquisa busca refletir criticamente sobre como essas ferramentas impactam a representatividade e a expressividade nas mídias sonoras.

Para aprofundar essas reflexões, foram produzidos dois episódios de podcast a partir do mesmo roteiro, adaptado por mim da obra *"Casal é Tudo Igual"*, de Luís Fernando Veríssimo. Um dos episódios contou com atores humanos, que passaram por um processo completo de preparação, criação de personagens e direção antes e durante a gravação. O outro foi realizado com vozes clonadas por inteligência artificial utilizando a ferramenta Eleven Labs. O terceiro episódio, que constitui o produto final da pesquisa, apresenta uma comparação entre as duas versões à luz da fundamentação teórica, permitindo observar o papel da direção de atores na construção da expressividade vocal. Também analisa o desempenho da IA na replicação de vozes e sua capacidade de atingir um nível de escuta e emoção comparável ao dos atores. Esse episódio final reúne ainda opiniões de especialistas em voz e tecnologia, além de ouvintes não especializados, ampliando o debate sobre os limites entre o humano e o artificial na atuação sonora.

2. Evolução da produção sonora: do rádio tradicional ao podcast

Desde os primórdios da humanidade, surgiu a necessidade de socialização, interação e comunicação entre os seres humanos. Comunicar-se, seja por gestos, sons, palavras, desenhos ou outras formas de expressão, tornou-se um ponto chave para que os indivíduos consigam expressar seus desejos e necessidades individuais e coletivas. A comunicação e as formas de se comunicar evoluíram juntamente com os seres humanos, criando grupos sociais e entidades como família, igreja, escola e outros grupos de socialização na vida pública e na sociedade. Com essa evolução e necessidade de se comunicar, surgiram vários meios de comunicação individual.

A história das comunicações evolui no mesmo trilho da história da humanidade. Pelo simples fato de que a última só existe porque de alguma forma foi relatada de pai para filho, de tribo para tribo, de cidade para cidade, de país para país por meio de indivíduos e de tecnologias que expandiram os recursos do corpo humano. Os meios de comunicação são extensões de nosso corpo, e suas mensagens, de nossos sentir e pensar. (Gontijo, 2004, p. 11)

Ou seja, os meios de comunicação, agindo como extensões do nosso corpo, evoluíram juntamente com a evolução do ser humano. A comunicação direta, que antes exigia a presença simultânea do transmissor e do receptor no mesmo ambiente, foi substituída pela necessidade de comunicar-se com mais pessoas. Os primeiros meios de comunicação de massa surgiram com o intuito de levar uma mensagem, seja de qual cunho fosse, para outros receptores. Para alguns autores, como Karen A. Foss e Stephen Littlejohn (Littlejohn; Foss, 2008, p. 42), os primeiros meios de comunicação de massa surgiram com o advento do cinema e do rádio já para outros estudiosos, como Wilson Dizard (Dizard, 2000, p. 27), os primeiros meios de comunicação de massa são os livros e jornais, com a invenção da prensa por Gutenberg, na Alemanha, no século XV. Os livros e panfletos, que antes eram comercializados de forma artesanal, copiados à mão, ganharam um instrumento importante de cópia. Com esse instrumento poderoso, surgiram também os jornais impressos e a necessidade de levar a informação além de apenas histórias, como se via nos livros, incluindo informações cotidianas e de cunho jornalístico. Depois vieram o rádio, o cinema, a televisão e os novos meios de comunicação que vêm surgindo, e agora, com as inteligências artificiais, outros ainda surgirão. Para entender um pouco essa rápida mudança nos meios, Jorge Zahar (2004) cita Wilson Dizard (2000) para falar sobre esse avanço:

As atuais mudanças são a terceira grande transformação nas tecnologias da mídia de massa nos tempos modernos. A primeira aconteceu no século XIX, com a introdução das impressoras a vapor e do papel de jornal barato. O resultado foi a primeira mídia de massa verdadeira: os jornais "baratos" e as editoras de livros e revistas em grande escala. A segunda transformação ocorreu com a introdução da transmissão por ondas eletromagnéticas (o rádio e a televisão). A terceira transformação na mídia de massa, que estamos presenciando agora, envolve uma transição para a produção, armazenagem e distribuição da informação e do entretenimento estruturadas em computadores. Ela nos leva para o mundo dos computadores multimídia, compact discs, bancos de dados portáteis, redes nacionais de fibras óticas, mensagens enviadas por fax de última geração, páginas de Web e outros serviços que não existiam há 20 anos. (Dizard, 2000 apud Zahar, 2004, p. 54).

Essa terceira transformação midiática, citada por Dizard, consolidou o avanço digital com a internet e as plataformas de streaming. O rádio, por exemplo, reinventou-se, tornando-se um meio versátil e flexível. Neste contexto, busca-se traçar o papel do rádio na formação de hábitos de consumo de áudio no Brasil. A tradição do brasileiro em ouvir rádio, seja para música, notícias ou entretenimento, criou uma cultura favorável ao surgimento e consolidação dos podcasts. Esse olhar permite entender como a internet transformou o consumo de música e informação, mantendo a conexão íntima com o áudio, agora adaptado às inovações tecnológicas.

O crescimento das plataformas de streaming e o isolamento social durante a pandemia contribuíram para essa consolidação e impulsionaram a popularização dos podcasts. No entanto, o avanço das inteligências artificiais traz novos desafios e possibilidades. A clonagem de vozes e a criação de vozes sintéticas levantam questões sobre autenticidade, direitos autorais e novas narrativas sonoras. Analisar essa transição permite compreender não só as mudanças tecnológicas, mas também as implicações culturais e éticas desse cenário em transformação.

2.1. O rádio e sua influência

O rádio, como meio de comunicação, teve um papel predominante não apenas no Brasil, mas em todo o mundo desde sua criação no final do século XIX. No decorrer do século XX, consolidou-se como um dos principais meios de comunicação de massa, oferecendo transmissões ao vivo de notícias, músicas, histórias fictícias e muito mais. Uma de suas características mais marcantes é a inclusão, alcançando também indivíduos que não

possuíam alfabetização suficiente para consumir os jornais escritos. Dessa forma, o rádio tornou-se uma ferramenta estratégica utilizada não apenas para entretenimento e informação, mas também para fins políticos e econômicos, contribuindo para a disseminação de ideias e a construção de hegemonias. Seu alcance instantâneo e massivo permitiu a formação de opinião e influência social em escala global.

A produção radiofônica, especialmente no período de ouro do rádio, seguia uma estrutura de produção linear, diferente das formas mais descentralizadas e mais democráticas de produção sonora contemporâneas, como o podcast. Segundo Eduardo Vicente,

Enquanto a produção radiofônica tradicionalmente opera sob rígidos padrões institucionais e uma estrutura hierárquica, os podcasts oferecem maior flexibilidade, permitindo aos produtores independentes explorarem formatos experimentais e uma estética sonora menos convencional (Vicente, 2018, p. 47).

Essa descentralização possibilita uma maior diversidade de vozes e narrativas, contribuindo para uma democratização do acesso à produção de conteúdo sonoro.

Um dos casos mais emblemáticos do poder do rádio ocorreu em 30 de outubro de 1938, quando Orson Welles, através da *Columbia Broadcasting System* (CBS) de Nova York, narrou uma suposta invasão marciana baseada na obra *A Guerra dos Mundos*, de H.G. Wells. A transmissão radiofônica, apresentada em formato de reportagem jornalística, gerou pânico entre os ouvintes que não haviam compreendido seu caráter fictício. Relatos da época indicam que milhares de pessoas rezaram, fugiram ou buscaram informações em outras fontes, evidenciando o impacto do rádio como meio de comunicação de massa e sua capacidade de mobilização social.

A força da rádio como meio de comunicação também foi impulsionada pela indústria musical. A possibilidade de sintonizar emissoras para ouvir músicas fez com que a rádio se tornasse ainda mais presente no cotidiano das pessoas, ampliando sua audiência e reforçando seu impacto na sociedade. Além de ser uma fonte de informação e entretenimento, a rádio se consolidou como um dos principais meios de consumo musical por décadas. Essa forte relação entre o rádio e a indústria musical teve um papel fundamental no avanço da comunicação sonora. Com o desenvolvimento tecnológico, surgiram novas formas de consumo, como tocas-fitas, leitores de fitas cassete e, posteriormente, dispositivos como MP3 players, iPods e telefones multimídia. A chegada da internet expandiu ainda mais as possibilidades, permitindo a transmissão de áudio por meio de blogs, reprodutores digitais e plataformas de streaming.

Além disso, essa relação impulsionou investimentos em pesquisa e novas tecnologias, fomentando o desenvolvimento da indústria cultural como um todo. Esses investimentos beneficiaram diretamente o rádio, garantindo sua adaptação às transformações tecnológicas e ao surgimento de novos formatos de transmissão sonora. A música, por sua vez, continua sendo um dos principais atrativos para a audiência, funcionando como porta de entrada para outros conteúdos sonoros. Muitas pessoas acessam plataformas de streaming em busca de música e acabam explorando também podcasts, assim como sintonizam rádios inicialmente pelo repertório musical e, conseqüentemente, consomem notícias e outros programas. Uma pesquisa realizada pelo Grupo Globo (2020) demonstra o impacto do consumo de áudio no Brasil, evidenciando sua relevância na descoberta e no engajamento com diferentes formatos sonoros. Conforme ilustrado no Gráfico 1, o áudio mantém grande força no país, consolidando-se como um elemento central no consumo de mídia sonora.

Figura 1: Gráfico - Consumo de áudio no Brasil



Fonte: Grupo Globo (2020)

Pelo gráfico, é possível observar a expressiva mudança nos hábitos de consumo de áudio, com um grande número de pessoas utilizando aplicativos como Deezer e Spotify para ouvir música. Além disso, os dados demonstram que o número de ouvintes de áudio digital no Brasil, independentemente do formato, cresceu cerca de 23% em apenas cinco anos, um aumento significativo que evidencia a transformação no modo como o público consome conteúdo sonoro.

Esse cenário de crescimento acelerado do consumo de áudio digital abriu espaço para novas formas de produção e distribuição de conteúdo sonoro, culminando na popularização dos podcasts.

2.2. O surgimento do podcast

A convergência do avanço das tecnologias digitais e a crescente demanda por formatos de comunicação mais flexíveis criou o cenário perfeito para o surgimento do podcast. Esse novo formato se diferencia do rádio tradicional ao possibilitar que qualquer pessoa produza e distribua conteúdo sonoro de forma independente, sem as amarras de uma estrutura hierárquica e centralizada. A combinação de dispositivos portáteis, softwares acessíveis e a popularização dos blogs preparou o terreno para uma revolução na produção e no consumo de áudio. Como ressaltava Ben Hammersley em seu artigo *Audible Revolution* (2004):

Com o benefício da retrospectiva, tudo parece bastante óbvio. Tocadores de MP3, como o iPod da Apple, em muitos bolsos, softwares de produção de áudio baratos ou de graça, e blogs, uma parte estabelecida da internet; todos os ingredientes estão lá para um novo boom no rádio amador. Mas como chamamos isso? Audioblogging? Podcasting? GuerillaMedia? (Hammersley, 2004, tradução nossa).

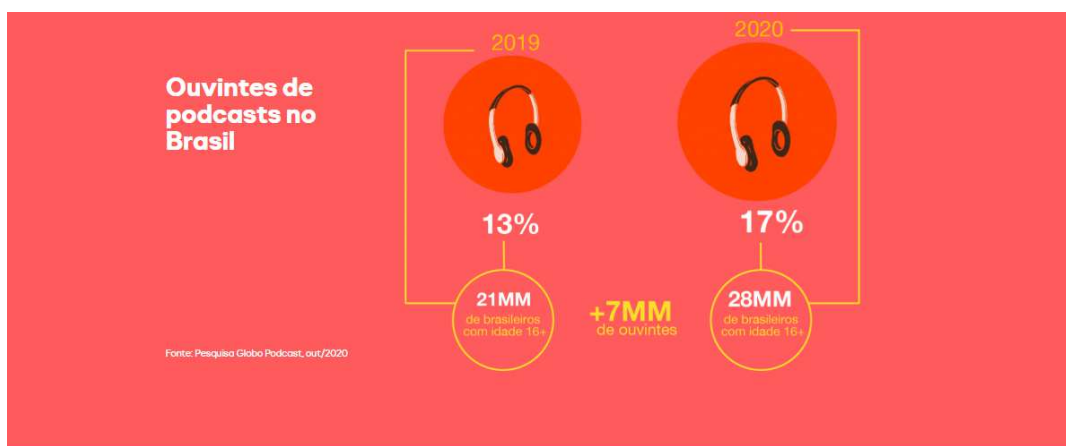
Eduardo Vicente (2018) menciona que, embora Hammersley tenha sido um dos primeiros a sugerir a denominação para esse novo modelo de produção sonora, o termo "podcast" inicialmente não estava associado à transmissão via RSS. Essa funcionalidade surgiu apenas em setembro de 2004, quando Dannie Gregoire a utilizou para descrever o processo desenvolvido por Adam Curry, permitindo que episódios fossem atualizados regularmente e baixados automaticamente nos dispositivos dos ouvintes. Esse avanço tecnológico consolidou o modelo que hoje conhecemos como podcasting.

O podcasting revolucionou a produção de conteúdo sonoro ao permitir que criadores independentes produzissem programas de qualquer local, com equipamentos relativamente

simples. Diferente do rádio convencional, que transmite para um público geral, os podcasts favorecem a formação de comunidades de ouvintes por meio de assinaturas e interações online, descentralizando o acesso e a distribuição do áudio.

O crescimento dos podcasts foi impulsionado pelo avanço das tecnologias digitais e pelo aumento do consumo sob demanda. Durante a pandemia de COVID-19, a audiência de podcasts cresceu consideravelmente, já que o formato possibilitava a produção e o consumo de conteúdo sem a necessidade de interação presencial (Vicente, 2018). Esse crescimento pode ser observado na pesquisa realizada pelo Grupo Globo (2020), que analisa a evolução do público de podcasts no Brasil entre 2019 e 2020.

Figura 2: Gráfico - Aumento de ouvintes de Podcast no Brasil



Fonte: Grupo Globo (2020)

Os dados evidenciam um aumento expressivo na audiência de podcasts. Em 2019, aproximadamente 21 milhões de brasileiros com mais de 16 anos ouviram podcasts, representando 13% da população. Em 2020, esse número subiu para 28 milhões, o equivalente a 17% dos brasileiros na mesma faixa etária. Esse crescimento de sete milhões de ouvintes em apenas um ano demonstra a declaração do podcast como um dos formatos de mídia sonora

mais relevantes da atualidade. Além disso, a pesquisa revela que 57% dos ouvintes passaram a consumir podcasts durante a pandemia, enquanto 43% já escutavam antes desse período. Dentre aqueles que já tinham o hábito de ouvir, 31% passaram a consumir ainda mais esse tipo de conteúdo. Esses dados reforçam a ideia de que o contexto pandêmico acelerou a adesão ao formato, consolidando-o como uma alternativa viável e atrativa dentro do ecossistema de mídia digital.

2.3. O crescimento dos podcasts ficcionais

Ancorado no crescimento do número de ouvintes e na produção de podcasts, observamos, nos últimos anos, um dos segmentos que mais se destacou nesse universo: as audioficções. Conforme ilustrado na Figura 3, que apresenta dados de uma pesquisa realizada pela *Nielsen Scarborough Podcast Buying Power* entre maio de 2019 e maio de 2023 nos Estados Unidos, o gênero de Ficção registrou um crescimento expressivo de 93% no número de ouvintes, consolidando-se como um dos segmentos que mais evoluiu nesse período. No Brasil, até o momento da elaboração deste TCC, ainda não há pesquisas específicas sobre gênero e crescimento, como as apresentadas naquele estudo. No entanto, embora a pesquisa tenha sido conduzida nos Estados Unidos, é possível refletir sobre esse crescimento e seu impacto, que, de certa forma, também se reflete no cenário brasileiro.

Figura 3: Gráfico - Aumento de ouvintes no mundo por gênero de podcast de 2019 a 2023

Podcast genres	May 2019 listeners	May 2023 listeners (ranked)	Change
Comedy	17,056,049	28,037,106	64%
News	17,677,927	23,858,773	35%
Society & Culture	14,326,250	23,376,726	63%
True Crime	10,382,429	20,794,054	100%
Education	11,677,723	19,281,312	65%
History	10,033,558	16,537,133	65%
Health & Fitness	11,522,093	16,368,138	42%
Music	10,414,129	15,095,480	45%
Business	10,513,916	14,904,118	42%
TV & Film	8,543,399	14,419,471	69%
Science	9,332,117	14,358,399	54%
Sports	10,132,777	14,076,280	39%
Arts	8,395,985	12,866,196	53%
Religion & Spirituality	7,705,373	11,680,003	52%
Leisure	7,179,375	11,462,748	60%
Fiction	5,628,916	10,863,847	93%
Technology	7,599,887	9,620,221	27%
Kids & Family	4,115,155	6,033,898	47%

Source: Nielsen Scarborough Podcast Buying Power, R2 2018 (May 2019) to R2 2022 (May 2023), Adults 18+

Fonte: Nielsen Scarborough Podcast Buying Power (2023)

A pesquisa também destaca que os gêneros de Comédia, Notícias, e Sociedade e Cultura continuam sendo os mais populares. No entanto, nos últimos cinco anos, a audiência de True Crime dobrou, enquanto a de Ficção quase alcançou o mesmo crescimento. Para os rumos deste TCC, é importante salientar o expressivo aumento de 93% nos podcasts ficcionais. Esse crescimento não apenas reflete uma maior demanda por narrativas sonoras imersivas, mas também indica uma expansão no mercado de trabalho para atores, roteiristas, diretores e produtores especializados nesse formato. Diferentemente dos podcasts jornalísticos e de entrevistas, que dependem fortemente de conteúdo informativo, os podcasts ficcionais exigem uma construção narrativa mais elaborada, com roteiros estruturados, design sonoro refinado e, sobretudo, atuações convincentes.

A ascensão das audioficções reconfigura a relação entre atores e mídias sonoras, abrindo novas possibilidades de trabalho para profissionais que antes estavam majoritariamente vinculados ao teatro, ao cinema e à televisão. A voz passa a ser o principal instrumento de atuação, exigindo técnicas específicas para transmitir emoções e desenvolver personagens sem o apoio visual. Com o aumento da demanda por produções ficcionais em áudio, torna-se fundamental compreender os desafios e particularidades da atuação vocal, bem como o papel do diretor na condução dessas performances.

