UNIVERSIDADE FEDERAL DA PARAÍBA

PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

DOUTORADO EM LINGUÍSTICA

PREDITORES ACÚSTICOS DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS: UM MODELO BASEADO EM MACHINE LEARNING

DEYVERSON DA SILVA EVANGELISTA

JOÃO PESSOA

2024

DEYVERSON DA SILVA EVANGELISTA

PREDITORES ACÚSTICOS DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS: UM MODELO BASEADO EM MACHINE LEARNING

Tese apresentada ao Programa Associado de Pós-Graduação em Linguística (PROLING) da Universidade Federal da Paraíba (UFPB), como pré-requisito para obtenção do título de doutor em Linguística.

Orientador: Dr. Leonardo Wanderley Lopes

JOÃO PESSOA

2024

DEYVERSON DA SILVA EVANGELISTA

PREDITORES ACÚSTICOS DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS: UM MODELO BASEADO EM MACHINE LEARNING

Tese apresentada ao Programa Associado de Pós-Graduação em Linguística (PROLING) da Universidade Federal da Paraíba (UFPB), como pré-requisito para obtenção do título de doutor em Linguística.

Aprovada em: 27 de fevereiro de 2024

Banca examinadora

_____

Prof. Dr. Leonardo Wanderley Lopes

Orientador

_____

Prof. Dr. Giorvan Ânderson dos Santos Alves – UFPB

Examinador interno

_____

Profa. Dra. Anna Alice Figueirêdo de Almeida – UFPB

Examinadora externa

_____

Profa. Drª Rosiane Kimiko Yamasaki Odagima – UNIFESP

Examinadora externa

_____

Profa. Drª. Vanessa Veis Ribeiro – UNB

Examinadora externa

JOÃO PESSOA

2024

UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS, LETRAS E ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA

PROLING

DECLARAÇÃO

Declaramos, para os devidos fins, que o aluno **DEYVERSON DA SILVA EVANGELISTA** foi APROVADO na DEFESA de TESE em LINGUÍSTICA/PROLING - João Pessoa - DOUTORADO ACADÊMICO do Curso de DOUTORADO, no dia 27 de fevereiro de 2024 às 14:00, no(a) AUDITÓRIO DO CCS, UFPB, cuja banca examinadora fora constituída pelos professores:

Doutor (a) LEONARDO WANDERLEY LOPES (Presidente)
Doutor (a) GIORVAN ANDERSON DOS SANTOS ALVES (Interno)
Mestre (a) FERNANDA PEREIRA FRANÇA (Externo ao Programa)
Doutor (a) ANNA ALICE FIGUEIREDO DE ALMEIDA QUEIROZ (Externo ao Programa)
Doutor (a) LARISSA NADJARA ALVES ALMEIDA (Externo à Instituição)
Doutor (a) ROSIANE KIMIKO YAMASAKI ODAGIMA (Externo à Instituição)
VANESSA VEIS RIBEIRO (Externo à Instituição)

A sua TESE intitulou-se:

PREDITORES ACÚSTICOS DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS: UM MODELO BASEADO EM MACHINE LEARNING

Esta declaração não exclui o aluno de efetuar as mudanças sugeridas pela banca nem vale como outorga de grau de DOUTORADO, de acordo com o definido na Resolução 079/2013-CONSEPE.

João Pessoa, 27 de fevereiro de 2024.

JAN EDSON RODRIGUES LEITE
COORDENADOR(A) DO PROGRAMA DE PÓS-GRADUAÇÃO EM LINGUÍSTICA- UFPB

Catalogação na publicação Seção de Catalogação e Classificação

E92p Evangelista, Deyverson da Silva.

Preditores acústicos das atitudes dos ouvintes em relação às vozes disfônicas: um modelo baseado em machine learning / Deyverson da Silva Evangelista. - João Pessoa, 2024.

186 f. : il.

Orientação: Leonardo Wanderley Lopes.
Tese (Doutorado) - UFPB/CCHLA.

UFPB/BC CDU 616.22-008.5(043)

Dedico este manuscrito à todas as pessoas que confiam
na minha atuação fonoaudiológica e me possibilitam exercer a arte do cuidar!

AGRADECIMENTOS

Aos meus pais Maria Selma e Mozart Ferreira e a minha irmã Thamyris Evangelista, que sempre serão os meus maiores motivadores e exemplos para a vida pessoal e profissional.

Aos meus amigos por estarem sempre comigo e por toda a compreensão diante da minha jornada acadêmica e profissional. De forma especial, registro a minha gratidão à uma amiga, confidente, parceira de negócios e grande profissional: Profª Drª Larissa Nadjara. Você sempre será uma grande inspiração!

Aos professores que marcaram a minha trajetória acadêmica até aqui: Muito obrigado. Em especial, quero registrar a minha gratidão ao meu orientador por 11 anos: Prof. Dr. Leonardo Wanderley Lopes. Seu empenho e zelo pela fonoaudiologia motivam a nova geração de fonoaudiólogos do país.

Aos meus pacientes, clientes e alunos: Vocês sempre serão uma das grandes motivações neste percurso.

“Poder é ter uma voz para inspirar,
influenciar e gerar mudanças.”

Julia Hart

RESUMO

OBJETIVO: Analisar a relação entre as medidas acústicas e o julgamento das atitudes dos ouvintes na análise de vozes disfônicas; e desenvolver um modelo de predição da valência das atitudes dos ouvintes para vozes disfônicas baseado em Machine Learning (ML). **MÉTODOS:** O estudo apresenta-se como retrospectivo, descritivo, transversal e quantitativo. Contou com 152 ouvintes brasileiros e 44 amostras vocais. Foram realizadas duas etapas: seleção das amostras vocais e extração das medidas acústicas por meio do *software* Praat, versão 5.3.77h, através do script VoxMore; e julgamento dos ouvintes, por meio da Escala de Julgamento de Atitudes Associadas a Vozes Disfônicas, que engloba 12 atributos e seus respectivos antônimos foram baseados em três dimensões básicas das atitudes: *avaliação*, *potência* e *atividade.* Os dados foram analisados de forma descritiva e inferencial, por meio de testes de associação, correlação, comparação de médias, regressão linear e da ML**. RESULTADOS:** Observou-se relação entre medidas acústicas, julgamento perceptivoauditivo (JPA), e atitudes dos ouvintes. No JPA, destacam-se os parâmetros grau geral, soprosidade e rugosidade em todas as atitudes, exceto calma, que se comportou de forma diferente, com relação apenas com soprosidade. As medidas acústicas que mais se relacionaram com a percepção de atitudes foram as cepstrais, de frequência e de perturbação, apresentando correlações moderadas e fortes com a valência em cada atitude, bem como diferenças nas médias dos parâmetros para valência positiva e negativa. O Grau geral, f_0 mínima e SNL influenciam no julgamento das atitudes dos ouvintes, de acordo com a análise de regressão. Dessa forma, quanto maior o desvio vocal e as alterações acústicas no sinal vocal mais negativo é o julgamento. A partir da ML, foi observado que as medidas CCP, *Shimmer*, f_0 máxima e f_0 mínima são preditivas no julgamento de ouvintes para vozes disfônicas. **CONCLUSÃO:** As medidas de f_0 máxima e f_0 mínima, CCP e shimmer foram preditoras da atitude dos ouvintes no julgamento das vozes disfônicas. Medidas cepstrais, de frequência e perturbação foram as que mais se relacionaram com as atitudes dos ouvintes na análise de vozes disfônicas, tanto em relação à valência quanto aos atributos. O maior componente de ruído e irregularidade vocal também influenciou na percepção dos ouvintes.

Palavras- chave: Voz; Acústica; Qualidade vocal; Atitude dos ouvintes; Distúrbios da Voz.

ABSTRACT

OBJECTIVE: To analyze the relationship between acoustic measurements and listeners' attitudes towards dysphonic voices and develop a model based on Machine Learning (ML) to predict the valence of listeners' attitudes towards dysphonic voices. **METHODS:** The study is retrospective, descriptive, cross-sectional and quantitative. It had 152 Brazilian listeners and 44 vocal samples. Two steps were carried out: selection of vocal samples and removal of acoustic measurements using the Praat software, version 5.3.77h, using the VoxMore script; and listeners' judgment, using the Scale of Judgment of Attitudes Associated with Dysphonic Voices, which encompasses 12 attributes and their transfer was based on three basic dimensions of attitudes: evaluation, potency and activity. The data were analyzed descriptively and inferentially, using association tests, mean comparison, linear regression and ML. **RESULTS:** We observed the relationship between acoustic measurements, auditory perceptual judgment (JPA), and listeners' attitudes. In JPA, the configurations of general degree, breathiness and roughness stand out in all attitudes, except calm, which behaved differently, in relation only to breathiness. The acoustic measures that were most related to the perception of attitudes were cepstral, frequency and disturbance, showing moderate and strong correlations with the valence in each attitude, as well as differences in the means of the parameters for positive and negative valence. The General Degree, minimum f_0 and SNL influence the judgment of listeners' attitudes, according to the regression analysis. Therefore, the greater the vocal deviation and acoustic changes in the vocal signal, the more negative the judgment. From the ML, it was observed that the CCP, Shimmer, maximum f_0 and minimum f_0 measures are predictive in the listeners' judgment of dysphonic voices. **CONCLUSION:** Cestral, frequency and disturbance measurements were those that most related to listeners' attitudes towards dysphonic voices, both in relation to valence and attributes. The greater component of noise and vocal irregularity also influences perception. The measures of maximum f_0 and minimum f_0, CCP and shimmer are predictors in the judgment of listeners' attitudes towards dysphonic voices.