2.4. Radionovelas no Brasil: a voz como protagonista da atuação

As radionovelas foram um dos formatos mais marcantes da era de ouro do rádio brasileiro, consolidando-se entre as décadas de 1940 e 1960 como um dos principais produtos de entretenimento da época. Inspiradas nos folhetins e novelas latino-americanas, elas não apenas moldaram a cultura popular, mas também estabeleceram um novo paradigma para a atuação e a direção de atores no contexto sonoro. Segundo Cristina Brandão e Guilherme Moreira:

As radionovelas permitiram o desenvolvimento de um modelo específico de interpretação, no qual os atores precisavam expressar emoções intensas apenas com o uso da voz, compensando a ausência de elementos visuais com variações tonais e dicção precisa”. (Brandão e Fernandes, 2014, p. 119).

A construção dos personagens nas radionovelas exigia um domínio técnico diferenciado por parte dos atores, uma vez que toda a expressividade precisava ser transmitida exclusivamente pela voz. Como destacam Da Cunha e Haussen (2003, p. 56), “a ausência de suporte visual fazia com que os intérpretes desenvolvessem técnicas vocais apuradas, utilizando variações tonais e pausas estratégicas para enfatizar emoções e criar identificação com o público”. Dessa forma, elementos como ritmo, dicção e entonação tornaram-se fundamentais para a caracterização dos personagens.

A direção de atores no rádio também apresentava particularidades. Diferente do teatro e do cinema, onde expressões faciais e corporais auxiliam na transmissão das emoções, a radionovela exigia que diretores e intérpretes priorizassem o peso dramático das palavras e a ambientação sonora. De acordo com Brandão e Fernandes (2014, p. 130), “os diretores tinham a função de equilibrar a intensidade da performance dos atores, evitando tanto exageros melodramáticos quanto interpretações monótonas que poderiam comprometer a imersão do ouvinte”. Nesse contexto, a sonoplastia e os efeitos sonoros eram essenciais para complementar a atuação, criando atmosferas que situavam os personagens no espaço e no tempo narrativo.

Com a chegada da televisão e a popularização das telenovelas, as radionovelas perderam espaço na mídia, mas sua influência na dramaturgia vocal permanece relevante. Em um contexto contemporâneo, a ascensão dos podcasts narrativos e das audioficções resgata diversas técnicas do rádio, demonstrando a atemporalidade desse formato. Vale ressaltar que as técnicas das radionovelas foram fundamentais para atores que passaram por esse meio. Diversos artistas consagrados do teatro e da televisão iniciaram suas trajetórias no rádio,

como Paulo Autran, Fernanda Montenegro e Tônia Carrero. A prática radiofônica foi essencial para que esses profissionais desenvolvessem pleno domínio da voz, da articulação e da dicção, qualidades indispensáveis para interpretar textos densos e, por vezes, difíceis de serem ditos. Essa formação lhes proporcionou uma habilidade técnica diferenciada, permitindo-lhes transitar com excelência por linguagens diversas, do palco à tela. O uso constante da voz como principal instrumento fez com que muitos desses intérpretes se destacassem posteriormente em outras mídias, justamente por sua precisão vocal e expressividade, características cada vez mais valorizadas nas produções sonoras atuais. Como aponta Vicente (2022, p. 12), “a experiência do rádio na dramaturgia vocal segue como referência para as novas mídias sonoras, que buscam explorar a voz como principal ferramenta de envolvimento do público”.

Assim, as radionovelas não apenas consolidaram um estilo específico de interpretação, mas também serviram de base para novas formas de contar histórias no ambiente digital. A fusão entre tradição e inovação na atuação sonora continua a evoluir, adaptando-se às transformações tecnológicas e às novas demandas do público.

3. O papel da direção de atores na construção da performance

A necessidade de comunicação sempre impulsionou o ser humano a buscar novas formas de expressão. Ao longo da história, a performance tornou-se um dos principais meios para transmitir histórias, valores e emoções. Desde os primeiros registros da comunicação humana, a expressão corporal e vocal já eram utilizadas como ferramentas para representar tanto a realidade quanto o imaginário.

No teatro grego, por exemplo, os atores incorporavam deuses, heróis e figuras humanas, utilizando máscaras e gestos marcantes para contar histórias que, muitas vezes, funcionavam como lições morais ou instrumentos de controle social. Narrativas sobre punições divinas reforçavam o temor à desobediência, ajudando a consolidar normas religiosas e políticas. Por outro lado, algumas peças destacavam heróis que desafiavam os deuses e triunfavam, abrindo espaço para questionamentos sobre autoridade e destino. Assim, o teatro não apenas entretinha, mas também incentivava a reflexão crítica sobre o poder e a sociedade.

Com o passar dos séculos, a atuação evoluiu juntamente com a sociedade, adaptando-se a novos formatos e linguagens. A direção de atores, enquanto uma prática

sistemática com foco na atuação, surge no século XX com o desenvolvimento do teatro moderno e do cinema, que tentavam, dentro de suas especificidades de linguagem, mostrar produtos que representassem ao máximo a realidade vivida como ela é. Tornou-se, assim, um elemento essencial nesse processo, garantindo que a interpretação estivesse alinhada à proposta narrativa da obra. O diretor atua como um mediador entre o roteiro e o ator, auxiliando na transformação das palavras em emoções autênticas e cenas envolventes.

Com a evolução das mídias, a atuação passou por transformações significativas. No cinema, por exemplo, a proximidade da câmera exigiu dos atores maior controle sobre a sutileza e a naturalidade, já que cada expressão e gesto são captados em detalhes. A televisão, por sua vez, com seu ritmo acelerado de produção, demandou técnicas que equilibrassem intensidade interpretativa e eficiência. No rádio, a atuação exigia um domínio expressivo ainda maior, pois os atores precisavam transmitir a história, as emoções e os sentimentos apenas por meio da voz, sem o suporte da imagem. No Brasil, como vimos no capítulo anterior, o rádio e a teledramaturgia tornaram-se pilares da cultura popular, influenciando gerações e construindo referências que fazem parte do imaginário coletivo.

Mais do que orientar performances e dizer ao ator o que deve ou não fazer, a direção de atores busca explorar as camadas psicológicas e emocionais de cada personagem, proporcionando interpretações mais profundas e realistas. Para isso, diferentes métodos foram desenvolvidos ao longo do tempo, combinando abordagens tradicionais e contemporâneas conforme as necessidades de cada produção. Esse trabalho não apenas enriquece a qualidade artística, mas também fortalece a conexão entre a obra e o público, tornando a experiência narrativa mais imersiva e impactante.

3.1 - O sistema Stanislavski: fundamentos e principais técnicas

A construção de um personagem e a direção de atores exigem técnicas e métodos que possibilitem uma atuação autêntica e envolvente. Para a elaboração deste TCC, foi de suma importância compreender abordagens que aprofundam a interpretação e permitem aos atores vivenciar plenamente seus personagens. Entre os diversos métodos existentes, o sistema desenvolvido por Constantin Stanislavski tornou-se uma das principais referências, influenciando gerações de artistas e diretores.

Stanislavski dedicou sua vida ao estudo da atuação, buscando tornar a performance mais orgânica e verdadeira. Para esta pesquisa, o livro *A Preparação do Ator* (1936) foi a

principal referência, pois apresenta um sistema que não apenas busca formar intérpretes comprometidos com a verdade cênica, mas também orienta a construção de personagens de maneira profunda e estruturada, explorando a psicologia, a emoção e a fisicalidade na atuação.

Na obra, ele desenvolve um conjunto de princípios que levam o ator a construir seu personagem de forma progressiva, partindo da vivência interna para uma interpretação mais autêntica. Entre os diversos conceitos abordados, destaca-se o superobjetivo, um dos métodos aplicados na preparação dos atores para a audiodicção. Esse princípio estrutura a jornada do personagem dentro da narrativa, definindo suas motivações e direcionando suas ações. O autor explica essa ideia da seguinte maneira:

Para desenvolver continuidade numa cena, o ator deveria achar o superobjetivo do personagem: o que é, acima de tudo, que o personagem procura, deseja, durante o curso da peça? Qual é a força motriz do personagem? Se ele pode estabelecer uma meta para a qual seu personagem se dirige, que deve lutar para alcançar, esta será para o artista um objetivo geral, uma linha sobre a qual dirigirá seu personagem." (Stanislavski, 2007, p. 17)

Esse conceito de superobjetivo visa dar mais coerência à atuação, garantindo que cada escolha do personagem tenha um propósito claro para que ele consiga suceder o seu propósito final. No contexto desta pesquisa, esse conceito foi essencial para a direção dos atores no podcast, pois ajudou a definir não apenas os desejos e objetivos de cada personagem, mas também os desafios e conflitos que enfrentaram ao longo da narrativa. Esse processo possibilitou trilhar caminhos na busca de interpretações mais profundas e envolventes, fortalecendo a conexão entre os intérpretes e suas histórias.

Outra técnica estudada e utilizada para fundamentar a aplicação na etapa prática deste trabalho é a memória afetiva, um dos elementos centrais do sistema de Stanislavski. Esse método parte do princípio de que as emoções estão profundamente ligadas às experiências vividas e autênticas, permitindo que o ator resgate sentimentos que já experimentou durante sua vida e os utilize em cena, incorporando-os ao seu personagem.

Stanislavski defendia que a interpretação não deveria ser forçada, mas construída a partir de vivências internas do próprio ator, transformando lembranças pessoais em material artístico. Ele afirma que "para despertar emoções verdadeiras em cena, o ator deve recorrer às suas próprias experiências e memórias, revivendo-as de maneira sincera e espontânea" (Stanislavski, 2007, p. 56).

A aplicação desse método segue etapas específicas. Primeiro, o ator deve identificar uma memória pessoal que tenha provocado uma emoção intensa, como alegria, tristeza ou medo. Em seguida, deve resgatar detalhes sensoriais dessa experiência, como cheiros, sons, texturas e imagens, para tornar a lembrança mais vívida. Assim, ao acessar essa emoção, ele pode transpor o sentimento para a cena, adaptando-o às circunstâncias do personagem.

Esse processo, quando bem conduzido, proporciona uma interpretação mais genuína e envolvente. No contexto desta pesquisa, a memória afetiva é um conceito essencial para a preparação dos atores, e foi usado como fundamento para acessar o lado mais emocional e psicológico dos atores e assim permitindo que cada um acessasse emoções compatíveis com as vivências de seus personagens. Embora os atores não tivessem necessariamente vivenciado essas experiências diretamente, eles puderam recorrer a observações de pessoas ao seu redor que passaram por relacionamentos tóxicos. Esse aprofundamento emocional, mesmo que aplicado brevemente, contribuiu para tornar a interpretação mais vívida e autêntica.

3.2 - A técnica Chubbuck: substituição e objetivo geral

A construção de um personagem exige do ator não apenas técnica, mas também um profundo envolvimento emocional e psicológico. Para além do sistema desenvolvido por Stanislavski, diversas abordagens foram criadas ao longo do tempo para aprimorar a atuação e permitir interpretações mais autênticas. Entre essas metodologias, destaca-se a Técnica Chubbuck, desenvolvida por Ivana Chubbuck, que segue uma linha de pensamento semelhante à de Stanislavski no que diz respeito à importância do envolvimento emocional do ator e ao uso de suas experiências pessoais como fonte de motivação para a cena.

No livro *O Poder do Ator* (2018), Chubbuck propõe um método estruturado em 12 etapas, que combina princípios da psicologia comportamental com técnicas de grandes mestres da atuação, como Constantin Stanislavski, Sanford Meisner e Uta Hagen. A autora enfatiza que a técnica não se limita ao aprendizado das práticas de atuação tradicionais, mas sim à aplicação de teorias psicológicas e comportamentais para criar performances autênticas. Conforme Chubbuck destaca: "A técnica não se trata apenas de aprender a atuar, mas de recorrer a teorias psicológicas e comportamentais para criar performances autênticas." (Chubbuck, 2018, p. 60).

Entre os elementos fundamentais desta abordagem, destacam-se os conceitos de *Substituição* e *Objetivo Geral*, que se assemelham aos princípios de Stanislavski, como o superobjetivo. A *Substituição* consiste no uso de experiências pessoais do ator para estabelecer uma conexão emocional genuína com o personagem. A partir dessa técnica, o ator substitui as circunstâncias da história por momentos reais vividos por ele, permitindo um acesso autêntico às emoções necessárias para a cena. Diferente da memória afetiva de Stanislavski, que se baseia na vivência de sentimentos passados, a Substituição visa impulsionar as ações concretas do personagem, tornando a emoção um meio para um fim narrativo.

Além disso, Chubbuck enfatiza a importância de se compreender o Objetivo Geral do personagem, que corresponde ao conceito de superobjetivo de Stanislavski. Esse objetivo central representa o desejo ou necessidade mais profunda do personagem ao longo da história, orientando suas escolhas e ações. Sobre esse ponto, a autora observa: "A Técnica Chubbuck oferece ao ator ferramentas para acessar emoções profundas e verdadeiras, resultando em performances impactantes." (Chubbuck, 2018, p. 105).

Ao compreender e aplicar essas técnicas, Ivana espera que os atores consigam construir interpretações mais coesas e envolventes. No contexto desta pesquisa, a Técnica Chubbuck foi utilizada, assim como o sistema de Stanislávski, como base teórica e fundamentação, na tentativa de aprimorar a atuação dos atores. Isso permitiu que cada intérprete explorasse suas experiências pessoais de forma direcionada, fortalecendo a conexão entre o personagem e a narrativa sonora, e ajudando a alcançar o objetivo do podcast.

3.3 - Harold Guskin: Como parar de atuar

Em contraste com as abordagens de direção de atores e técnicas que vimos anteriormente, que focam na exploração profunda das emoções e experiências do personagem, Harold Guskin, em *Como Parar de Atuar* (2010), propõe uma abordagem que desafia essas convenções. Guskin acredita que a chave para uma atuação autêntica não está em um processo de construção e preparação rigorosa, mas, sim, em um movimento oposto: "A chave para uma atuação autêntica é parar de tentar atuar e começar a viver verdadeiramente as circunstâncias do personagem" (Guskin, 2010, p. 20). Para ele, a verdadeira performance surge quando o ator abandona o esforço consciente de criar um personagem e, ao invés disso, permite que as palavras e as emoções fluam de forma orgânica.

A principal proposta de Guskin, que foi também um dos exercícios propostos para os atores, gira em torno da ideia de "tirar o texto do papel", uma metáfora que ilustra o processo de internalizar o texto de forma natural e intuitiva. Segundo o autor, a preparação excessiva e o estudo mecânico do roteiro podem resultar em performances rígidas e artificiais. Em vez disso, Guskin sugere que o ator deve explorar o texto de maneira mais espontânea, permitindo que as palavras ganhem vida através de suas próprias experiências: "Em vez de decorar mecanicamente, o ator deve explorar o texto de forma intuitiva, permitindo que as palavras ganhem vida através de suas próprias experiências" (Guskin, 2010, p. 35).

Esse conceito se distancia da ênfase nas emoções pessoais e no uso da memória afetiva, características de métodos como os de Stanislavski e Chubbuck. Para Guskin, a preparação deve ser uma "despreocupação" em relação ao esforço de atuar, permitindo ao ator responder de maneira genuína ao momento presente. Como ele coloca: "A preparação excessiva pode levar à rigidez; é essencial manter a espontaneidade e a abertura para o momento presente" (Guskin, 2010, p. 50). Ao abraçar a imperfeição e a vulnerabilidade, o ator permite que a performance seja um reflexo de sua verdade interna, sem forçar uma narrativa ou emoção específicas.

No processo de preparação para o podcast em questão, as ideias de Guskin foram incorporadas para contrabalançar as abordagens mais tradicionais, que focam no envolvimento emocional profundo com o personagem. Ao permitir que os atores desconectarem um pouco das técnicas ligadas ao emocional e a construção de uma identidade do personagem, buscou-se um espaço para a espontaneidade e a naturalidade nas falas e interações.

Além disso, outro ponto que Guskin destaca em seu livro é a importância de confiar no próprio instinto do ator. Em um ambiente de performance, isso permite respostas imediatas e sinceras às situações improvisadas que surgem: "Confiar no próprio instinto permite que o ator responda de maneira genuína às situações improvisadas que surgem em cena" (Guskin, 2010, p. 65). Essa abordagem mais intuitiva e espontânea, sem o peso da preparação excessiva, leva à autenticidade. Guskin sugere que a preparação excessiva pode gerar uma rigidez no ator, bloqueando a fluidez necessária para que a cena se desenvolva naturalmente: "A preparação excessiva pode levar à rigidez; é essencial manter a espontaneidade e a abertura para o momento presente" (Guskin, 2010, p. 50).

Para Guskin, a atuação genuína surge quando o ator se permite ser vulnerável e aberto às emoções que emergem durante a performance. Como ele afirma: "A verdadeira conexão

com o personagem surge quando o ator se permite ser vulnerável e aberto às emoções que emergem durante a performance" (Guskin, 2010, p. 80).

A crescente digitalização das produções artísticas amplia a necessidade de técnicas que favoreçam a adaptação dos atores a novas linguagens e formatos narrativos. No caso das produções sonoras, essa demanda se intensifica, pois a voz se torna o principal veículo expressivo, resgatando elementos das radionovelas e, ao mesmo tempo, incorporando novas possibilidades de imersão. A proposta de Guskin dialoga diretamente com esse cenário, pois ao valorizar a espontaneidade e a conexão genuína com o texto, permite que a atuação no ambiente digital ocorra de maneira mais fluida e orgânica.

Atualmente, a expansão das plataformas digitais trouxe novos desafios e possibilidades. Webséries, produções independentes e podcasts ampliaram as formas de atuação, exigindo dos intérpretes uma adaptação constante a diferentes formatos e estilos narrativos. No caso dos podcasts e produções sonoras, a voz tornou-se o principal recurso expressivo, resgatando a tradição das radionovelas e explorando novas formas de envolvimento do público.