Keywords: Voice; Acoustics; Vocal quality; Attitude of listeners; Voice Disorders.

LISTA DE QUADROS

LISTA DE TABELAS

LISTA DE FIGURAS

SUMÁRIO

1 INTRODUCÃO

A voz é um dos principais veículos para a comunicação humana, consistindo em uma manifestação complexa das características biológicas, psicológicas e sociais de um indivíduo. Ela reflete não apenas a mensagem que deseja ser transmitida, mas também revela nuances da identidade social, profissional e afetiva do falante (ZHAO *et al*., 2019; PUTS *et al*., 2007). As nuances vocais fornecem ao ouvinte pistas sutis sobre a atratividade sexual, condição física e outras características que se entrelaçam com os elementos verbais e acústicos da fala, influenciando a percepção e o julgamento do ouvinte (SUIRE *et al*., 2018; LORTIE *et al*., 2018; KLOFSTAD *et al*., 2015; BABEL *et al*., 2014; GELFER & BENNETT, 2013; KREIMAN & SIDTIS, 2011).

A capacidade de comunicação clara e eficaz não só gera impressões positivas no ouvinte, mas também é essencial para o sucesso profissional, especialmente para aqueles que dependem da voz em seu trabalho, o que representa uma faixa significativa da população ativa (TITZE, LEMKE & MONTEQUIN, 2022). As percepções dos ouvintes, baseadas em pistas vocais, podem desencadear avaliações positivas ou negativas em relação à personalidade do falante, sua competência, saúde e outros atributos (EVANGELISTA *et al*., 2022; CORBARI, 2013). Esses julgamentos são moldados por fatores culturais, valores sociais e estereótipos preexistentes (BESTELMEYER *et al*., 2010).

No âmbito da comunicação interpessoal, além de pesquisas com indivíduos vocalmente saudáveis, há um esforço para compreender como a presença de um distúrbio de voz pode impactar na performance comunicativa e nas interações sociais e profissionais de indivíduos disfônicos (WAARAMAA *et al*., 2021; LATOSZEK *et al*., 2017). Nesse sentido, o próprio conceito de disfonia envolve qualquer dificuldade ou distúrbio na produção vocal que prejudica na transmissão da mensagem verbal e emocional do indivíduo, repercutindo negativamente nos aspectos psicossociais do falante (ISETTI, XUEREB & EADIE, 2014; COSTA V., *et al*., 2012; HOUTLE, *et al*., 2011; BEHLAU, 2005).

A gênese das disfonias pode estar associada a fatores orgânicos, comportamentais ou da interação entre eles. A sua manifestação é multidimensional e, por isso, o diagnóstico deve ser multifacetado e envolver a história clínica, a autoavaliação vocal, o julgamento perceptivoauditivo (JPA), a análise acústica, a avaliação aerodinâmica e o exame visual da laringe (PATEL *et al*., 2018). O processo

de avaliação está focado em compreender o mecanismo fisiopatológico envolvido na produção vocal e o seu impacto na vida do sujeito.

Por outro lado, nenhum desses itens da avaliação é capaz de capturar o impacto de uma voz disfônica no interlocutor. O grau do desvio vocal e dos parâmetros relacionados à rugosidade, soprosidade e tensão, por exemplo, podem predizer atitudes negativas dos ouvintes em relação a sujeitos disfônicos (EVANGELISTA *et al*., 2022; AMIR & LEVINE-YUNDOF, 2013; ALLARD & WILLIAMS, 2008).

As medidas acústicas também têm sido associadas às atitudes dos ouvintes. Homens com frequência fundamental (f_o) mais baixas tendem a ser avaliados mais positivamente, associados a atributos de competência e atratividade (ANDERSON *et al*., 2014; SKRINDA *et al*., 2014; MCALEER, TODOROV & BELIN, 2014). Por outro lado, vozes femininas com f0 mais elevadas são frequentemente relacionadas a jovialidade e fertilidade (BABEL, MCGUIRE & KING, 2014; PISANSKI & RENDALL, 2011; BLIGH & ROBINSON, 2010). Além disso, a relação harmônico-ruído está associada à avaliação positiva ou negativa de falantes. Vozes com maior componente de ruído têm maior probabilidade de receber julgamento negativo por parte dos ouvintes (BABEL *et al*., 2014; LATINUS *et al*., 2011; BRUCKKERT *et al*., 2010; BAUMANN *et al*., 2008).