A atuação, em todas as suas vertentes, continua a se reinventar. Seja no teatro, no cinema, na televisão ou nas mídias digitais, o compromisso com a verdade emocional e a força da narrativa permanece como a essência do trabalho do ator. A direção, por sua vez, segue sendo a ponte que possibilita que essas histórias sejam contadas da maneira mais impactante possível.

4. Voz, som e *viewpoints* vocais

Na construção de uma obra sonora como um podcast ficcional, o sentido da história e a imersão na narrativa não se limitam à interpretação dos atores nem ao uso da fala. Ouvir um conteúdo sonoro envolve mais do que apenas captar as palavras: é preciso compreender a relação entre som, voz e intenção. O som, enquanto elemento estruturante da narrativa, contribui para a criação de atmosferas, marcações temporais e sensações espaciais, atuando como uma força dramática. A voz, por sua vez, carrega não apenas conteúdo verbal, mas também expressividade, subjetividade e presença. Em formatos como o podcast, nos quais a imagem está ausente, a qualidade da voz torna-se fundamental para a construção de sentido e emoção. Como destaca Vicente (2018, p. 45), o som atua como componente estrutural da dramaturgia, organizando o tempo e o espaço narrativos por meio de ritmos, silêncios, pausas e camadas de ambiência. Nesse contexto, a voz emerge não apenas como instrumento de fala, mas como corpo expressivo, capaz de transmitir emoção, presença e intenção dramática. Como afirma Cristina Brandão, “a voz na dramaturgia radiofônica é a principal forma de atuação, sendo o veículo da emoção, da intenção e da subjetividade do personagem” (Brandão; Fernandes, 2014, p. 124).

Este capítulo explora a relação entre voz, som e o conceito de *Viewpoints Vocais*, propondo uma reflexão crítica sobre sua interação. Inicialmente, aborda a voz como matéria expressiva e subjetiva, destacando seu caráter singular na construção de significados. Em seguida, examina o papel do som e da paisagem sonora como elementos estruturantes da narrativa, analisando sua função composicional. Por fim, apresenta os *Viewpoints vocais* como dispositivo analítico, que será usado na investigação proposta. O objetivo com essa divisão é compreender como a clonagem de voz pode afetar não apenas a identidade sonora, mas também a intenção, a emoção e a autenticidade performativa nos podcasts de audioficção.

Para realizar a análise comparativa do objeto de estudo, os dois episódios iniciais do podcast, compreende-se que será preciso partir da comparação entre as vozes humanas dos atores no primeiro episódio e suas versões sintéticas no segundo, buscando contrastes significativos no ritmo, na intenção e na organicidade da fala. Para sistematizar essa escuta crítica, recorre-se à adaptação da metodologia dos *Viewpoints*, proposta originalmente por Anne Bogart e Tina Landau (2005) no campo da atuação, observando como tempo, espaço, ritmo e dinâmica podem ser identificados e analisados na performance sonora das audioficções.

4.1. A voz como identidade

A voz carrega em si camadas de identidade, emoção e contexto. Não se trata apenas de um som qualquer, mas de uma assinatura corporal e subjetiva que expressa identidade, intenção e presença. Como aponta Stanislavski (2003, p. 123), toda ação vocal está imbricada à intenção emocional do personagem, sendo impossível dissociar a técnica de verdade dramática. A entonação, a respiração, o ritmo, os silêncios e até mesmo as falhas carregam informações que ativam a escuta empática do espectador. Essa limitação da inteligência artificial em relação à voz também se intensifica em momentos de maior carga emocional ou de espontaneidade, como as risadas. A risada é um fenômeno involuntário, corporal e emocional, que escapa à lógica da previsibilidade algorítmica. Isso evidencia o que Bogart e Landau (2012) apontam ao discutir o trabalho do ator como algo relacional, enraizado no tempo presente e na escuta do outro. A IA, nesse ponto, permanece fora da dinâmica interpretativa, sem relação real com o momento ou com o ambiente sonoro ao redor criado pelo podcast para dar ambiência para a história.

4.2. Escuta, som e ruídos: para além da fala

Como observa Vicente (2022), a atuação em uma audioficção depende de uma construção orgânica que envolve não apenas a voz, mas sua inserção em um ritmo coletivo de interpretação, ruídos, silêncios e pausas. Essa organicidade ainda não pode ser simulada por tecnologias de clonagem vocal de forma plena, uma vez que tais ferramentas atuam com base em padrões repetitivos e previsibilidade, enquanto o trabalho do ator é, por natureza, vivo, sensível e imprevisível.

Ainda que a fala seja o fio condutor em grande parte das mídias sonoras, ela não atua de forma isolada. Segundo Vicente (2018), o som não deve ser entendido como simples suporte da palavra, mas como parte essencial da narrativa. Ele respira com a cena, acompanha o ritmo dos corpos, ainda que invisíveis, e estrutura o tempo e o espaço da escuta. Nesse sentido, qualquer alteração na naturalidade dessa paisagem sonora, como a inserção de vozes artificiais, afeta diretamente a experiência do ouvinte.

Brandão e Fernandes (2014) ressaltam a importância da organicidade na dramaturgia radiofônica. Cada ruído, hesitação e respiração compõem um quadro expressivo que vai além do texto escrito. Quando a voz clonada remove ou padroniza esses elementos, a narrativa

perde densidade e emoção. A artificialidade ocupa o lugar da expressividade, tornando a escuta mais técnica e menos sensível. Isso quebra a ilusão ficcional, pois o ouvinte identifica uma dissociação entre a expressão vocal e os códigos corporais e emocionais esperados de uma performance humana.

Diante dessa quebra de ritmo, assim como de interferências mecânicas e ruídos típicos de vozes sintetizadas, a escuta passa a demandar uma atenção redobrada. Como reforça Vicente (2022), os elementos que constroem a narrativa sonora vão além do texto verbalizado. Em um contexto onde a voz é simulada, torna-se essencial que o ouvinte perceba não apenas o conteúdo do que é dito, mas também como esse conteúdo é transmitido. Ruídos, pausas, entonações e silêncios ganham ainda mais relevância, sobretudo quando comparados com suas versões automatizadas, que frequentemente falham em reproduzir essas nuances.

A escuta crítica, portanto, torna-se fundamental. Em vez de apenas consumir a narrativa, o ouvinte é convidado a refletir sobre suas camadas. O que é autêntico e o que é simulado? Qual emoção está sendo comunicada e qual se perdeu no processo de clonagem? Essas perguntas levantam não apenas os limites atuais das tecnologias de voz, mas também questões profundas sobre a natureza da atuação e da criação artística sonora.

Esse cenário, no entanto, não fecha as possibilidades. Deixa em aberto a hipótese de que, futuramente, sistemas mais avançados consigam reproduzir não apenas a voz em sua superfície, mas também aspectos mais sutis e complexos, como o ritmo interno da fala, as pausas espontâneas, a pontuação vocal, os silêncios significativos e os ruídos que compõem a performance viva. Ainda permanece o desafio de reproduzir, de forma convincente, o ambiente sonoro completo das audioficções. Tendo em vista que as construções desse estilo de podcast geralmente são compostos por camadas de sons que interagem com a voz dos atores: trilhas, ruídos de fundo, sons ambientes e elementos diegéticos, todos cuidadosamente organizados para criar uma ambiência narrativa coerente. As vozes clonadas, por sua vez, também ainda não conseguem interagir com esse ambiente de forma natural e fluida. Muitas vezes, são inseridas de forma rígida, sem se adaptarem ao espaço sonoro ou responderem ao cenário ao redor. O resultado pode ser uma voz que soa deslocada, que não reage ao som da panela ao fundo, à batida de uma porta ou à trilha tensa que conduz o momento. Assim, a ambientação perde potência narrativa, e a experiência auditiva se enfraquece.

4.3. Viewpoints vocais: um método de escuta

Diante das limitações observadas nas vozes clonadas, especialmente em relação à presença emocional, ao ritmo natural e à integração com o ambiente sonoro, torna-se necessário desenvolver um instrumento analítico capaz de captar essas nuances. Não se trata apenas de evidenciar falhas técnicas, mas de compreender como a voz humana, quando viva, se relaciona com o tempo, o espaço e com o outro. A partir dessa perspectiva, propõe-se uma abordagem que desloca o foco da fala para a escuta como forma de composição. É nesse contexto que se insere a proposta dos *Viewpoints Vocais*, inspirada na metodologia cênica desenvolvida por Anne Bogart e Tina Landau (2005), aqui adaptada à análise de performances vocais em narrativas sonoras.

A técnica dos *Viewpoints* foi inicialmente desenvolvida por Mary Overlie e, posteriormente, expandida por Bogart e Landau como um sistema de criação cênica baseado na improvisação e na percepção do tempo e do espaço. A proposta organiza a atuação a partir de categorias como tempo, forma, gesto, duração e relação com o ambiente. Com isso, os atores são estimulados a compor em cena de modo intuitivo, coletivo e atento ao outro. No entanto, a potência dessa abordagem não se limita ao corpo físico visível. Ao considerar a voz como um corpo sonoro, os *Viewpoints* podem ser transpostos para o campo da escuta, transformando a escuta ativa em ferramenta de análise e criação.

Nesse sentido, os *Viewpoints Vocais* se apresentam como uma adaptação metodológica voltada à escuta crítica e sensível da performance vocal. Eles servem tanto para analisar a atuação quanto para refletir sobre a recepção sonora, permitindo observar aspectos que ultrapassam o texto falado. São seis os *Viewpoints Vocais* propostos neste trabalho: ritmo, timbre, volume e intensidade, articulação, presença e espaço sonoro. O ritmo diz respeito ao fluxo da fala, às pausas, acelerações e cadências que constroem musicalidade e emoção. O timbre refere-se à textura única de cada voz, sua cor sonora e a forma como ela ocupa o espaço acústico. O volume e a intensidade dizem respeito às variações dinâmicas que reforçam a intenção dramática. A articulação observa a clareza, a dicção e a definição da fala. A presença refere-se à sensação da voz como corpo vivo, carregada de intenção e afeto. Já o espaço sonoro avalia a percepção de profundidade, ambiência e espacialidade da voz dentro da mixagem.

A clonagem vocal tende à homogeneização da performance, produzindo falas lineares e com baixa variação expressiva. Faltam hesitações, quebras rítmicas, inflexões emocionais e

pausas significativas que caracterizam a atuação humana. A ausência dessas microvariações limita a expressividade e evidencia os limites da tecnologia no campo da performance. Como afirmam Bogart e Landau, “presença cênica não é apenas estar no palco, mas estar escutando o outro, disponível à ação compartilhada” (Bogart; Landau, 2005, p. 70). Transpondo essa ideia para o campo sonoro, entende-se que uma voz verdadeiramente presente é aquela que escuta antes de falar. Essa escuta se manifesta na entonação, nas pausas e nos silêncios. A voz clonada, por sua vez, não responde ao ambiente emocional da cena, apenas reproduz.

A análise por meio dos Viewpoints Vocais evidencia, portanto, que a clonagem de voz, ainda que tecnicamente impressionante, apresenta uma limitação essencial. Falta-lhe escuta, intenção e presença. A comparação entre a atuação de atores reais e suas versões clonadas não é apenas uma questão técnica, mas também uma reflexão ética sobre o que significa estar presente com a voz. Assim, os Viewpoints Vocais revelam não apenas os limites atuais da clonagem de voz, mas também a centralidade da escuta como ato criativo. Escutar é compor. E é pela escuta que se percebe aquilo que a máquina ainda não pode simular: a presença humana na voz.

5. IA: origem, conceito e contexto atual

A Inteligência Artificial (IA) é um campo da ciência da computação dedicado ao desenvolvimento de sistemas capazes de simular habilidades cognitivas humanas, como aprendizado, raciocínio e tomada de decisão. Sua aplicação se expandiu para diversas áreas, incluindo comunicação, produção audiovisual e geração de conteúdo digital, o que impacta diretamente o campo das mídias sonoras. Segundo Dennis dos Santos Gomes em seu livro denominado “Inteligência Artificial: conceitos e aplicações”:

A Inteligência Artificial pode ser compreendida como um conjunto de técnicas computacionais voltadas para a simulação de comportamentos inteligentes, permitindo que máquinas executem tarefas que tradicionalmente requerem intervenção humana. (Gomes, 2010, p. 235)

Apesar da busca por máquinas que possam replicar habilidades humanas ser antiga, foi apenas no século XX que a IA passou a se consolidar como disciplina científica.

Porém, com os avanços computacionais e o aumento exponencial da coleta e processamento de dados, a IA deixou de ser apenas um campo de estudo teórico para se tornar uma tecnologia amplamente empregada em dispositivos e plataformas acessíveis ao público em geral. O cotidiano de grande parte da população mundial já é mediado por sistemas inteligentes, mesmo que de forma imperceptível. Ferramentas como assistentes virtuais (a Siri da Apple, a Alexa da Amazon e o Google Assistente), recomendações personalizadas em plataformas como Spotify e Netflix, IAs de navegação e geolocalização, como o Waze e o Google Maps, estão entre os exemplos mais difundidos da presença da IA no dia a dia, é que muita das vezes nem são lembradas ou tratadas mais como inteligência artificial, como no caso dos aplicativos de geolocalização por exemplo.

A presença dominante da inteligência artificial se intensifica ainda mais com o surgimento das IAs generativas, capazes de criar conteúdos inéditos a partir de comandos humanos. A chamada IA generativa refere-se a modelos que, em vez de apenas processarem informações, são capazes de gerar textos, imagens, músicas, vozes e vídeos. Entre os exemplos mais populares e mais usados estão as textuais, como o ChatGPT, da empresa americana OpenAI, seu concorrente chinês DeepSeek, o Claude AI, da empresa também americana Anthropic, além do Gemini do Google e do Meta AI do WhatsApp. Dentre as IAs que geram imagens a partir de descrições textuais temos o DALL-E da OpenAI, o Midjourney e o Stable Diffusion, que também geram imagens realistas ou artísticas e o Adobe Firefly,

criada pela adobe e integrada aos softwares de design da mesma. Atualmente, temos também os sistemas de clonagem vocal, como o Descript, Resemble AI, iSpeech, Veed AI e ElevenLabs, capazes de sintetizar vozes humanas com alta fidelidade.

Esses sistemas de IA generativa utilizam modelos de aprendizado profundo (*deep learning*) e são treinados com grandes volumes de dados. Alguns estão conectados à internet e realizam buscas em tempo real, enquanto outros operam com bancos de dados massivos, constantemente alimentados pelos próprios usuários que interagem com essas plataformas. Esse processo de produção e abastecimento de dados não se dá apenas por meio da internet, mas também pelos dados fornecidos diretamente pelos usuários nas plataformas, como mencionado anteriormente. A integração da IA em atividades cotidianas revela uma tendência irreversível: a naturalização da inteligência artificial na vida social. Seja por meio das atividades já citadas como streamings e aplicativos de localização ou até mesmo nas sugestões de compras online, da moderação automática em redes sociais, da tradução de idiomas em tempo real, da produção de conteúdos automatizados ou da simulação de vozes humanas em audiolivros e podcasts, a Inteligência Artificial não apenas complementa, mas, em muitos casos, pode até substituir certas funções humanas.

Nesse contexto, segundo Stuart Russell e Peter Norvig, dois dos principais teóricos contemporâneos da área, a IA pode ser classificada de acordo com sua abordagem funcional, como sistemas que pensam como humanos, que agem como humanos, que pensam racionalmente ou que agem racionalmente, o que demonstra sua amplitude conceitual e aplicação prática (Russell; Norvig, 2013). Isso evidencia que a IA não é uma tecnologia única, mas um ecossistema de soluções cada vez mais interdependentes e sofisticadas. A crescente convivência com essas inteligências nos obriga a repensar a mediação tecnológica nas relações humanas, bem como as implicações culturais, sociais e profissionais desse novo cenário. Como destaca Gomes (2010), a IA não apenas automatiza processos, mas redefine os modos de produção, percepção e interação comunicacional no século XXI.

5.1. Bancos de dados, viés algorítmico e os desafios éticos da clonagem

Essa reconfiguração dos modos de produção e interação mediada por inteligências artificiais não ocorre sem contradições. À medida que esses sistemas se integram ao cotidiano, emergem desafios éticos significativos, especialmente em relação aos dados que os alimentam. Entre as principais preocupações estão as questões morais e éticas associadas à

forma como esses sistemas são desenvolvidos, treinados e utilizados. Um dos pontos mais debatidos é o viés algorítmico, isto é, a tendência das inteligências artificiais de reproduzirem preconceitos e desigualdades já presentes nos dados de origem. Como os modelos são treinados com grandes volumes de informações extraídas de ambientes digitais marcados por desigualdades sociais, raciais, de gênero e regionais, acabam muitas vezes por reforçar ou até amplificar essas assimetrias. Buolamwini (2018, p. 10), pesquisadora do *MIT Media Lab*, afirma que as inteligências artificiais não apenas refletem os dados que recebem, mas também as limitações e os preconceitos de seus desenvolvedores.

No Brasil, essa crítica ganha força nos estudos de Tarcízio Silva (2020), que aponta falhas recorrentes em sistemas de reconhecimento facial e classificação de gênero, sobretudo quando aplicados a rostos negros e femininos, evidenciando um racismo algorítmico estrutural. No contexto desta pesquisa, essa constatação torna-se ainda mais preocupante ao se considerar tecnologias como as inteligências artificiais de clonagem de voz. A reprodução automatizada da fala pode suprimir identidades sonoras minoritárias, privilegiando padrões vocais dominantes como vozes brancas, masculinas e de sotaque neutro, em detrimento da diversidade linguística, regional, cultural e ancestral, tanto no português quanto em outras línguas.

Essa realidade técnica que marginaliza diversidades vocais encontra respaldo teórico no conceito de mediações algorítmicas desenvolvido por Winkes e Longhi (2022). As autoras demonstram que "os algoritmos, indo além da razão instrumental, tornaram-se vetores sociais e constituidores de sentidos, pois tensionam e são tensionados pelas dinâmicas sociais estabelecidas na web" (Winkes; Longhi, 2022, p. 152). O Mapa do Sistema de Mediações Algorítmicas apresentado pelas pesquisadoras oferece uma ferramenta conceitual fundamental para compreender como as tecnologias de clonagem de voz operam dentro de estruturas de poder preexistentes. Segundo o modelo proposto, as mediações algorítmicas envolvem quatro dimensões centrais: institucionalidade, tecnicidades, temporalidades e fluxos, atravessadas por submediações como narrativas, algoritmos, sociabilidade e cidadania. No contexto das IAs de clonagem vocal, a dimensão da institucionalidade revela como as grandes corporações tecnológicas estabelecem padrões dominantes que privilegiam determinados tipos de voz. As tecnicidades evidenciam que as escolhas técnicas aparentemente neutras, como quais datasets usar para treinamento ou quais métricas priorizar, carregam "escolhas humanas e institucionais" (Winkes; Longhi, 2022, p. 157) que perpetuam hierarquias vocais.

As autoras alertam que "quanto mais tempo um sujeito se detiver em um aplicativo ou plataforma digital, mais dados sobre seus hábitos, gostos e comportamentos serão coletados" (Beer, 2009; Couldry; Mejias, 2018; O'neil, 2016; Silveira, 2019; Srnicek, 2017 Apud Winques; Longhi, 2022, P. 158). No caso da clonagem de voz, isso significa que os algoritmos não apenas reproduzem vozes, mas também constroem perfis sonoros baseados em dados que refletem desigualdades sociais, privilegiam vozes que já possuem maior representatividade digital. Particularmente relevante é a observação das autoras de que os algoritmos "fazem classificações por meio de formas e conjuntos de modelos variados: histórico de navegação, informações geográficas, sexo, idade, comportamento, preferências políticas, profissão, rendimento" (Winques; Longhi, 2022, p. 158). Essa classificação automatizada, quando aplicada à reprodução vocal, pode reforçar estereótipos sobre como determinados grupos "devem" soar, marginaliza vozes que não se adequam aos padrões hegemônicos.

O conceito de mediações algorítmicas demonstra, portanto, que as tecnologias de clonagem de voz não são ferramentas neutras de reprodução sonora, mas dispositivos que operam dentro de "regimes de poder e de conhecimento" (Winques; Longhi, 2022, p. 157) específicos. Elas mediam não apenas a relação entre sujeito e tecnologia, mas também as formas como determinadas identidades sonoras são validadas, reproduzidas ou silenciadas no ecossistema digital contemporâneo. A incorporação das inteligências artificiais no cotidiano revela, assim, uma dualidade constitutiva que evidencia sua natureza não neutra ao se configurar como produto de mediações culturais, sociais e institucionais.

5.2. Clonagem de voz: experiência do ouvinte e transformações no mercado sonoro

O avanço da Inteligência Artificial (IA) tem provocado transformações profundas nos modos de produção e recepção de conteúdos midiáticos, especialmente nas mídias sonoras. O desenvolvimento de softwares capazes de clonar vozes humanas com alto grau de fidelidade representa uma dessas inovações tecnológicas com implicações diretas para o mercado de produções sonoras, como rádio, podcasts, publicidade, dublagem e dramaturgia sonora. É preciso, portanto, compreender essas transformações a partir de uma reflexão teórica que aborde o funcionamento das tecnologias de clonagem de voz, seus impactos no campo da sonoro e os debates em torno da percepção de autenticidade vocal e da emoção na recepção dessas vozes sintéticas.

A clonagem de voz, ou *voice cloning*, é uma técnica de síntese vocal baseada em Inteligência Artificial e aprendizado profundo, na qual algoritmos aprendem os padrões específicos de uma voz humana a partir de amostras gravadas para reproduzi-la artificialmente. Essas tecnologias procuram replicar variações como entonação, timbre, ritmo e articulação da fala, criando vozes sintéticas com alto grau de verossimilhança. Elas são empregadas para automatizar produções, acelerar processos criativos e reduzir custos operacionais, mas levantam desafios éticos e jurídicos, especialmente no uso indevido da voz de terceiros.

No Brasil, esse fenômeno ainda está em expansão, mas já começa a influenciar o mercado de produções sonoras. A aplicação da clonagem de voz tem se mostrado particularmente relevante no contexto publicitário e em produções digitais, onde há demanda crescente por personalização e agilidade. No entanto, conforme analisa Rosa (2023, p.29), a ausência de regulamentação específica no país cria um cenário de vulnerabilidade. No âmbito da pesquisa pode-se observar que essa insegurança jurídica e a falta de regulamentação e leis de proteção, impactam diretamente o ramo sonoro para profissionais cuja principal ferramenta de trabalho é a voz. Dubladores, locutores, radialistas e atores de voz podem ter suas identidades vocais replicadas e manipuladas sem consentimento ou compensação, o que gera implicações éticas e legais ainda em debate no meio acadêmico e jurídico.

Para além das questões técnicas e mercadológicas, a clonagem de voz envolve aspectos sensíveis da experiência estética e emocional dos ouvintes. A voz humana é portadora de afeto, presença e intencionalidade, funcionando como um vetor de empatia na escuta (Chion, 1999, p. 52). A substituição por vozes geradas por IA pode modificar a percepção do público, principalmente em contextos ficcionais, documentais ou jornalísticos, onde a presença humana é associada à verdade e à emoção.

Embora a discussão sobre esses impactos seja mais avançada internacionalmente, no Brasil o debate ainda amadurece. Durante a pandemia, a busca por autenticidade sonora manteve-se valorizada entre produtores e ouvintes, mesmo com o uso crescente de tecnologias digitais, indicando o desejo de preservar a “alma” da performance vocal ligada à subjetividade e expressividade humanas (Vicente, 2022, p. 88).

Dessa forma, refletir sobre a inteligência artificial e a clonagem de voz nas mídias sonoras exige mais do que compreender o funcionamento técnico dessas tecnologias. É preciso considerar também como essas transformações reconfiguram sentidos culturais da voz, da escuta e da presença humana no som. Em um cenário onde as fronteiras entre o

natural e o artificial se tornam cada vez mais tênues, o debate ético e estético sobre a autenticidade, a emoção e a representação torna-se essencial. Trata-se, portanto, de pensar não apenas o futuro das vozes que ouvimos, mas também os modos como sentimos e atribuímos valor a elas em uma paisagem sonora marcada pela presença das máquinas.

6. Memorial Descritivo: Por dentro do Jogo

Neste capítulo, serão explorados e detalhados os caminhos percorridos para a execução do projeto proposto. Serão apresentadas as etapas fundamentais que orientaram o desenvolvimento, destacando os métodos empregados, as decisões tomadas ao longo do processo e os fundamentos que embasaram cada escolha. Essa análise busca oferecer uma visão ampla e estruturada sobre o planejamento e a execução do projeto, evidenciando a relação entre os objetivos inicialmente estabelecidos e os resultados alcançados. Além disso, serão enfatizadas as estratégias utilizadas para garantir que os objetivos fossem atingidos de forma eficiente e alinhada às expectativas. Todo o percurso foi construído com base em análises criteriosas, fundamentações teóricas e práticas, assegurando a qualidade e a coerência do trabalho.

6.1. Decisão de jogar: a seleção do tema

Minha escolha pelo curso de Rádio, TV e Internet surgiu da minha grande paixão por rádio e fotografia. O rádio foi o elemento que me conquistou desde muito pequena. Lembro do meu pai cozinhando e ouvindo jogos pelo rádio que ficava perto do fogão. Eu, sentada perto dele, ouvia atentamente, imaginando aquele campo enorme, ouvindo os nomes dos jogadores que eu nem conhecia e criando mentalmente os rostos que corriam pelo campo. Era encantador. Minha mãe também sempre ouvia rádio enquanto arrumava a casa ou lavava roupa. O rádio fez parte da programação dos diversos churrascos e festas da minha infância. Cresci imitando locutores falando sobre peças automotivas, ouvindo jogos e cantando os hits que tocavam.

Junto a essa paixão, veio também a fotografia, especialmente na adolescência. Acredito que ela tenha surgido quando comecei a usar óculos. Com meus altos graus de miopia, o mundo nunca tinha sido tão colorido antes, e eu passei a me encantar por paisagens, prédios, flores, pássaros e pela beleza do cotidiano. Comecei a registrar tudo com a câmera digital da minha avó. Lembro bem de ter escolhido o curso por esses motivos, mas ao longo da trajetória, encontrei pessoas, temas e termos que eu não conhecia, frequentei projetos, aulas e espaços que me despertaram interesses por muitas outras coisas.

Fiquei fascinada pela beleza da direção de arte, especialmente maquiagem e figurino, pela grandiosidade e impacto que a direção sonora tem, e também pelo marketing e design. O interesse por cinema, edição, atuação e direção de atores, que já me atraía pelos anos de teatro

antes de entrar na faculdade, se intensificou. Enfim, acho que todos os caminhos que percorri e as experiências que vivi durante o curso moldaram a pessoa e a profissional que sou hoje.

Quando comecei a pensar no meu trabalho de conclusão de curso, queria unir minhas paixões com temas voltados para o futuro. Queria fazer uma pesquisa sobre algo que me entusiasmasse, algo que me fizesse pensar sobre como as coisas serão daqui para frente. Foi então que surgiu a ideia de criar um podcast, dirigindo atores e usando inteligência artificial para clonar suas vozes em um experimento, comparando essas duas versões. A escolha envolvia tecnologia, atualidade e temas que eu particularmente gosto, além de algo que me motivasse a continuar pesquisando no futuro, como um possível projeto de mestrado. Para mim, o mais difícil foi escolher um único tema entre tantas coisas que aprendi e que gosto de estudar.

Escolha feita? Agora, era hora de planejar essa aventura do trabalho de conclusão. Como amante de jogos, não pude deixar de imaginar esse processo como um grande jogo, com players, fases, peças e possibilidades de escolha. Como em um tabuleiro de jogo, cada decisão de personagem faria a diferença no final. Chegava então o momento de escolher o jogo da vez e começar a jogar.

6.2. Definição da estória: escolhendo o tabuleiro

A escolha do roteiro de ficção que serviria como base para o primeiro e o segundo episódio deste trabalho, assim como para toda a análise, foi feita a partir de um gosto pessoal por crônicas e pela maneira sutil com que alguns autores conseguem contar histórias de forma simultaneamente engraçada e dramática.

O primeiro roteiro selecionado retratava a vida de um casal comum, que normalmente se dava bem, mas estava passando por uma fase morna no relacionamento. Tudo muda quando a mulher chega em casa, mexe nas coisas do casal e encontra uma calcinha que não é dela. O acontecimento gera um grande tumulto, levando-a a confrontar o marido e questioná-lo sobre a traição. Ele, por sua vez, nega insistentemente até, finalmente, admitir que pertence a outra mulher, o que a faz duvidar ainda mais, diante de tantas versões anteriores. No fim, ele assume que a peça, na verdade, é dele.

Esse roteiro parecia perfeito para mim: leve, dramático e engraçado, um contraste ideal para explorar na direção dos atores. Além disso, trazia nuances que eu suspeitava que a

inteligência artificial não conseguiria reproduzir. No entanto, por um problema relacionado à autoria e à impossibilidade de identificar a origem do texto, precisei desistir dessa escolha.

Com o desejo de encontrar um roteiro que permitisse explorar altos e baixos emocionais, parti em busca de uma segunda crônica. Foi então que me deparei com um autor que já conhecia e cuja escrita sempre admirei. A história trazia uma curva dramática bem construída e um plot twist envolvente, o que me levou a escolhê-la para adaptação.

Como meu objetivo era transformá-la em um roteiro ficcional voltado exclusivamente para o áudio, uma audioficção, algumas adaptações foram necessárias. Durante esse processo, percebi que certos elementos precisavam ser atualizados devido ao tempo em que a crônica foi escrita e às mudanças nos costumes e tecnologias. Um exemplo disso foi a substituição de ligações feitas por telefone fixo por interações via WhatsApp, além da atualização de algumas expressões que caíram em desuso, como "só faltava continuar rebolando nas discotecas com as amigas", substituída por referências mais contemporâneas, como sair para festas ou fazer compras online.

6.3. A formação do elenco: as peças do jogo

Para a seleção do elenco, eu buscava atores com experiência prévia, pois o tempo para trabalho e desenvolvimento era curto. Além disso, sabia que conseguiria extrair mais e obter um melhor rendimento com pessoas que já estavam acostumadas aos conceitos de personagem, cena, história e roteiro. A familiaridade com esses elementos permitiria um processo mais dinâmico e produtivo.

Seguindo essa lógica, também priorizei atores que já se conheciam e haviam atuado juntos antes. Essa escolha garantiria um entrosamento natural, facilitando a construção das relações entre os personagens. Por fim, optei por não realizar uma seleção aberta, mas sim escolher diretamente os atores e verificar sua disponibilidade, sem comunicá-los previamente.

Dessa forma, os atores escolhidos foram Emanuelle Serdeira, estudante de Jornalismo na UFJF, ex-integrante do Grupo Divulgação e com experiência em atuação para podcasts, para o papel de Marina e Matheus Omar, bacharel em Publicidade e Propaganda pela UniAcademia e ator do Grupo Divulgação, para o papel de Bruno.

6.4. Ensaios: arrumando as peças no tabuleiro

Com os atores definidos, era hora de colocar em prática os conceitos de direção de atores que aprendi com Stanislavski e Harold. O primeiro encontro ocorreu online, e enviei previamente o roteiro para que eles lessem. Durante a reunião, discutimos suas primeiras impressões sobre a história e os personagens. Como exercício, pedi que me enviassem um relato sobre suas percepções dos personagens apenas com base no texto, antes de lerem minhas anotações sobre características e história de cada um. Dessa forma, evitei que suas interpretações fossem influenciadas diretamente pelas minhas ideias como diretora e autora da adaptação.

No segundo encontro, ainda online, trabalhei diferentes formas de leitura do roteiro. Passamos o texto interpretando com extrema raiva e, depois, com alegria exagerada. O objetivo desse exercício era mostrar que não existe apenas uma maneira de ler o texto, incentivando-os a explorá-lo de forma mais livre e flexível. Para a próxima atividade, pedi que escolhessem uma música que combinasse com seu personagem, para escutarem antes dos ensaios e levarem para discutirmos. A ideia era criar uma relação mais próxima com os personagens, ajudando-os a entender sua essência, forma de andar e gostos musicais.

No terceiro ensaio, que aconteceu presencialmente, iniciamos ouvindo e analisando as músicas escolhidas pelos atores. Discutimos as emoções que as canções evocavam e como elas poderiam influenciar suas interpretações. Em seguida, realizamos um alongamento e um exercício corporal: cada ator deveria demonstrar como imaginava que seu personagem andaria. Depois, passamos o texto duas vezes enquanto caminhávamos em círculo, uma técnica que auxilia na memorização e fluidez, pois o movimento contínuo ajuda a relaxar o corpo e soltar a fala.

Esse ensaio foi o mais longo e produtivo, repleto de atividades diversificadas. Um dos momentos mais interessantes foi quando propus uma dinâmica com o jogo de Uno. Primeiramente, jogamos como pessoas comuns, para que eu pudesse observar a naturalidade de cada um. Em seguida, pedi que jogassem como se fossem seus personagens. O terceiro jogo trouxe uma camada emocional mais profunda: pedi que interpretassem seus personagens jogando contra seus ex-parceiros, o que resultou em expressões intensas de raiva, impaciência e frustração.

Após uma breve pausa para aliviar a tensão gerada, fizemos a última rodada do jogo. Dessa vez, os atores deveriam jogar como se estivessem na noite do jantar retratado no

roteiro, após a refeição, conversando e interagindo naturalmente como seus personagens. Esse foi o momento mais impressionante do ensaio: os atores conseguiram mergulhar completamente em seus papéis, demonstrando uma compreensão profunda de suas dinâmicas e relações.

Esses exercícios foram essenciais para construir a química entre os atores e explorar as camadas emocionais dos personagens, preparando-os para uma interpretação mais autêntica e envolvente. Depois da primeira leitura, levantamos e passamos o texto mais duas vezes, caminhando pela sala e utilizando o exercício de Harold Guskin, retirado do livro *Como Parar de Atuar* (Guskin, 2012). A ideia era desprender o texto da folha e trazer mais naturalidade para a interpretação.

Em seguida, fizemos um exercício de improvisação. Pedi para que os atores simulassem um encontro inesperado entre seus personagens em um mercado, como se fossem dois estranhos que se esbarram pela primeira vez. O resultado foi espontâneo, porque eles realmente se chocaram na improvisação e deixaram a cena fluir naturalmente. Como já haviam discutido gostos em comum na construção dos personagens, a improvisação acabou seguindo esse caminho. Comentaram sobre risoto e até combinaram um jantar.

No segundo improviso, propus que simulassem uma cena em que os personagens já estavam juntos e, por um descuido, a luz havia sido cortada porque um deles esqueceu de pagar a conta. Eles desenvolveram um diálogo improvisado sobre de quem era a culpa e como resolveriam a situação. Durante a conversa pós-exercício, perceberam que entraram tanto na cena que reagiram de forma diferente do que fariam na vida real. O Bruno, que interpretava Omar e assumiu ter esquecido de pagar, comentou: "Achei que você ia brigar comigo". A Marina, que fazia Manu, respondeu: "Eu também poderia ter esquecido facilmente". Esse momento revelou como a relação dos personagens se estabeleceu de maneira mais leve e madura do que alguns relacionamentos vividos pelos próprios atores.

Depois desse processo, passamos o texto mais três vezes, caminhando pelo espaço. Ao final do ensaio, percebi que eles haviam aprofundado ainda mais a construção dos personagens. Finalizei com uma tarefa para casa: pedi que refletissem sobre seus personagens e anotassem o que haviam descoberto de novo sobre eles.

Na quinta-feira, dia do último ensaio, começamos relendo o texto para entender o que havia sido absorvido do dia anterior. Depois, introduzi o primeiro exercício baseado no conceito de superobjetivo de Constantin Stanislavski, retirado do livro *A Preparação do Ator* (Stanislavski, 2019). Pedi que se acomodarem confortavelmente, fechassem os olhos e

respirassem profundamente. A partir daí, comecei a guiá-los com perguntas para que explorassem os objetivos dos personagens: "O que seu personagem busca de verdade? Quais são os maiores obstáculos que ele enfrenta para alcançar isso?". O superobjetivo é a espinha dorsal do personagem, aquilo que guia todas as suas ações ao longo da história, um desejo maior que vai além dos objetivos pontuais de cada cena. Esse exercício ajudou os atores a compreenderem melhor a motivação dos personagens e a tomarem decisões mais coerentes dentro da narrativa.