Embora haja uma reconhecida associação entre as medidas acústicas e o julgamento de atitudes dos ouvintes, a maioria dos estudos realiza análise das medidas acústicas individualmente (PETTY *et al*., 2022; AUNG & PUTS, 2020; SUIRE *et al*., 2019 2018; ARMSTRONG *et al*., 2019; SEBESTA *et al*., 2017; TSANTANI *et al*., 2017). Durante as interações humanas, os ouvintes utilizam diferentes pistas acústicas da voz do ouvinte no processo (inconsciente ou não) de atribuir atitudes positivas ou negativas ao falante. Além disso, tais pistas acústicas interagem e podem ter pesos diferentes na predição das atitudes dos ouvintes. Poucos estudos têm utilizado modelagens estatísticas ou computacionais que permitam compreender a interação entre diferentes medidas acústicas da voz do falante, na determinação da atitude de um ouvinte (ZEHENG *et al*., 2020; RAYMOND *et al*., 2019; SHIRAZI *et al*., 2018).

Nesse contexto, as técnicas de inteligência artificial (IA) têm desempenhado um papel fundamental no reconhecimento de emoções e intenções a partir da voz (BERGNER *et al*., 2023; TOLMEIJER *et al*., 2021; LEE SEO-YOUNG *et al*., 2019).

Com o avanço da tecnologia, surgiram inúmeras aplicações inovadoras nessa área, que vão desde a análise de sentimentos em interações sociais até aplicações em saúde.

A IA constitui um domínio da ciência computacional dedicado à simulação de processos cognitivos humanos por meio de sistemas computacionais (SANTOS *et al*., 2019). Seu escopo abarca o desenvolvimento de algoritmos e modelos que facultam às máquinas a realização de tarefas que, se executadas por seres humanos, seriam qualificadas como manifestações de inteligência. Esse campo busca dotar entidades artificiais com habilidades para o processamento de linguagem natural, percepção sensorial, aprendizado, raciocínio abstrato e geração de artefatos criativos (ROBIN J. *et al*., 2020).

Machine Learning (ML), um subsetor da IA, envolve a criação e aperfeiçoamento de algoritmos e modelos estatísticos que dotam as máquinas da capacidade de aprender empiricamente (BERARDI *et al*., 2017). Distinto da programação convencional, na qual a lógica de decisão é explicitamente codificada, o ML permite que as máquinas otimizem seu comportamento por meio da detecção de padrões em conjuntos de dados volumosos.

A ML é um catalisador do avanço da Inteligência Artificial, conferindo a sistemas computacionais a habilidade de adaptar-se autonomamente e responder às complexidades ambientais e demandas humanas com crescente precisão e eficiência. A ML, enquanto ramificação pragmática da IA, tem fomentado avanços significativos em múltiplos domínios científicos, entre eles a Linguística e a Saúde. A inserção de ML nestas áreas tem não apenas potencializado a análise de dados complexos, mas também revolucionado as metodologias e aplicações práticas, promovendo um salto qualitativo no modo como compreendemos e interagimos com fenômenos linguísticos e de saúde.

Na Linguística, ML tem contribuído para a expansão do campo conhecido como Processamento de Linguagem Natural (PLN), permitindo a automatização e aperfeiçoamento de tarefas como tradução automática, análise de sentimentos, sumarização de textos e reconhecimento de fala (LE GRÉZAUSE, 2017; STYLER, 2015). Algoritmos de ML são treinados com grandes corpora de texto para modelar a linguagem humana, identificando padrões gramaticais, semânticos e pragmáticos que imitam a complexidade do idioma. Isso viabiliza a criação de interfaces homem-

máquina mais intuitivas, como assistentes virtuais que respondem a comandos verbais e sistemas que oferecem feedback linguístico personalizado. A capacidade de processar grandes volumes de texto em tempo real também tem viabilizado novas abordagens em pesquisas lexicográficas e sociolinguísticas, onde a detecção de tendências de uso linguístico e a emergência de neologismos podem ser rastreadas e analisadas com precisão sem precedentes.

Ademais, ML tem sido crucial para desvendar estruturas linguísticas inerentes a idiomas pouco documentados ou para a construção de recursos educacionais que se adaptam ao estilo de aprendizado individual de cada estudante (STYLER, 2015). Com o uso de técnicas como o aprendizado supervisionado e não supervisionado, é possível desenvolver modelos que compreendem e geram linguagem com um nível de sofisticação que era inimaginável há poucas décadas.