Depois, fizemos um exercício de memória afetiva, também baseado no método de Stanislavski. Pedi que relaxassem novamente e pensassem em uma lembrança boa, algo marcante. Fui guiando com perguntas sensoriais: "O que você vê ao seu redor? Há sons nessa lembrança? Há cheiros?". No final, perguntei como se sentiam e se essa lembrança gerava alguma reação física. Depois, repetimos o processo com uma memória não tão boa.

Em seguida, pedi que pensassem em experiências pessoais que se assemelhavam às vividas pelos personagens e como reagiram nessas situações. Depois, questionei como seus personagens agiriam nessas mesmas circunstâncias. O objetivo era conectar as emoções próprias às do personagem, sem que precisassem reviver memórias traumáticas. Desde o início, deixei claro que o exercício não deveria causar desconforto ou despertar gatilhos. Era apenas uma forma de explorar emoções dentro de um espaço seguro de experimentação.

Fizemos uma pausa para relaxar, tomar água e aliviar a tensão. Para encerrar, propus que passássemos o episódio mais uma vez e dei uma última tarefa: pedi que, até o dia da gravação, continuassem refletindo sobre seus personagens e ouvissem novamente a música que haviam escolhido como trilha emocional. Também sugeri que, ao chegarem para a gravação, entrassem na sala já incorporando seus personagens, como se estivessem prestes a subir no palco. Como todos são atores de teatro, achei que essa analogia ajudaria na preparação.

O objetivo não era que tudo saísse perfeito de primeira. Eu já sabia que gravaria mais de uma vez e que eu estaria ali para dirigir eles, para dar suporte de perto. Mas queria que sentissem o texto de um jeito mais orgânico, como algo que vinha de dentro, e não apenas como falas decoradas.

6.5 - Dia da gravação: play - que o jogo comece

Chegou o tão aguardado dia da gravação com os atores. A ansiedade era compartilhada entre todos os envolvidos. Havia uma expectativa visível no ar, pois sabíamos que estávamos prestes a criar algo diferente. Para essa etapa, foi combinado que o encontro ocorreria na sala de rádio da faculdade. No entanto, naquele dia, eu estava debilitada por uma virose, o que me impediu de ter contato próximo com os atores por precaução quanto ao risco de contágio. Ainda assim, consegui conversar com eles antes da gravação e conduzi os preparativos iniciais.

Começamos com os exercícios indicados e já feitos em aula pela coorientadora do TCC, a Marcinha, primeiro um exercício de alongamento corporal, com o objetivo de evitar tensões físicas, permitindo que o corpo estivesse solto, com a musculatura ativada e o fluxo sanguíneo circulando adequadamente. Em seguida, realizamos um aquecimento vocal, fundamental para garantir que as vozes estivessem relaxadas e prontas para a interpretação. Complementamos com um exercício de relaxamento geral, visando alinhar corpo e mente, o que favorece a expressividade e concentração durante a performance.

O roteiro foi gravado integralmente cinco vezes. A cada nova tomada, utilizei os intervalos para aprofundar a direção de atuação. Busquei orientar os atores quanto ao tom de voz, à variação de entonação e aos detalhes que poderiam ser ajustados para alcançar uma performance mais fluida, natural e carregada de emoção. A escuta atenta a cada versão gravada permitiu uma evolução notável na entrega interpretativa.

Ao final da sessão, optamos por gravar separadamente alguns trechos-chave, como uma risada conjunta entre os personagens, a reação de Marina ao som do vidro quebrando e uma tosse seca do personagem Bruno. Esses elementos foram registrados de forma isolada com o intuito de estarem disponíveis na pós-produção, caso fosse necessário reforçar a ambientação sonora ou reconstruir trechos específicos na edição final.

6.6 - Pós-jogo: hora de rever a partida

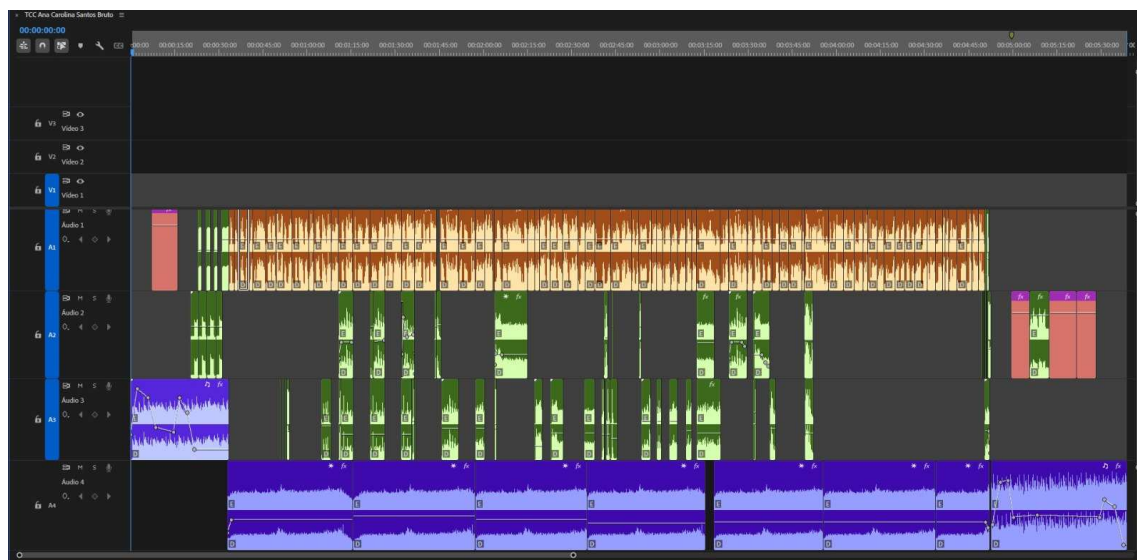
Após a gravação da primeira parte do experimento, iniciou-se a fase de edição. As cinco gravações realizadas foram ouvidas com atenção, e constatou-se que a última versão foi a que melhor atingiu o ponto-chave desejado. O primeiro passo da edição foi separar os áudios por personagem, uma vez que a proposta envolvia o tratamento individual de cada voz, além da inserção de efeitos sonoros. A ideia era simular uma chamada telefônica, em que cada personagem estivesse em sua casa, com ambientes sonoros distintos.

Para a personagem Marina, foi criada uma atmosfera de cozinha, com sons de utensílios, fritura, panela sendo mexida e alimentos em cozimento. Alguns desses sons foram captados pessoalmente em casa (foley), como o som de frango sendo refogado, enquanto outros foram obtidos em bancos de áudio gratuitos disponíveis online, como o YouTube. No caso do personagem Bruno, que já havia realizado a ação de cozinhar em trechos anteriores da narrativa, optou-se por um ambiente mais tranquilo, com ruídos de rua, carros passando, latidos ao fundo e o som de uma porta batendo, simulando o deslocamento pela casa durante a ligação. Assim como em chamadas reais, os sons ambientes de cada personagem eram ativados apenas no momento em que falavam, promovendo uma alternância entre os cenários sonoros conforme o diálogo se desenvolvia.

Em uma narrativa sonora, os efeitos e a trilha desempenham função estrutural, conferindo ritmo, emoção e tridimensionalidade à escuta, sobretudo em um projeto com ambientação voltada para o terror. Por isso, uma das etapas mais aguardadas foi a criação da paisagem sonora. Foram explorados diversos bancos gratuitos para selecionar *background sounds* (BGs) adequados à cena. Alguns efeitos, como o de vidro quebrando, foram criados manualmente por meio da técnica de foley, visto que os disponíveis nos bancos pareciam artificiais ou destoavam da atmosfera pretendida. Essa etapa foi particularmente prazerosa, pois permitiu aplicar técnicas aprendidas nas aulas de direção sonora com o professor Talison.

Com os sons prontos, todo o material foi organizado na timeline do Adobe Premiere, realizando a sincronização entre falas e efeitos. Recorreu-se ao Audacity para a aplicação de efeitos de reverberação e equalização, simulando vozes com a qualidade típica de chamadas telefônicas, com aspecto mais mecânico e filtrado. Foram também ajustados elementos como ruídos de fundo, projeção vocal, frequência de graves e agudos. Após essa etapa, o primeiro episódio da audioficção foi finalizado. A seguinte timeline se obteve no final do processo de edição.

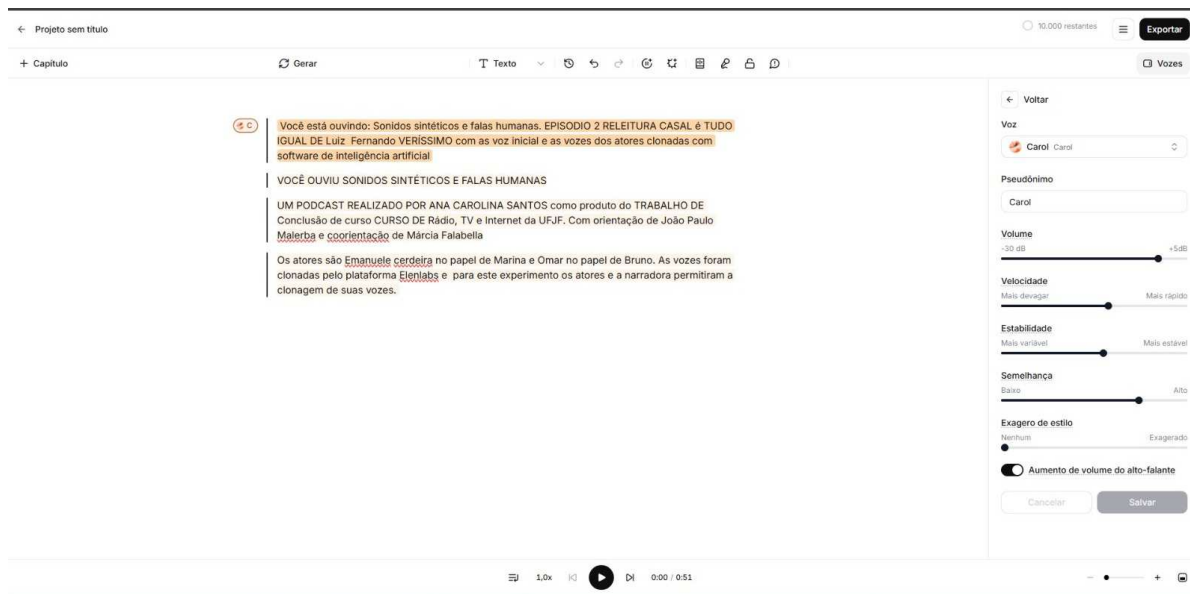
Figura 4: Tela de edição do episódio um no Adobe Premiere



Fonte: Imagem autoral (2025)

Consequentemente, com o episódio 1 pronto dava-se início ao processo de criação do episódio 2. Dessa forma, o passo seguinte consistiu na clonagem das vozes dos atores. Desde o início do processo e da apresentação da ideia do projeto para os atores, eles foram incentivados a enviar áudios de referência. Inicialmente, foi planejado utilizar a plataforma Play.ht, que oferece recursos avançados de clonagem com controle sobre a entrega vocal. Então adquiri um plano pago e os testes preliminares demonstraram bons resultados. No entanto, ao tentar processar as vozes definitivas do episódio, a plataforma apresentou instabilidades frequentes, como erros de carregamento e falhas na renderização. Diante disso, optei por migrar para o Eleven Labs, cuja lógica de funcionamento não era muito parecida com a plataforma que pretendia usar inicialmente, mas que também apresentou ótimos resultados em termos de clonagem. No Eleven Labs, as vozes foram clonadas a partir do envio de materiais dos atores (incluindo ensaios e gravações finais), gerando clones para Marina (Manu), Bruno (Omar) e também para a narração inicial e final (voz da autora). Como é possível observar nas imagens a seguir:

Figura 5: Página da clonagem de vozes inicial e final no Eleven Labs.



Fonte: Imagem autoral (2025).

O Eleven Labs, assim como outras IAs de clonagem de voz, utiliza o texto para reproduzir o que você deseja dizer com a voz que o usuário clonou. Dessa forma, na imagem acima, podemos observar à esquerda o texto usado no início e no final do episódio 1, adaptado para o episódio 2. Nessa parte, notei que o uso de letras maiúsculas e minúsculas impactava um pouco na maneira como a IA lia e reproduzia a voz, então utilizei esse recurso no texto. No canto direito, é possível observar a aba de ajuste da voz, como também é mostrado na imagem abaixo.

Figura 6: Aba de ajuste de voz no Eleven Labs

Fonte: Imagem autoral. (2025)

Na imagem acima, está a parte de edição da voz inicial e final, na qual aparece minha voz, responsável por apresentar e encerrar o episódio. Tanto a minha voz quanto a dos atores, após serem clonadas, podiam ser ajustadas por meio dessa aba. A voz clonada, gerada a partir de áudios da pessoa original, podia ser editada em aspectos como volume, velocidade da fala, estabilidade, semelhança com os áudios originais e estilo exagerado. Essa etapa consistia principalmente em dosar esses elementos para obter um resultado mais natural e próximo da voz original. É importante ressaltar que, embora fosse intuitivo imaginar que aumentar ao máximo o grau de semelhança deixaria a voz mais parecida, na prática isso nem sempre funciona. Quando o nível de semelhança era elevado demais, a voz acabava perdendo em

outros aspectos, como se fosse uma imagem que, ao ser ampliada além do limite, perde nitidez. Apesar da facilidade, a plataforma oferecia menos controle sobre elementos expressivos, como entonação e emoção, permitindo apenas ajustes gerais como gênero, faixa etária e nacionalidade. O software também possui botões que permitem ajustes e edições, conforme ilustrado na imagem abaixo:

Figura 7: Aba de ajuste de voz no Eleven Labs

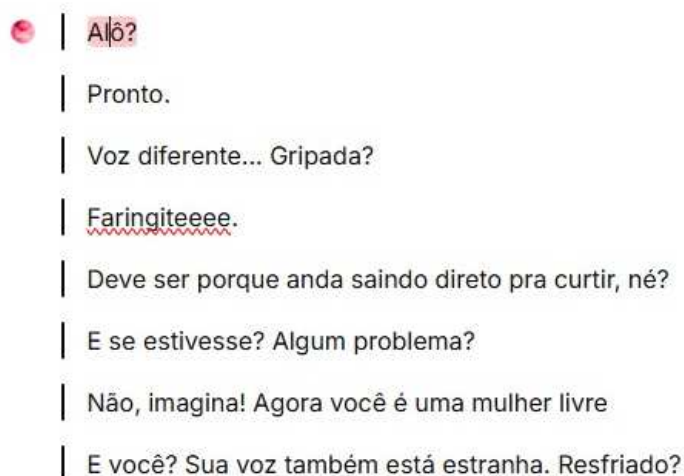


Fonte: Imagem autoral (2025)

Os botões da imagem, em ordem da esquerda para direita, significam: modificador de texto, no qual permite apontar se o texto é normal, se é um título e nomear o parágrafo de texto. Em seguida, tem o relógio, que é o histórico de versões, que permite que você acesse versões passadas, no qual eu usei especialmente quando eu fazia alguma modificação, principalmente na semelhança e na voz em si dos atores, e depois precisava voltar pra versão anterior nos casos em que ficava ruim a atual com a mudança. Então, ter esse histórico de versões foi importante. Em seguida, tem as setinhas para desfazer a ação e refazer. Em seguida, o botão de inserir quebra, ou seja, ajudar o texto. Usei essa função principalmente quando sentia que a IA estava indo muito direto, sem pausas, que com os atores eu trabalhei muito também: as pausas, os silêncios, como forma interpretativa e de gerar intenção de fala e de interpretação. O próximo botão era o de inserir efeito sonoro, que, particularmente, eu quase não usei por conta de inserir depois no Premiere. Então, essa parte queria só as vozes brutas. Usei ele mais na fase de teste pra ver como ficaria. O próximo botão me ajudou bastante, que é o que tem tipo um livrinho. Ele é o editor de pronúncias e nele você consegue adicionar um dicionário, no qual, no caso, eu adicionei um dicionário criado por mim mesmo com a ajuda de IA, do Claude e do Manus IA, com gírias e falas cariocas para ajudar ele a moldar melhor o sotaque da Marina (Manu), e usei também o dicionário de palavras em português. Essa parte me ajudou a alterar várias palavras em inglês que ele colocava, principalmente números. Eu até deixei um dos números propositalmente no podcast com esse intuito, justamente de mostrar essa dificuldade que o software nativo em inglês tem de interpretar o português e não trocar as palavras. Ou seja, no texto deixei claro que era a fala “número 2”, por exemplo, mas na hora de clonar, a IA substituiu por “two”. O próximo botão também foi um dos mais úteis e usados, com esse microfoninho. Ele é o modo ator, no qual

you can record or send an audio saying exactly what you want the sentence to be, that the AI will try to reproduce that intonation. Then, I used it well, adding parts that the actors said in the original podcast with their voices and, in my initial and final parts, I recorded them again, sending fresher audios. And, finally, we have the button with a lock, which is to block a paragraph, which you can use when you write a paragraph, but you don't want it to be read. And, finally, it is to indicate that it is not to generate a determined paragraph. This button is more used in the initial phase, which is to place the text and separate it into paragraphs and speeches. The next image shows a part of the script for episode 2 in ElevenLabs:

Figura 8: Página do ElevenLabs com o texto do episódio 2



○		Aiô?
		Pronto.
		Voz diferente... Gripada?
		<u>Faringiteeee.</u>
		Deve ser porque anda saindo direto pra curtir, né?
		E se estivesse? Algum problema?
		Não, imagina! Agora você é uma mulher livre
		E você? Sua voz também está estranha. Resfriado?

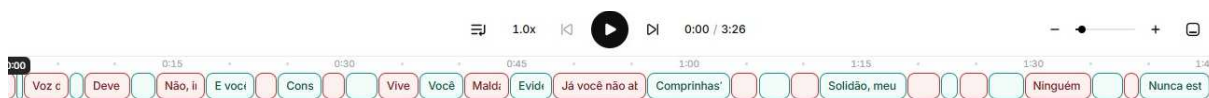
Fonte: Imagem autoral (2025)

This image shows how I used the text in artificial intelligence. I had to separate each sentence of the actors and indicate for the AI who should speak each one. In the image, the first to say "oi" is the character Omar, represented by the pink dot. The next image shows how the voice of each character appeared for me in the corner of the screen.

Figura 9: Página do ElevenLabs com o texto do episódio 2

Fonte: Imagem autoral (2025)

A voz de Omar (Bruno) é marcada com a cor rosa, e a de Emmanuelle (Marina) com a cor verde. Como mostrei anteriormente, conseguiria se quisesse ir adaptando a intensidade da fala e outros elementos, assim como fiz com a minha própria voz, que aparece na abertura e no encerramento dos episódios, apresentando e finalizando a narrativa. Na imagem a seguir, é possível observar também como aparecia a breve linha do tempo que indicava o tempo da gravação.

Figura 10: Sessão de linha do tempo e player do ElevenLabs

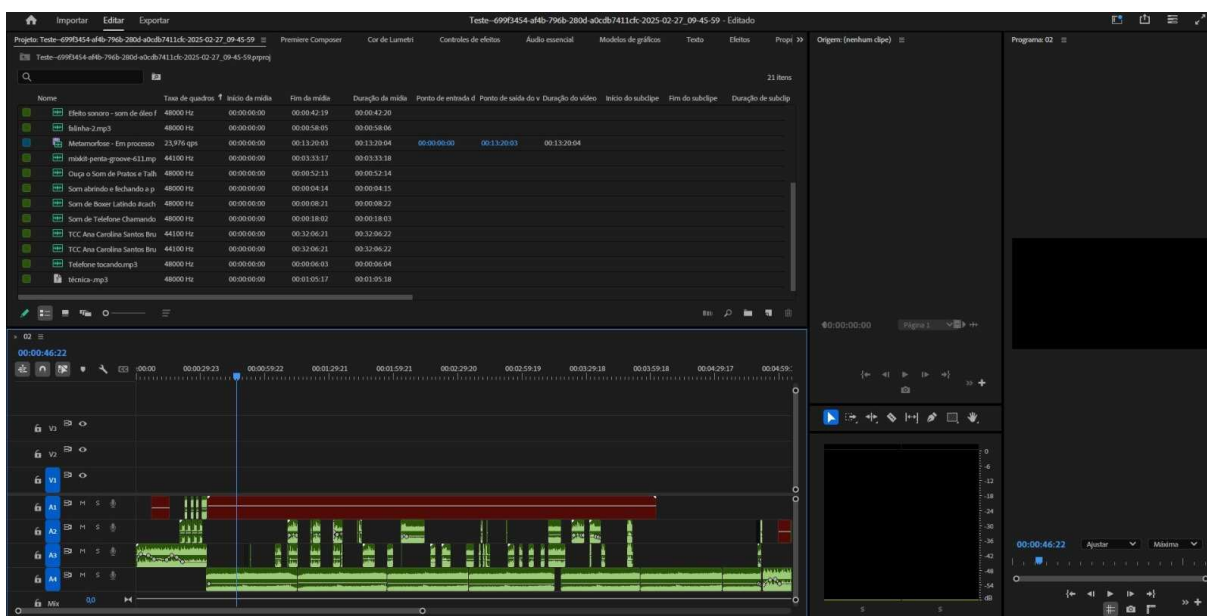
Fonte: Imagem autoral (2025)

Nessa parte eu conseguia pausar, avançar, retroceder, acelerar e reproduzir apenas uma parte selecionada do áudio. E na parte inferior ao player, aparece uma mistura das cores rosa e verde, junto ao início das falas. Essas cores representam o que os atores diziam e quem estava falando, de acordo com a cor associada a cada voz.

Durante a geração dos diálogos, surgiu uma nova dificuldade: não era possível alternar automaticamente entre vozes em um mesmo texto. Foi necessário separar cada fala manualmente, processá-las de forma individual e remontá-las posteriormente no Premiere, como em um quebra-cabeça sonoro. O resultado inicial foi insatisfatório. As vozes soavam artificiais, sem a expressividade dos atores originais. Decidi então, refazer o processo, estudando melhor os recursos oferecidos pela IA. Foram consultados tutoriais, ajustados parâmetros e testadas diferentes formas de input textual para alcançar maior naturalidade nas falas geradas. Reconhecendo as limitações atuais da inteligência artificial em replicar performances humanas com precisão, optou-se por explorar ao máximo os recursos

disponíveis, sem perder de vista a intenção original do roteiro. Com as vozes reprocessadas, a ambientação sonora foi reorganizada e o episódio reeditado. O resultado final, ainda que com limitações, representou o melhor aproveitamento possível dos recursos disponíveis, respeitando as intenções narrativas e sonoras do projeto. A imagem a seguir é da tela do Adobe Premiere após a edição completa do segundo episódio.

Figura 11: Tela de edição do episódio dois no Adobe Premiere



Fonte: Imagem autoral (2025)

6.7 - Análise: Revendo as partidas e comparando os resultados

Com os episódios finalizados, iniciou-se uma das etapas mais cruciais da pesquisa: a análise comparativa entre os episódios gravados com os atores e aqueles realizados por meio da clonagem de voz. O objetivo foi compreender as diferenças de desempenho vocal, expressividade e naturalidade entre as interpretações humanas e os resultados gerados por inteligência artificial. Essa etapa da pesquisa teve início com a escuta atenta dos dois episódios, com foco em marcas de regionalidade, como o sotaque carioca da personagem Marina (interpretada por Emanuele). Desde o início, a proposta era que a atriz mantivesse seu sotaque natural, pois a personagem era originalmente do Rio de Janeiro, e essa informação seria transmitida ao ouvinte por meio da voz, sem a necessidade de uma explicação no

roteiro. A análise revelou que a IA apresenta dificuldades significativas para replicar o "s" puxado, típico do sotaque carioca, enfraquecendo esse traço identitário da personagem.

Outro aspecto observado foi a pontuação. A IA, em diversos momentos do episódio 2, não respeitou vírgulas ou pausas previstas no roteiro. Em situações de exclamação, a voz se tornava artificialmente aguda, com um tom forçado e robótico. A modulação vocal, que acompanha a progressão emocional da narrativa, também se mostrou comprometida. No decorrer da história, as personagens passam de uma tensão inicial para um embate mais direto, com acusações e confrontos, até chegar a um desfecho mais leve e cômico. Embora a IA tente acompanhar essas variações de tom, em trechos específicos como quando Marina acusa Bruno de nunca ter lavado sequer uma louça e ele reage em tom mais alto e defensivo, a resposta gerada pela IA se mostra superficial, sem captar a carga emocional e a nuance da interpretação humana.

Nos episódios analisados, percebe-se que a IA não consegue reproduzir uma risada de forma convincente. O som se torna mecânico, fora de tempo, sem a vibração do corpo que torna a risada algo verdadeiramente humano. A IA também apresentou dificuldade em reproduzir expressões como o "ai" dito por Marina ao ver o copo de sua mão cair no chão ou a risada compartilhada entre os personagens ao final, quando percebem que tudo não passou de um mal-entendido. Nesses momentos, o resultado soou nitidamente robótico e artificial. Mesmo ouvintes que não conheciam os atores ou os episódios originais puderam identificar a ausência de naturalidade.

São seis os Viewpoints Vocais propostos neste trabalho: ritmo, timbre, volume e intensidade, articulação, presença e espaço sonoro. O ritmo diz respeito ao fluxo da fala, às pausas, acelerações e cadências que constroem musicalidade e emoção. O timbre refere-se à textura única de cada voz, sua cor sonora e a forma como ela ocupa o espaço acústico. O volume e a intensidade dizem respeito às variações dinâmicas que reforçam a intenção dramática. A articulação observa a clareza, a dicção e a definição da fala. A presença refere-se à sensação da voz como corpo vivo, carregada de intenção e afeto. Já o espaço sonoro avalia a percepção de profundidade, ambiência e espacialidade da voz dentro da mixagem. Ao aplicar esses pontos de escuta ao episódio 2 do podcast analisado, torna-se evidente que as vozes sintéticas, embora bem construídas tecnicamente, não ativam plenamente todos esses elementos.

Para reforçar a análise, foram realizadas escutas experimentais com ouvintes externos, sem familiaridade prévia com o projeto, os intérpretes ou as tecnologias de IA. Os participantes foram escolhidos de forma a representar diferentes faixas etárias e contextos sociais: uma senhora de 71 anos, uma criança de 11 anos e uma universitária de 23 anos. Nenhuma delas havia escutado o episódio 1 ou sabia que se tratava de vozes clonadas. O objetivo era sair do campo da percepção individual de pesquisadora e ampliar a análise para perspectivas diversas da população. Há uma diferença importante: a estudante de 23 anos, Julia, conhecia a voz original dos atores por já ter contato prévio com eles. Já a senhora de 71 anos e a criança de 11 não sabiam que se tratava de inteligência artificial e também não conheciam as vozes originais. O que mais me chamou atenção nessa entrevista foi a diferença nas reações. Ana Júlia (11 anos) não identificou que era IA. Ivone (71 anos) sentiu algo estranho, mas só levantou a hipótese de ser inteligência artificial após ser questionada sobre essa possibilidade. Ana Julia, mesmo após ser perguntada diretamente, afirmou que o áudio era humano e disse que não imaginava outra possibilidade.

Esse fato levanta uma questão relevante, tratada também na parte teórica da pesquisa: até que ponto as pessoas estão seguras diante de uma ferramenta tão poderosa, capaz de causar engano e confusão com vozes reais? Do meu ponto de vista como pesquisadora em inteligência artificial, o áudio clonado se mostra evidente. Apresenta um som marcado, sem naturalidade e com aspecto robotizado. Mesmo com o uso de todos os recursos disponíveis, com dedicação total, com a combinação de mais de uma IA e com uma pesquisa prévia sobre as ferramentas mais adequadas para clonagem, o resultado não atingiu a qualidade esperada por mim anteriormente. A minha expectativa em relação à máquina era maior.

No entanto, essa entrevista e análise com pessoas sem conhecimento aprofundado sobre inteligência artificial revela um alerta importante. O relato de Júlia ajuda a reforçar essa preocupação. Ela conhece as vozes reais dos atores e observa algo relevante: apesar de notar certa semelhança, percebe uma aproximação mais clara na voz do Omar do que na da Manu. Na minha percepção, a voz da Manu apresenta maior instabilidade e dificuldade para a inteligência artificial. O caso de Renata Lanna, citado no próprio podcast, exemplifica esse problema. A voz gerada ficou bastante próxima da original, como comentaram também Márcia Falabella e Talita Magnolo. Renata, ao contrário, demonstrou surpresa ao descobrir

que sua voz havia sido clonada, mas reconheceu a diferença em detalhes como pausas, vírgulas e entonação. Isso indica que, em certos contextos, conhecer a própria voz ou ouvir alguém com frequência torna a percepção da clonagem mais clara. Para Renata, essa diferença ficou evidente, enquanto para mim, Marcia e Talita, a semelhança se destacou. Isso me leva a um novo questionamento sobre a qualidade do áudio e o tipo de software utilizado. Não consegui identificar qual aplicativo ou ferramenta foi usada na clonagem do caso de Renata, mas, ao comparar com o podcast experimental, acredito que a diferença na qualidade esteja relacionada tanto ao programa e à verba disponível quanto ao estilo de locução. A locução de Renata segue uma linha marcada, com tom político e de convencimento, enquanto no episódio do podcast a atuação dos atores varia bastante, com momentos de calma, humor, ápice e reconciliação. Também considero que a IA pode ter interpretado de forma distinta por lidar com duas vozes diferentes, uma masculina e uma feminina, sendo a feminina marcada por traços regionais do sotaque carioca.

Quanto ao episódio 3, sua elaboração ocorreu de forma fluida e bem estruturada, tanto na etapa de roteiro quanto na edição. Além das conversas com Marcinha e Talita, realizei entrevistas com Renata Lanna, Ana Julia, Julia e Ivone. Desde o início, havia clareza sobre os temas que desejava abordar em cada fala, os pontos que precisava evidenciar, os aspectos a esclarecer e as reflexões que gostaria de provocar. As entrevistas contribuíram diretamente para o desenvolvimento do episódio e também para a análise final do projeto. Cada uma apresentou uma perspectiva específica, desde quem conhece profundamente as vozes dos atores até quem escutava o material sem qualquer referência prévia. Essas escutas possibilitaram identificar diferentes formas de recepção e percepção das vozes geradas por inteligência artificial, confirmando algumas hipóteses técnicas iniciais e, ao mesmo tempo, abrindo espaço para novos questionamentos éticos e estéticos.

Na fase de edição, utilizei o software Adobe Premiere. Selecionei cuidadosamente os áudios e trilhas de fundo que ajudariam a construir uma atmosfera mais leve, favorecendo o envolvimento do ouvinte sem comprometer a clareza e a consistência da narrativa. A intenção era tornar o episódio mais acessível, com um equilíbrio entre o conteúdo analítico e a fluidez sonora. A organização do material se assemelhou a um grande quebra-cabeça, em que cada elemento precisava ocupar seu lugar com precisão para que o resultado final fizesse sentido. Assim como nos demais episódios, a construção exigiu atenção aos detalhes, escolhas técnicas bem fundamentadas e sensibilidade na costura das falas e ambientes sonoros. O

episódio finalizou um ciclo importante, reunindo aspectos técnicos, reflexões críticas e experiências práticas que atravessaram toda a pesquisa.

Dessa forma, o episódio 3 sintetiza os principais caminhos percorridos ao longo da pesquisa, articulando elementos técnicos, experiências de escuta e reflexões críticas sobre a aplicação da inteligência artificial em mídias sonoras. A combinação entre análise e prática tornou possível explorar com mais profundidade os limites e as possibilidades dessa tecnologia, especialmente no que se refere à clonagem de voz e à recepção por parte do público. As entrevistas, as decisões de edição e a montagem final reforçaram a complexidade do tema e evidenciaram o quanto aspectos técnicos e subjetivos se entrelaçam quando se trata de voz, autenticidade e representação.

7 - Considerações finais

Este trabalho teve como principal objetivo investigar a relação entre direção de atores, inteligência artificial sonora e clonagem de voz em produções ficcionais em áudio, por meio de uma abordagem teórico-prática que uniu pesquisa acadêmica e experimentação criativa. Ao longo do desenvolvimento, foi possível perceber como a tecnologia tem reconfigurado práticas de produção, percepção e recepção da voz, afetando diretamente profissionais da comunicação, como atores, dubladores, locutores e criadores de conteúdo sonoro.

A fundamentação teórica foi dividida em quatro eixos principais. O primeiro tratou da evolução dos formatos sonoros, do rádio ao podcast. O segundo abordou a direção de atores e as técnicas de construção vocal. O terceiro focou na expressividade da voz e nos modos de escuta. O quarto discutiu a inteligência artificial e a clonagem de voz, com atenção às questões éticas e profissionais envolvidas. Esses eixos se conectaram ao longo do trabalho e contribuíram para uma análise mais ampla das transformações nas mídias sonoras. A pesquisa dialoga com autores de diferentes áreas, como Vicente, Stanislavski, Chubbuck, Rosa e Gomes, que ofereceram fundamentos teóricos para sustentar a reflexão crítica e a criação do podcast experimental.

No aspecto prático, a criação dos episódios do podcast e as experiências com clonagem vocal serviram como espaço de experimentação para observar os efeitos, desafios e limites da inteligência artificial no campo sonoro. Ficou evidente que, embora a IA consiga replicar características técnicas da voz humana, ela ainda carece de naturalidade e profundidade interpretativa. A ausência de intenção, escuta e presença torna sua atuação artificial, soando, em muitos momentos, mecânica e limitada. Dessa forma, percebe-se que a tecnologia ainda não é capaz de construir personagens, nem de sustentar narrativas com a complexidade emocional e expressiva que o trabalho humano exige.

No mercado sonoro foi possível perceber que ainda é difícil que a IA substitua completamente atores e dubladores. Pois falta-lhe a capacidade de escutar o outro, de reagir ao contexto, de compreender o texto além da superfície. A IA pode repetir falas, mas não interpreta. Ela não é capaz de transmitir emoção, nem estabelecer conexão com o interlocutor.

Ainda dentro do campo prático, as análises e entrevistas reforçaram a urgência de discutir o uso da IA com base em princípios éticos. Questões como a preservação da identidade vocal, o direito à imagem sonora, a possível exclusão de minorias por preconceitos algorítmicos e a valorização do trabalho humano se mostram centrais. A ausência de regulamentação

específica no país cria um cenário de vulnerabilidade jurídica que impacta diretamente profissionais cuja principal ferramenta de trabalho é a voz. Esta insegurança legal se reflete na precarização das condições de trabalho e na dificuldade de estabelecer parâmetros claros para o uso ético da tecnologia no setor audiovisual. Ainda existem lacunas legais quanto ao uso de vozes clonadas, o que representa riscos profissionais e jurídicos. Ao mesmo tempo, a tecnologia oferece caminhos criativos, desde que usada com responsabilidade e consciência crítica.

A pesquisa do TCC em si, evidencia que o som, como linguagem expressiva e identidade pessoal, está no centro de uma tensão entre inovação e autenticidade. Neste contexto, onde as fronteiras entre o natural e o artificial se tornam progressivamente mais difusas, emerge a necessidade de um debate crítico sobre os critérios que definem autenticidade, representação e valor emocional na experiência sonora. Trata-se de repensar não apenas as vozes que habitam nossa paisagem sonora, mas fundamentalmente os processos pelos quais atribuímos significado e legitimidade a essas manifestações vocais em um ambiente crescentemente mediado pela tecnologia. Esse cenário exige escuta crítica tanto de quem produz quanto de quem consome conteúdo sonoro. Este estudo pode colaborar com debates sobre ética, regulação e direitos autorais no uso da inteligência artificial na voz. Também destaca a importância de envolver artistas, técnicos, juristas e pesquisadores na criação de práticas que aliem tecnologia e responsabilidade.

Por fim, é essencial reconhecer que a relação entre voz humana e inteligência artificial requer cuidado e responsabilidade. O futuro das mídias sonoras depende da capacidade de equilibrar inovação técnica com respeito às singularidades humanas, garantindo que as vozes reais ou clonadas sejam instrumentos de expressão legítima e ética. O desafio atual é que toda a sociedade, incluindo produtores, profissionais e público, desenvolva uma compreensão crítica e consciente sobre a presença crescente da inteligência artificial na voz. A adaptação a essa nova realidade passa pelo reconhecimento das possibilidades e dos riscos da clonagem vocal, preservando a autenticidade e o valor da expressão humana. Apenas com essa consciência coletiva será possível garantir um uso ético e responsável, protegendo os direitos e a diversidade nas mídias sonoras.