Na área da Saúde, a aplicação de ML tem grande potencial para reformular a prática clínica. O diagnóstico assistido por computador, por exemplo, tem se beneficiado enormemente da capacidade de ML em identificar padrões em imagens médicas, como radiografias, ressonâncias magnéticas e tomografias computadorizadas (RICHENS *et al*., 2020; CHOUDHURY *et al*., 2019; DE BRUIJINE, 2016). Algoritmos são capazes de detectar estruturas fora do padrão, como tumores, com uma precisão que, em muitos casos, supera a de clínicos experientes. Essa habilidade é particularmente valiosa para a triagem precoce de doenças, possibilitando intervenções mais rápidas e aumentando as chances de sucesso no tratamento.

Além disso, ML tem sido aplicado no desenvolvimento de terapias personalizadas, especialmente no campo da oncologia (FIELD *et al*., 2021; BERTSIMAS & WIBERG, 2020; CUOCOLO *et al*., 2020). Através da análise de grandes conjuntos de dados genômicos, os modelos de ML podem identificar quais pacientes são mais propensos a responder a determinados tratamentos, orientando a medicina de precisão. No gerenciamento de doenças crônicas, como diabetes e hipertensão, dispositivos vestíveis integrados com algoritmos de ML permitem o monitoramento contínuo dos pacientes, oferecendo recomendações em tempo real e ajustando tratamentos baseados em dados fisiológicos coletados ao longo do tempo.

Os modelos preditivos de ML também são uma ferramenta poderosa na saúde pública para a modelagem de surtos de doenças infecciosas, ajudando a prever e

mitigar a propagação de epidemias. Além disso, na pesquisa de novos medicamentos, as técnicas de ML aceleram a descoberta de compostos farmacológicos, analisando enormes bibliotecas de moléculas para prever quais são mais prováveis de serem eficazes como medicamentos.

Em resumo, Machine Learning está desempenhando um papel transformador tanto na Linguística quanto na Saúde. Na Linguística, está expandindo as fronteiras do possível no entendimento e na interação com a linguagem humana. Na Saúde, está melhorando o diagnóstico, o tratamento e a prevenção de doenças, além de personalizar o cuidado ao paciente e otimizar a eficácia dos sistemas de saúde pública. À medida que essas tecnologias continuam a evoluir, podemos esperar que sua integração nessas áreas se torne ainda mais profunda e revolucionária.

Modelos baseados em ML podem contribuir na avaliação inicial e diagnóstico precoce de pacientes com queixas vocais, além de ser balizador do efeito do tratamento vocal em indivíduos com disfonia (LEITE; MORAES; LOPES, 2022; UMENO *et al*., 2022, 2020; FAGHERAZZI *et al*., 2021; CALIFF *et al*., 2018; COHEN *et al*., 2012). Além disso, as técnicas de ML têm desempenhado um papel fundamental no reconhecimento de emoções e intenções a partir da voz, e com aplicações em áreas como marketing e comunicação profissional (SARA M. *et al*., 2017; YU G. *et al*., 2016).

No reconhecimento de emoções, algoritmos de ML podem analisar padrões na entonação, no ritmo e nas características acústicas da voz para identificar estados emocionais (TARUNIKA *et al*., 2018). Isso tem sido útil em sistemas de assistentes virtuais, onde a IA pode detectar se um usuário está frustrado, feliz ou irritado para ajustar as respostas de acordo com essas emoções, oferecendo um atendimento mais personalizado (SHIRAMIZU *et al*., 2022).

Além disso, no campo da saúde mental, a detecção de emoções pela voz pode ser uma ferramenta valiosa. Pesquisas exploram como a IA pode identificar sinais de depressão, ansiedade ou estresse por meio da análise vocal, auxiliando na triagem precoce e no suporte psicológico (ESPINOLA C. W. *et al*., 2021; SHIN D., *et al*., 2021). As pesquisas sobre os biomarcadores extraídos da voz e da fala, a partir das técnicas de ML possibilitam diagnosticar, prever e monitorar algumas condições de saúde (J. ROBIN *et al*., 2020). Há resultados satisfatórios para algumas alterações como a disfonia espasmódica, disartria, déficits cognitivos leves, desvios da qualidade

vocal, hiperfunção laríngea e a doença de Alzheimer (KOJIMA T., *et al*., 2022; H. LIN *et al*., 2020; J. M. TRACY *et al*, 2020; K. LOPES *et al*., 2013; RITCHINGS R. T., *et al*., 2002).