REFERÊNCIAS

- BEER, David. The social power of algorithms. *Information, Communication & Society*, v. 20, n. 1, p. 1-13, 2017.
- BOGART, Anne; LANDAU, Tina. *The viewpoints book: a practical guide to viewpoints and composition*. New York: Theatre Communications Group, 2005.
- BOGART, Anne; LANDAU, Tina. *The viewpoints book: a practical guide to viewpoints and composition*. 2. ed. New York: Theatre Communications Group, 2012.
- BRANDÃO, Cristina; FERNANDES, Guilherme Moreira. O passado e o presente da dramaturgia radiofônica no Brasil. *Rádio-Leituras*, v. 5, n. 1, p. 118-140, 2014.
- BUOLAMWINI, Joy. Gender shades: intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, v. 81, p. 77-91, 2018.
- CHION, Michel. *The voice in cinema*. New York: Columbia University Press, 1999.
- CHUBBUCK, Ivana. *O poder do ator: a técnica Chubbuck em 12 etapas: do roteiro à interpretação viva, real e dinâmica*. Tradução de Carolina Muniz. São Paulo: Novo Conceito, 2014.
- COULDRY, Nick; MEJIAS, Ulises A. *The costs of connection: how data is colonizing human life and appropriating it for capitalism*. Stanford: Stanford University Press, 2018.
- CUNHA, Magda; HAUSSEN, Doris Fagundes. *Rádio brasileiro: episódios e personagens*. Porto Alegre: EDIPUCRS, 2003.
- DIZARD, Wilson. *A nova mídia: a comunicação de massa na era da informação*. 3. ed. Rio de Janeiro: Jorge Zahar Editor, 2000.
- FOSS, Karen A.; LITTLEJOHN, Stephen W. *Teoria da comunicação*. São Paulo: Cengage Learning, 2009.
- GOMES, Dennis dos Santos. Inteligência artificial: conceitos e aplicações. *Revista Olhar Científico*, v. 1, n. 2, p. 234-246, 2010.
- GONTIJO, Silvana. *O livro de ouro da comunicação*. Rio de Janeiro: Ediouro, 2004.
- GRUPO GLOBO. *Pesquisa de consumo de áudio no Brasil*. Rio de Janeiro: Grupo Globo, 2020. Disponível em: <https://shre.ink/tER1>. Acesso em: 1 ago. 2025
- GUTIERREZ, Daniel. *A história que eu cresci*. Participação de Mel Lisboa e Seu Jorge. Paciente 63, Spotify Studios, 22 jul. 2023. 17 min. Disponível em: <https://open.spotify.com/episode/4JULJQ8MPe0bRrFN3DJr5G>. Acesso em: 30 jul. 2025.

GUSKIN, Harold. Como parar de atuar. Tradução de Diego Monteiro. Rio de Janeiro: Apicuri, 2010.

HAMMERSLEY, Ben. Audible revolution. The Guardian, 2004. Disponível em: <https://www.theguardian.com/media/2004/feb/12/broadcasting.digitalmedia>. Acesso em: 30 jul. 2025.

LITTLEJOHN, Stephen W.; FOSS, Karen A. Teorias da comunicação humana. 9. ed. Porto Alegre: Artmed, 2008.

NAVAS, Débora. Rádio em transição: convergência e reconfiguração do meio. In: BRANDÃO, Cristina; NAVAS, Débora; FERNANDES, Guilherme Moreira (org.). Emergências periféricas em práticas midiáticas. São Paulo: ECA-USP, 2018. p. 59-74.

NIELSEN SCARBOROUGH. **Podcast buying power report 2023**. New York: Nielsen, 2023.

O'NEIL, Cathy. Weapons of math destruction: how big data increases inequality and threatens democracy. New York: Crown, 2016.

PAIVA, Raquel. O espírito comum: comunidade, mídia e globalismo. São Paulo: Paulus, 2006.

ROJAS, Julio. Um corpo na noite. Roteiro de Fábio Yabu e Leonel Caldela. Direção de Gustavo Kurlat. Produção de Alexandre Ottoni e Deive Pazos. Interpretação de Selton Mello. França e o Labirinto, Spotify Studios, 22 ago. 2023. Podcast. Disponível em: <https://open.spotify.com/episode/2BvqwdfLnUT3mq54fDJUXU?si=0f1e6ef5554e49d3>. Acesso em: 30 jul. 2025.

ROSAS, Ana Carolina Gomes. A tutela da voz no mundo da inteligência artificial: aspectos atuais da sua regulamentação no Brasil e na Europa. 2023. Trabalho de Conclusão de Curso (Graduação em Direito) – Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2023.

RUSSELL, Stuart; NORVIG, Peter. Inteligência artificial. 3. ed. Rio de Janeiro: Elsevier, 2013.

SILVA, Tarcízio. Racismo algorítmico: inteligência artificial e discriminação nas redes digitais. São Paulo: Dandara, 2020.

SILVEIRA, Sergio Amadeu da. Economia da intrusão e demo-cracia digital. In: COCCO, Giuseppe; HOPSTEIN, Graciela (org.). As multidões e o império entre globalização da guerra e universalização dos direitos. Rio de Janeiro: DP&A, 2019. p. 363-376.

SRNICEK, Nick. Platform capitalism. Cambridge: Polity Press, 2017.

STANISLAVSKI, Constantin. A preparação do ator. Tradução de Pontes de Paula Lima. Rio de Janeiro: Civilização Brasileira, 2003.

STANISLAVSKI, Constantin. A preparação do ator. São Paulo: Civilização Brasileira, 2019.

VERÍSSIMO, Luís Fernando. Casal é tudo igual.. Porto Alegre: L&PM, 1981.

VICENTE, Eduardo. Audioficção na pandemia: estratégias de produção e busca por uma linguagem sonora atual. *Insólita: Revista Brasileira de Estudos Interdisciplinares do Insólito, da Fantasia e do Imaginário*, v. 2, n. 2, p. 1-15, 2022.

VICENTE, Eduardo. Do rádio ao podcast: as novas práticas de produção e consumo de áudio. In: BRANDÃO, Cristina; NAVAS, Débora; FERNANDES, Guilherme Moreira (org.). *Emergências periféricas em práticas midiáticas*. São Paulo: ECA-USP, 2018. p. 43-58.

WINQUES, Kérley; LONGHI, Raquel Ritter. Mediações algorítmicas: o mapa do sistema como proposta teórico-metodológica. *Revista Famecos*, Porto Alegre, v. 29, n. 1, p. 152-168, jan./abr. 2022.

WINQUES, Kérley; LONGHI, Raquel Ritter. Mediações algorítmicas: o mapa do sistema como proposta teórico-metodológica. *Revista Famecos*, Porto Alegre, v. 29, n. 1, p. 152-168, jan./abr. 2022.

WINQUES, Kérley; LONGHI, Raquel Ritter. Algoritmos e mediações: reflexões sobre comunicação digital. *Estudos em Comunicação*, Covilhã, v. 36, n. 2, p. 155-172, maio/ago. 2023.n. 2, p. 151–172, 2022. DOI: 10.11606/issn.1982-8160.v16i2p151-172. Disponível em: <https://revistas.usp.br/matrizes/article/view/183743>. Acesso em: 1 ago. 2025.

ZAHAR, Jorge. *Comunicação e novas tecnologias*. Rio de Janeiro: Jorge Zahar Editor, 2004.

APÊNDICE 1 - Roteiro da Audioficção



ANA CAROLINA DE SOUZA SANTOS - RTVI - UFJF - TCC II - 2025.1

ORIENTAÇÃO: PROF. JOÃO PAULO MALERBA

COORIENTAÇÃO: PROFA. MÁRCIA FALABELLA

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

1	TEC	ENTRA BG em fade in
2		
3	TEC	[Som de celular tocando, seguido de clique de atendimento]
4		
5	PER1	Bruno: Alô?
6		
7	PER2	Marina: Pronto.
8		
9	PER1	Bruno: (estranhando) Voz diferente... Gripada?
10		
11	PER2	Marina: Faringite.
12		
13	PER1	Bruno: Deve ser porque anda saindo direto pra curtir, né?
14		
15	PER2	Marina: (seca) E se estivesse? Algum problema?
16		
17	PER1	Bruno: (tosse leve) Não, imagina! Agora você é uma mulher livre.
18		
19		
20	PER2	Marina: (irônica) E você? Sua voz também está estranha.
21		Resfriado?
22		
23	PER1	Bruno: Constipado
24		

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

25	PER2	Marina: (rindo leve) Constipado? Mas é um velho mesmo, quem
26		fala assim hoje em dia?
27		
28	PER1	Bruno: A gente aprende...
29		
30	PER2	Marina: (som de copo batendo na mesa) Tá vendo? A separação
31		serviu pra alguma coisa.
32		
33	TEC	[Som de panela sendo mexida ao fundo, um leve chiar de óleo]
34		
35	PER1	Bruno: Viver sozinho é bom. A gente cresce.
36		
37	PER2	Marina: Você sempre viveu sozinho. Até quando casado, só fez o
38		que quis.
39		
40	PER1	Bruno: Maldade sua. Deixei de lado várias coisas quando a
41		gente se casou.
42		
43	PER2	Marina: (fundo som de fritura aumentando de intensidade)
44		Evidente! Só faltava continuar indo pra festa com os brothers.
45		
46	PER1	Bruno: Já você não abriu mão de nada. Streaming, compras
47		online, horas no WhatsApp com as amigas...
48		
49	TEC	[Silêncio, ao fundo som de copo caindo e quebrando]
50		
51	PER2	Marina: Comprinhas? A única coisa que comprei em quinze anos
52		de casamento foi um par de brincos.
53		

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

54	PER1	Bruno: Quinze anos? Achei que fosse menos.
55	PER2	Marina: A memória dos homens é um caso de polícia!
56		
57	PER1	Bruno: Mas passar horas no WhatsApp...
58		
59	PER2	Marina: Solidão, meu caro. Trabalhar fora, cuidar da casa,
60		preparar jantar pro HERÓI que chega à noite. Convenhamos...
61		
62	PER1	Bruno: Você nunca reclamou disso.
63		
64	PER2	Marina: E você perguntou alguma vez?
65		
66	PER1	Bruno: (cortando) Lá vem você de novo. Sempre a mesma
67		ladainha.
68		
69	PER2	Marina: Isso se chama diálogo! Nunca conversamos de verdade.
70		
71	PER1	Bruno: Ninguém fala nada na hora. Depois reclamam que faltou
72		diálogo. Mulheres são de outro mundo.
73		
74	PER2	Marina: E vocês homens só podem ser de outra dimensão.
75		
76	TEC	[Silêncio curto, som de porta batendo ao fundo e de cachorro
77		latindo distante]
78		
79	PER1	Bruno: E então, como vai a vida?
80		
81	PER2	Marina: Nunca estive tão bem. Livre pra pensar. Ninguém pra me
82		dizer o que fazer.

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

83		
84	PER1	Bruno: E isso é bom?
85		
86	PER2	Marina: Depois de quinze anos de jornada, é um alívio.
87		
88	PER1	Bruno: Eu nunca fui autoritário!
89		
90	PER2	Marina: Também nunca foi compreensivo.
91		
92	PER1	Bruno: Jamais disse que era perfeito.
93		
94	PER2	Marina: Limitado e omissos.
95		
96	PER1	Bruno: Você nunca foi irônica assim.
97		
98	PER2	Marina: A gente aprende também.
99		
100	PER1	Bruno: Sempre te apoiei.
101		
102	PER2	Marina: Foi no segundo mês de casamento que você lavou a única louça da sua vida. Inesquecível.
103		
104		
105	PER1	Bruno: Do que você está falando?
106		
106	PER2	Marina: (impaciente) Copa do Mundo? Vinte caipirinhas pra um bando de marmanjos? Meu sonho de vida.
107		
108		
109	PER1	Bruno: Eu detesto futebol!
110		

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

111	PER2	Marina: (surpresa) O quê? Alexandre, você tá louco?
112	PER1	Bruno: Alexandre? Meu nome é Bruno!
113		
114	TEC	[Silêncio, ao fundo barulho de panela caindo no chão e um leve "ai"]
115		
116		
117	PER1	Bruno: De onde você está falando?
118		
119	PER2	Marina: 9123-4567.
120		
121	PER1	Bruno: Não é 9123-4576?
122		
123	PER2	Marina: Não.
124		
125	PER1	Bruno: (rindo) Desculpe, engano.
126		
127		[Ambos riem alto]
128		
129	PER1	Bruno: Então você faz uma ótima caipirinha?
130		
131	PER2	Marina: Modéstia à parte. Mas prefiro vinho tinto.
132		
133	PER1	Bruno: Mesmo? Vinho é minha bebida favorita.
134		
135	PER2	Marina: E detesta futebol?
136		
137	PER1	Bruno: 22 homens correndo atrás de uma bola? Deus me livre,
138		Ridículo!
139		

Programa: Episódio 01 - Releitura do texto "Casal é Tudo Igual" de Luís Fernando Veríssimo

Adaptação, edição e produção: Ana Carolina Santos Tempo: 5 minutos

140	PER2	Marina: Bem, eu vou preparar o jantar.
141	PER1	Bruno: Que pena, o meu já está pronto. Risoto. Minha
142		especialidade.
143		
144	PER2	Marina: (intrigada) Mentira! Meu prato favorito.
145		
146	PER1	Bruno: A porção dá pra dois. Tem Chianti também. Não gostaria
147		de...
148		
149	PER2	Marina: (interrompendo) Adoraria!
150		
151	PER1	Bruno: E você mora aonde?
152		
153	PER2	Marina: Na Brasil, esquina com a São Gonçalo.
154		
155	PER1	Bruno: (animação) Moro a dois quarteirões de você. Pegar ou
156		largar?
157		
158	PER2	Marina: (rindo) Tô passando aí, Bruno.
159		
160	PER1	Bruno: Combinado, vizinha!
161		
162	TEC	[Som de panela sendo tampada, cachorro latindo novamente, e
163		riso leve dos dois enquanto a conexão é encerrada. Som de
164		chamada sendo encerrada e sobe BG - Trilha crescente
165		otimista.]
166		

APÊNDICE 2 - Perfil dos personagens

Bruno - Omar

- **Idade:** 32 anos
- **Profissão:** Personal Chef, apaixonado por gastronomia contemporânea e vinhos.
- **Personalidade:** Descontraído e espirituoso, Bruno é o tipo de pessoa que transforma situações comuns em momentos leves e bem-humorados. Apesar do bom humor, ele ainda carrega resquícios de introspecção causados pela separação recente. Acha graça na ironia e não se esquivava de provocações, mas há momentos em que ele demonstra certa vulnerabilidade. Ama cozinhar e se orgulha disso, usando a cozinha como uma forma de expressar afeto e se conectar com as pessoas.

Sugestão de rumo de interpretação

Bruno:

- Mostre descontração natural, mas insira nuances de hesitação quando a conversa começa a ficar mais pessoal.
- Use pausas estratégicas para dar a impressão de alguém que está saboreando cada momento, seja no diálogo ou na cozinha.
- No final da conversa, insira uma leve animação e entusiasmo, como alguém que redescobriu uma faísca de conexão.
- Traga nuances de um homem que está tentando se reconectar com sua espontaneidade após um casamento pragmático com Renata. Isso pode ser expresso na leve ironia e no humor em momentos mais relaxados, mas com um tom que sugira alívio e curiosidade ao encontrar alguém como Marina.

Marina - Manu Serdeira

- **Idade:** 31 anos
- **Profissão:** Designer de Interiores, criativa e apaixonada por transformar espaços.
- **Personalidade:** Marina é direta e irônica, mas sempre com charme. Ela aprendeu a valorizar sua independência após a separação, mas isso não significa que deixou de se emocionar e refletir sobre o passado. Gosta de conversas profundas, mas também sabe brincar com as situações. Marina é prática, mas tem um lado sonhador que emerge em momentos inesperados, especialmente quando algo desperta seu interesse. Adora um bom vinho e usa isso como uma espécie de ritual de relaxamento.

Sugestão de rumo de interpretação

Marina:

- Dê a ela um tom de alguém que acabou de sair de um relacionamento controlado com Gustavo e está redescobrando a leveza e a liberdade de ser ela mesma. No início, use um toque de desconfiança nas respostas, mas permita que isso suavize à medida que a conversa avança.
- No início, mantenha uma postura levemente defensiva, como se estivesse testando os limites da conversa.
- Conforme a conversa avança, deixe a barreira cair aos poucos, revelando um lado mais espontâneo e encantador.
- No final, misture surpresa e empolgação contida, como alguém que não esperava encontrar tanta sintonia.

Legenda:

Dir. = Diretora

At.= atores

Tec.Som = Técnico de Som dá rádio (responsável por operar a mesa de som e gravar os episódios)

APÊNDICE 4 - Moodboard dos personagens

MoodBOARD TCC

MARINA



#FFDADA

cor preferida

Emmanuelle Serdeira

carol, minhas impressões sobre o personagem são de que ela parece ser bastante incisiva, bem irônica e debochada, pq em diversos momentos ela repete o que o outro personagem fala de forma que quase a ataca ele, pode ser um pouco de mágoa também, que no final parece que transforma quase em um alívio e uma esperança

primeiras impressões da atriz após a primeira leitura do roteiro

primeiras impressões da atriz após a primeira leitura do roteiro

A playlist tocando no celular da Marina:



BRUNO



Matheus Omar

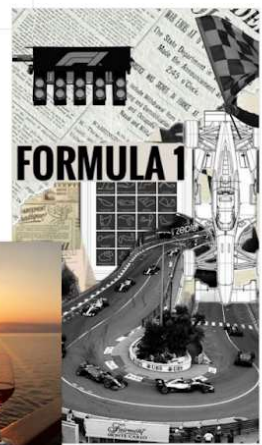
#02007A

Cor preferida

"Eh então eu achei tipo ele é um cara que é tipo uma pessoa mais independente assim e tal e que apesar dele parecer tipo entrasse bem resolvido ali que ele ele menciona que que amadureceu e tal eh eu acho que tipo uma pessoa que guarda muitas coisas pra ele sabe tipo e e acabava não até um momento que ele cobra também nunca falou nada e tal eh que eu consigo lembrar agora é isso?"