Quanto ao reconhecimento de intenções, sistemas baseados em IA são capazes de analisar os padrões linguísticos e vocais para determinar as intenções associadas às situações de comunicação. Isso é aplicado em assistentes virtuais para entender comandos e solicitações dos usuários, adaptando as respostas de acordo com a intenção expressa.

A incorporação de algoritmos de ML na análise da interação entre características acústicas vocais e as subsequentes atitudes dos ouvintes representa um avanço notável no estudo interdisciplinar da Fonética, Psicolinguística e Psicologia Social. Tais modelos computacionais têm a capacidade de extrair e processar variáveis acústicas detalhadas da fala, como f_0, intensidade, timbre e outros parâmetros espectrais, e correlacioná-las com as reações afetivas e cognitivas dos ouvintes. Através do treinamento supervisionado em datasets anotados, esses modelos podem predizer com precisão a valência das atitudes dos ouvintes, baseando-se em características vocais objetivamente mensuráveis.

No contexto clínico, o emprego de ML na avaliação de pacientes disfônicos é de suma importância. A disfonia pode afetar adversamente a comunicação e levar a repercussões negativas na interação social, profissional e emocional dos pacientes. O uso de modelos preditivos de ML para quantificar as reações dos ouvintes à disfonia pode fornecer insights valiosos sobre o impacto dessa condição na vida diária dos pacientes, além de servir como um indicador objetivo da eficácia das intervenções terapêuticas vocais.

A análise longitudinal dessas atitudes utilizando ML pode também facilitar a monitorização da evolução do paciente sob tratamento, permitindo ajustes finos e personalizados das abordagens terapêuticas. A identificação de padrões acústicos específicos — por exemplo, o grau de soprosidade ou a regularidade do *jitter* — que são mais fortemente associados a avaliações negativas por parte dos ouvintes, pode direcionar a terapia vocal para focar em aspectos específicos da produção vocal.

Para pacientes disfônicos, sistemas de feedback em tempo real baseados em ML podem prover uma avaliação objetiva da percepção social de sua voz, potencializando a autogestão e a conscientização das próprias habilidades

comunicativas. A disponibilidade de feedback imediato e quantitativo sobre melhorias na qualidade vocal pode servir como um reforço positivo, contribuindo para o aumento da adesão ao tratamento e para a melhoria da autoeficácia.

A integração futura desses modelos em plataformas digitais de fácil acesso, como aplicativos móveis e dispositivos *wearable*, tem potencial para aumentar a acessibilidade e a praticidade do monitoramento vocal. Esta tecnologia poderia, idealmente, oferecer notificações proativas quando variações vocais patológicas são detectadas, encorajando a prática de exercícios de reabilitação vocal ou a busca por avaliação clínica especializada.

A implementação de técnicas ML no estudo das reações dos ouvintes à qualidade vocal pode emergir como uma abordagem inovadora na avaliação e tratamento de pessoas com disfonia. Além disso, a IA demonstra capacidade ímpar de detectar nuances acústicas que moldam as percepções dos ouvintes, permitindo previsões precisas sobre as atitudes em relação a vozes tanto saudáveis quanto disfônicas. Essa abordagem analítica e preditiva promete não só otimizar a assistência à saúde vocal, mas também expandir o conhecimento científico sobre o impacto social da disfonia, apoiando-se em parâmetros quantitativos que transcendem e complementam a observação perceptual e comportamental.

Além disso, o uso de técnicas de ML para prever as atitudes dos ouvintes ao escutar um falante pode ser promissora em outros contextos. Por exemplo, através da análise de dados de áudio e de interações sociais, algoritmos podem inferir as reações dos ouvintes, identificando se estão engajados, entediados, concordando ou discordando. Isso é valioso em áreas como marketing, onde compreender as reações do público pode influenciar estratégias de comunicação e vendas.

Por outro lado, as pesquisas que abordam a relação entre a atitude dos ouvintes e as medidas acústicas, concentram-se em países europeus e norte-americanos. Tal fato limita a generalização dos achados, pois sabe-se que a língua nativa pode influenciar no julgamento e nas preferências das características vocais do falante, modificando percepções esperadas e desejáveis em contextos comunicativos (IRANI, ABDALLA, HUGHES, 2014; BORSEL, JANSSENS, BODT, 2009). No Brasil, há uma tendência de realizar estudos sobre julgamento de atitudes e sua relação com a percepção da qualidade vocal, mas não há evidências sobre os julgamentos do

ouvinte e a sua relação com medidas acústicas com indivíduos disfônicos nativos do português brasileiro, associando-o com o método ML.

Dessa forma, a presente pesquisa tem como hipóteses que:

I. Existe associação entre atitudes dos ouvintes e medidas acústicas obtidas nas amostras vocais de falantes disfônicos e não disfônicos nativos do português brasileiro.

II. Há medidas acústicas que predizem a valência (positiva/negativa) das atitudes dos ouvintes em relação às vozes de falantes disfônicos e não disfônicos nativos do português brasileiro.

A presente tese de doutorado ambiciona trazer contribuições significativas em diversas esferas, intercalando avanços clínicos, tecnológicos, sociais, metodológicos, educativos e de pesquisa. No espectro clínico, espera-se que o desenvolvimento de um modelo preditivo baseado em ML para a análise da relação entre medidas acústicas e atitudes dos ouvintes frente a vozes disfônicas resulte em um diagnóstico mais acurado e intervenções mais céleres e apropriadas à singularidade do perfil vocal do paciente. A aplicação dessa ferramenta preditiva promete não apenas a personalização do tratamento, mas também a mensuração objetiva dos resultados terapêuticos através da monitorização da evolução das percepções dos ouvintes em relação à voz tratada.

Do ponto de vista tecnológico, a tese tem o potencial de impulsionar o desenvolvimento de softwares inovadores, que possam ser utilizados tanto por pacientes para o auto-monitoramento da condição vocal quanto por profissionais da saúde para aprimorar o acompanhamento clínico. Além disso, insere-se no crescente campo de aplicação da IA na Linguística e em interface com a saúde, demonstrando como algoritmos podem ser treinados para discernir e interpretar variáveis complexas de características humanas, neste caso, as propriedades acústicas da voz.

No contexto social e psicológico, a investigação poderá elucidar como diferentes qualidades vocais afetam a percepção social, contribuindo para a compreensão de preconceitos auditivos e estereótipos vocais. Tal entendimento é essencial para desenvolver estratégias que visem mitigar os impactos negativos da disfonia na qualidade de vida dos indivíduos afetados, tanto no âmbito emocional quanto no social.

Na dimensão metodológica, esta pesquisa exemplifica a modelagem interdisciplinar, integrando conhecimentos de linguística, acústica, psicologia, fonoaudiologia e ciência da computação, e fornece um avanço para a Fonética e Linguística, ao detalhar como as propriedades acústicas da fala estão relacionadas com a percepção social e com as atitudes dos ouvintes.

No âmbito educativo, a tese oferece um caso de estudo valioso para o ensino de IA mostrando como problemas complexos e multidimensionais podem ser abordados por meio de métodos computacionais avançados. Adicionalmente, promove a capacitação de linguistas e profissionais de áreas correlatas em técnicas de análise de dados e IA expandindo suas habilidades e aptidões profissionais.

Finalmente, em termos de pesquisa, os resultados obtidos podem pavimentar o caminho para futuras investigações, estabelecendo novas perguntas de pesquisa e abordagens metodológicas tanto para estudos clínicos quanto para investigações fundamentais sobre a voz humana e a percepção auditiva. Em suma, a tese propõe-se a ser uma contribuição interdisciplinar e inovadora, com aplicabilidade direta nos estudos de linguística e na prática clínica, com potencial de catalisar novas direções na pesquisa científica.

2 OBJETIVOS

2.1 GERAL

Analisar a relação entre as medidas acústicas e o julgamento das atitudes dos ouvintes na análise de vozes disfônicas; e desenvolver um modelo de predição da valência das atitudes dos ouvintes para vozes disfônicas baseado em *Machine Learning* (ML).

2.2 ESPECÍFICOS

- Verificar se existem diferenças nas medidas perceptivo-auditivas e acústicas entre vozes julgadas com valência positiva e negativa;
- Averiguar se existem diferenças nas medidas perceptivo-auditivas e acústicas entre vozes julgadas com valência positiva e negativa em cada atitude investigada;
- Analisar se existe correlação entre as atitudes dos ouvintes em relação às vozes disfônicas e não disfônicas, e as medidas perceptivo-auditivas e acústicas;
- Investigar se as medidas acústicas e perceptivo-auditivas são capazes de predizer a variabilidade no julgamento de atitudes dos ouvintes;
- Identificar as medidas capazes de predizer a valência das atitudes dos ouvintes em relação às vozes dos falantes disfônicos e não disfônicos nativos do português brasileiro;
- Desenvolver um modelo preditivo para as atitudes dos ouvintes em relação a vozes disfônicas, de acordo com a valência positiva e negativa.