Primeiras impressões do ator depois da primeira leitura do roteiro

A playlist tocando no celular do Bruno:



Filmes que gosta muito



Filmes que gosta muito:



APÊNDICE 4 - ORDEM DO DIA – GRAVAÇÃO EP 01 (17/01)

Horário	Atividade	Duração	Responsável
17:00	Chegada e preparação	-	Todos
17:10	Aquecimento corporal	10 min	atores + diretora
17:20	Aquecimento vocal	10 min	atores + diretora
17:30	Motivação e imersão nos personagens	10 min	DiretorA
17:40	Gravação 1 (Roteiro completo)	40 min	At.+Dir.+ Tec.Som.
18:20	Pausa (água/banheiro)	10 min	Todos
18:30	Gravação 2 (Ajustes de atuação)	40 min	At.+Dir.+ Tec.Som.
19:10	Pausa (água/banheiro)	10 min	Todos
19:20	Gravação 3 (Versão final)	40 min	At.+Dir.+ Tec.Som.
20:00	Término das gravações	-	Todos

Legenda:

Dir. = Diretora

At.= atores

Tec.Som = Técnico de Som da rádio (responsável por operar a mesa de som e gravar os episódios)

APÊNDICE 5 - Cronograma de Produção

Data	Fase	Atividades Detalhadas
06/01 - 09/01	Pré-Produção	- Reserva de estúdios na FACOM - Encontro online com atores (apresentação do projeto, moodboard, personagens) - Primeira leitura do roteiro e feedback dos atores
13/01 - 19/01	Produção	- Contato com entrevistados (episódio 3) - Ensaio presencial com atores - Gravação do 1º episódio - Desenvolvimento da parte teórica
20/01 - 26/01	Pós-Produção	- Decupagem do EP 1 - E abastecendo a IA para clonagem de vozes para o 2º episódio - Continuação da parte escrita teórica
27/01 - 02/02	Pós-Produção e produção	- Edição do episódio 1 - produção do episódio 2 - Organização de materiais para memorial (fotos, roteiros, anotações)
10/02 - 16/02	Produção	- Gravação do 3º episódio - Ensaio com técnicas de Stanislavski (13/02 online e 15/02 presencial)
17/02	Gravação Chave	- Dia de gravação do episódio 1 (cronograma detalhado: aquecimento, 3 takes de 40min, pausas)
17/02 - 24/02	Pós-Produção	- Edição do 3º episódio - Revisão final do memorial e teoria
01/06 - 25/06	Site	- Desenvolvimento do site do projeto
04/06 - 14/07	Escrita TCC	- Capítulo 5: Som, Voz e Viewpoints Vocais (acréscimos) - Capítulo 6: Memorial (relatos de edição e processo)
09/06 - 14/07	Edição Final	- Finalização do episódio 3 (produto final)
16/06 - 30/06	Revisão TCC	- Capítulo 4: IA (acréscimos) - Capítulo 3: Direção de Atores (acréscimos)
14/07 - 20/07	Revisão	- Revisão e refinamento do episódio 3
14/07 - 24/07	Revisão Final	- Últimos ajustes no TCC e materiais
26/07	Entrega	- Submissão do TCC

APÊNDICE 6 - Roteiro base episódio 3



ANA CAROLINA DE SOUZA SANTOS - RTVI - UFJF - TCC II - 2025.1

ORIENTAÇÃO: PROF. JOÃO PAULO MALERBA

COORDENAÇÃO: PROFA. MÁRCIA FALABELLA

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

1	TEC.	ENTRA vinheta de abertura em fade in
2		
3	LOC.	Ana Carolina: Este é o "Sonidos Sintéticos e Falas Humanas",
4		um podcast que investiga os impactos da Inteligência
5		Artificial no campo sonoro, nos processos criativos e na
6		atuação vocal. Episódio 3 - O que a IA ainda não consegue
7		dizer?
8		
9	TEC.	CAI vinheta de abertura em fade out e entra BG - Trilha suave,
10		digital, com clima reflexivo
11		
12	LOC.	No episódio de hoje, vamos abordar uma questão sensível: o que
13		diferencia uma voz real de uma voz criada por inteligência
14		artificial? Você vai ouvir relatos de quem teve a própria voz
15		clonada, reflexões com especialistas e uma análise comparativa
16		entre vozes humanas e vozes sintetizadas. Para começar, escute
17		o relato da jornalista Renata Lana
18		
19	RENATA	Eu sou jornalista, tenho pós-graduação em Marketing Digital e
20		estudo inteligência artificial há uns dois anos. Eu tenho uma
21		agência de comunicação já tem 11 anos, mas comecei minha
22		carreira no rádio, depois passei pela televisão em várias
23		funções: produção, repórter, apresentadora, âncora.

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

24	LOC.	Renata que teve sua voz clonada, conta agora para gente um pouco de como tudo aconteceu.
25		
26		
27	RENATA	Período de campanha política. Eu tava gravando pra um cliente, eu era responsável pelo texto, roteiro e voz. Aí o marqueteiro viu e perguntou se podia copiar o modelo do roteiro e se eu toparia gravar. Eu disse que sim, só que ele falou que o funcionário dele já tinha clonado a minha voz. Ele disse: "Nossa, Renata, clonaram sua voz aqui", e me mandou o áudio. Fiquei em choque. Nunca ouvi falar de brasileiro que tenha a voz clonada, só americano. E eu nem sabia o que fazer. Falei: "Mas eu tô aqui, posso gravar". Ele falou: "Não, já tem a clonagem. Quero saber se posso usar em três campanhas e a ideia do seu roteiro". Eu falei: "Tá, pode, dê um valor".
31		
32		
33		
34		
35		
36		
37		
38	TEC.	Sobe bg
39		
40	LOC.	Chamei para uma conversa e um análise as doutoras Márcia Falabella, atriz, professora na facom que traz ao podcast um olhar sobre a direção de atores, som e voz e a Talita magnolo doutora e pesquisadora de IA, nesse primeiro momento escutamos a voz real de Renata e depois a clonagem da voz da Renata e discorremos sobre.
41		
42		
43		
44		
45		
46		[Sobe som da voz real da Renata em campanha política de um candidato de ervália]
47		
48		
49		[Desce som da voz real e sobe som da voz clonada]
50		
51		
52	TALITA	Gente, eu não vi muita diferença não. Assustador.

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

53	MÁRCIA	Esse, realmente Normalmente conseguimos ver, mas essa ficou
54		muito próxima.
55		
56	TALITA	E que o jeito que ela fala também né marcinha? Ela tem uma
57		força de fala muito grande.
58		
59	MÁRCIA	Talvez aceleraram a fala dela, o que mascara pontos de
60		artificialidade.
61	TALITA	Sim, impressionante
62		[SOBE BG EM FADE IN]
63	TEC	
64		
65	LOC.	Renata teve a sorte de ser avisada e até remunerada pelo uso
66		da clonagem. Mas como ela mesma aponta, esse não é o padrão. A
67		tecnologia avança, e a legislação ainda não acompanha. O que
68		nos leva à pergunta: a IA pode mesmo substituir o humano?
69		
70	TEC	[DESCE BG EM FADE IN]
71		
72	RENATA	Tive sorte de ser avisada. Mas o padrão é clonar sem avisar.
73		IA já está substituindo pessoas. Entrei em grupos de
74		locutores, que tem vozes famosas, a voz da globo, vários nomes
75		importantes e todos disseram: processa. Mas eu que já estudava
76		IA há algum tempo, parei e pensei, vi que estavam lutando
77		contra algo que não controlam. O mercado clona e nem dá
78		satisfação. Eu ainda tive a sorte de ser avisada e remunerada.
79		E com isso eu tive a virada de chave para perceber que a
80		inteligência artificial já está impactando o trabalho das
81		pessoas. Porque antes eu só ouvia falar, mas quando acontece
		com você a percepção é outra, é diferente. Ai você percebe que

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

82		já é real, que já chegou,. A Inteligência artificial,
83		sinceramente no ponto que ela tá indo ela substitui facilmente
84		umas 2 , 3 pessoas em um setor. Isso já está acontecendo, é
85		realidade.
86		
87	LOC.	Marcinha e Talita também dão a suas opiniões sobre.
88		
89	MÁRCIA	Eu acho que ainda tem um longo caminho para atingirmos para
90		uma voz humana ser substituída por uma maquina.
91		
92	TALITA	Eu acho que isso já é possível, só que eu acho que existem
93		várias plataformas que fazem isso acontecer
94		
95	MÁRCIA	Acho que a palavra ainda é chave, então é possível que a
96		máquina chegue em um ponto. O que a talita fala nós notamos a
97		embocadura é diferente, a pegada é diferente. O que eu preciso
98		então? é que a máquina entenda isso, na clonagem dela parece
99		aceleraram a fala, e isso pode provavelmente ter mascarado
100	TEC	algum ponto de artificialidade que eu notaria.
101		
102	LOC.	[SOBE Trilha eletrônica suave em fade in]
103		
104		A fala humana é viva, orgânica, e nasce da escuta. Cada voz
105		carrega ritmo, intenção, corpo e história. Com base nisso,
106		utilizamos neste episódio uma ferramenta de análise chamada
106		Viewpoints Vocais, adaptada da metodologia cênica de Anne
107		Bogart e Tina Landau. Os Viewpoints Vocais propõem que a
108		escuta seja o ponto de partida para analisar a presença da
109		voz. São seis os aspectos observados: Ritmo, Timbre, Volume e
		intensidade, Articulação, Presença, Espaço sonoro. Agora, você vai
		ouvir dois trechos do mesmo roteiro. O primeiro foi
		interpretado por atores reais. O segundo, gerado por

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

110		inteligência artificial. Preste atenção
111		
112	TEC	SOBE versão com atores - aprox. 40 segundos] DESCE trecho
113	TEC	SOBE trecho [versão com IA - aprox. 40 segundos]DESCE trecho
114		
115	LOC.	Agora, vamos à análise crítica a partir dos Viewpoints Vocais:
116		No ritmo A fala humana varia, respira, hesita.
117		[entra trecho]
118		
119		A IA tende à linearidade e à monotonia.
120		
121		[entra trecho]
122		No Timbre, a textura vocal dos atores carrega identidade.
123		
124		[entra trecho]
125		As vozes clonadas ainda soam padronizadas e pouco
126		personalizadas.
127		
128		[entra trecho]
129		
130		O volume e intensidade Nos atores, há nuances dramáticas
131		claras.
132		[entra trecho]
133		
134		Na IA, não há modulação emocional verdadeira.
135		
136		Ao que diz a articulação. A dicção humana permite quebras,
137		vírgulas expressivas, pausas naturais.
138		[entra trecho]

Programa: Sonidos Sintéticos e Falas Humanas – Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

139		A IA fala com fluidez técnica, mas sem intenção dramática.
140		
141		[entra trecho]
142		
143		
144		A presença vocal é a escuta do outro, é responder
145		emocionalmente à cena.
146		
147		[entra trecho]
148		A IA não, ela apenas reproduz.
149		
150		[entra trecho]
151		
152		
153		Ao que diz o espaço sonoro. A voz humana interage com a
154		ambiência.
155		
156		[entra trecho]
157		A voz sintética, mesmo bem mixada, soa deslocada, sem
158		organicidade.
159		
160		[entra trecho]
161		A tecnologia impressiona, mas ainda carece de algo
162		fundamental: a escuta do outro.
163		
164	LOC.	Márcia Fallabella também compartilha sua visão sobre a
165		construção do som, da fala e dessa composição a partir de
166		diferentes viewpoints vocais.
	TEC	[SOBE trilha eletrônica suave em fade in]

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

MÁRCIA

A essência da fala, a fala começa na respiração. Eu vou respirar e essa respiração faz as minhas cordas vocais vão vibrar. A respiração é o combustível da fala, esse é o primeiro ponto. Depois a gente tem a entonação, é o desenho sonoro da fala que eu trabalho. Ritmo, quando a gente está falando aqui, alguns momentos aqui eu acelero o ritmo, alguns momentos eu falo mais cadenciado, porque estou pensando qual a próxima coisa que eu vou falar, estou elaborando. Então esse ritmo altera, nós temos pessoas que falam mais rápido e muitas vezes não conseguimos entender. Nisso você perde o colorido da fala, uma fala mais rica em sonoridades. Você tem também a dicção, a articulação, o desenho de quando eu falo, a minha boca ela se movimenta, eu tenho pausas como recursos. A gente tem nossas musicalidades, isso tem haver com a nossa língua, isso tem haver com a nossa formação cultural com o grupo social que convivemos, então a gente fala isso como um produto do nosso meio, e o modo que falamos pra gente é normal. Então temos um estranhamento de sotaque, aí quando você fala: quero que você mantém o sotaque carioca é interessante porque o carioca fala "paistel", ele coloca um "i" ele acrescenta, o mineiro corta, então assim você tem as variações. É saber, o que eu quero? Então eu acho que a gente precisa, a máquina precisa, e eu acho que vai conseguir chegar, a esse conhecimento desses elementos, desses viewpoints vocais que eu posso trabalhar em combinações absolutamente diferentes de acordo com a intenção que eu quero dar a fala.

LOC.

Agora para saber como a IA é percebida por pessoas fora do

Programa: Sonidos Sintéticos e Falas Humanas – Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

		meio acadêmico e de estudo de IA, trazemos 3 pessoas.
	TEC	[SOBE trilha eletrônica suave em fade in]
	JULIA	Me chamo Júlia, tenho 22 anos e conheço bem o Omar e a Emanuele, já fazem alguns anos. E eu achei muito parecido, a do Omar então foi a mais parecida, fiquei até assustada . Já a da Emanuelle achei muito parecida nos primeiro segundos mas depois ela foi mudando conforme ela foi falando e ficando mais diferente. Mas gostei muito, ficou muito legal e isso me assustou como a IA pode sim pegar nossas vozes e transformar e fica muito parecido.
	ANAJU	Me chamo Ana Julia, tenho 11 anos, e achei bem legal, eu acho que foi feito por pessoas reais e não achei nada estranho. Feito por ia? Não é possível , eu não acredito
	IVONE	Me chamo Ivone , tenho 71 anos, gostei do diálogo ficou bem interessante e bem real , parece que foi feito por pessoas reais, só no final que teve uma pausa muito grande então achei estranho essa pausa aí achei que pudesse ser inteligência artificial mesmo.
	LOC.	Se, do ponto de vista técnico, a inteligência artificial ainda apresenta limitações vocais, o relato de Júlia, que conhece bem a voz original dos atores, mostra como a escuta pessoal é capaz de perceber aproximações e nuances que escapam aos critérios mais rígidos da análise. No caso da voz de Omar, por

Programa: Sonidos Sintéticos e Falas Humanas – Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

exemplo, há algo que soa familiar, ainda que não idêntico. Já os depoimentos de Ivone e Ana Júlia, mesmo sem conhecimento específico sobre IA, demonstram como essa tecnologia tem alcançado públicos de diferentes idades. Esses relatos reforçam a urgência de regulamentar o uso de IAs sonoras e da clonagem de voz, diante do risco de manipulação e uso indevido de identidades vocais.

Marcinha e Talita compartilharam suas percepções sobre os desafios e limitações do uso atual da inteligência artificial.

TEC

[SOBE trilha eletrônica suave em fade in]

[ENTRA BG]

TALITA

A gente tem que considerar que os softwares de IA têm a língua inglesa como base. Então qualquer coisa que vamos fazer na língua portuguesa não vai dar bom, porque não é a língua mãe. Então eles todos seus trejeitos, seus modos de falar, de gesticular, que não é do mesmo jeito que a gente fala. Então também temos essa dificuldade da língua.

Eu acho também que quando falamos de direção de atores e isso eu to falando de um lugar com zero experiência, marcinha pode me ajudar melhor nesse quesito, mas quando você vai preparar um ator você dá a ele toda uma contextualização, então você vai contar toda a história, olha você foi casada, você tem esse medo, você tem essa essa sensação de liberdade, então o ator já vem preparado, com essa bagagem. E a ia não, ela atende comandos mas ela não tem essa bagagem, que é uma bagagem subjetiva, aí eu já li autores falando que ela é boa em fingir sentimentos mas ela jamais vai sentir, então ela

Programa: Sonidos Sintéticos e Falas Humanas – Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

		pode simular o amor, a dor do divórcio, a surpresa mas ela não vai conseguir chegar aos pés do sentimento humano.
	MÁRCIA	E que eu falo que um diretor de atores você trabalha com os atores para chegar naquela intenção, naquela inflexão, na ideia que o texto transmite, agora você chega na Ia e quais comandos é possível, porque assim, hoje pra você usar bem a IA você precisa saber fazer a pergunta certa. e você não tem isso
	TALITA	Aí é preciso de uma legislação, é por isso que às vezes debatemos tanto, é preciso formas de usar. E a verdade que o que a IA faz não é novidade, se pensarmos um pouco atrás temos o photoshop, as mixagens ,até os fotógrafos que fazia a encenação de cenas de guerra , a manipulação de imagens e de sonora sempre existiu. O que tá pegando pra gente é justamente essa facilidade, essa rapidez e essa terra sem lei que estamos navegando, diante de pessoas que vão usar pro bem e pro mal.
	MÁRCIA	Eu sempre falo assim , é inteligência artificial porque é artificial.
	LOC.	"A presença cênica não é apenas estar no palco, mas estar escutando o outro, disponível à ação compartilhada". No campo sonoro, isso significa que uma voz verdadeiramente viva é aquela que escuta antes de falar. A escuta se manifesta nas pausas, nas hesitações, nos silêncios carregados de intenção. A clonagem vocal, embora tecnicamente avançada, ainda não escuta. E por isso, ainda não compõe. Porque escutar é também criar.

Programa: Sonidos Sintéticos e Falas Humanas - Episódio 3: O desafio de replicar a essência humana em uma voz artificial

Adaptação, edição e produção: Ana Carolina Santos Tempo: 23 minutos

TEC

[DESCE BG]

[SOBE trilha eletrônica inicial e final em fade in]

LOC

Este foi o terceiro episódio do podcast "Sonidos Sintéticos e Falas Humanas", uma produção de conclusão de curso em Rádio, TV e Internet da UFJF, com orientação do professor João Paulo Malerba. Agradeço à jornalista Renata Lana, à professora Marcinha, à professora Thalita e a todas entrevistadas. Obrigada!