3 FUNDAMENTACÃO TEÓRICA

3.1.1 Julgamento de atitudes: conceitos e aspectos metodológicos

As primeiras concepções descritas na literatura sobre atitudes datam do século XX e advém das áreas de Psicologia Experimental e Social, definindo atitude como uma disposição interna do indivíduo perante a um grupo social, objeto ou situação que repercute na tomada de decisão (ALLPORT, 1967).

As atitudes, de um modo geral, desempenham uma função importante no comportamento do indivíduo e culmina em uma reação favorável ou desfavorável, a partir de uma maneira organizada e coerente de pensar, agir e reagir a qualquer acontecimento no ambiente. Portanto, a atitude se expressa através de pensamentos e crenças, sentimentos e emoções, bem como tendências para reagir às situações (LAMBERT, 1963).

Na área da linguística, o conceito de atitude engloba diversas dimensões, relacionada a variedades dialetais e estilos de fala, permeando desde as atitudes relacionadas ao aprendizado de uma língua, até as atitudes com relação a grupos, comunidades, minorias, entre outros (SAVILLE, 2003). A atitude pode manifestar-se como preferências e atribuições realizadas pelos falantes em contextos comunicativos (PERLOFF, 2008).

Duas teorias permeiam os estudos das atitudes linguísticas: Teoria Mentalista e a Teoria Behaviorista. Na corrente Mentalista, a atitude é vista como um elemento que se apresenta entre um estímulo e a reação que um indivíduo manifesta, fazendo-se necessário recorrer a técnicas indiretas, para que haja a extração de respostas do estado mental (CORBARI, 2013; ALLPORT, 1967). Na corrente Behaviorista, a atitude é interpretada como uma conduta, uma reação ou resposta a um estímulo, isto é, a uma língua, uma situação ou a características sociais determinadas (CORBARI, 2013).

Essas teorias se diferem pelo fato de as atitudes possuírem, ou não, mais de um componente em sua constituição. Para os mentalistas, as atitudes possuem componentes afetivos, cognitivos e conativos, enquanto para os behavioristas, possui apenas o componente afetivo (AGHEYISI e FISHMAN, 1980).

Diante dos conceitos existentes na literatura, o material teórico desta tese refletirá a atitude como uma forma de pensar, sentir e agir, com base na voz ou fala de um indivíduo e assim gerar respostas positivas ou negativas (ALLPORT, 1967; GARRET, 2010), a partir de três componentes: cognição, afeto e comportamento.

As atitudes são cognitivas na medida em que contêm ou compreendem crenças sobre o mundo, suas relações e significado social. São afetivas porque envolvem sentimentos sobre o objeto da atitude, a partir de um valor que é atribuído a ele, e consideradas comportamentais através da predisposição para agir mediante à alguma situação, além da sua relação com os julgamentos cognitivos e afetivos (GARRET, 2010).

Deste modo, entende-se a atitude conforme a figura abaixo:

FIGURA 01: Elementos constituintes na definição da atitude

FONTE: GARRET, 2010, adaptado para o português brasileiro.

O componente cognitivo apresenta a maior relevância, uma vez que dele advém os conhecimentos e pré-julgamentos dos falantes, bem como crenças, estereótipos, expectativas sociais, entre outros. Já o componente afetivo centra-se em juízo de valores sobre as características da fala, a partir da variedade dialetal do falante, a associação com traços de identidade e o sentimento do ouvinte em relação

ao grupo ao qual pertence o falante. Por fim, há o componente comportamental que presume a intenção e o plano da ação para determinados contextos e circunstâncias do indivíduo, e influência na interação dos interlocutores em diferentes ambientes (GARRET, 2010).

Existem duas importantes fontes de atitudes: A experiência e o ambiente pessoal. Elas são adquiridas e modificadas pela observação ou reações diante de algum acontecimento, por exemplo (AGUILERA, 2008). A literatura descreve diversos aspectos que podem ser relacionados à atitude conforme descrito no quadro 01:

QUADRO 01: Diferentes conceitos relacionados à atitude

HÁBITOS	Diferem das atitudes na medida que são observadas como rotinas de comportamentos, contudo, as atitudes não são vistas como ações comportamentais essencialmente.
VALORES	São concepções mais gerais e globais e podem culminar em uma gama de atitudes.
CRENÇAS	Corresponde ao componente cognitivo das atitudes.
OPINIÕES	Apresenta características discursivas, no entanto, para algumas atitudes este recurso é mais difícil de ser utilizado.
ESTEREÓTIPOS	Moldam os processos cognitivos das atitudes e funcionam como filtro sobre as percepções a respeito de um indivíduo ou grupo, a partir de julgamentos positivos ou negativos.

FONTE: GARRET, 2010; CARGILE, *ET AL*., 1994; BAKER, 1992.

Perloff e colaboradores (2008), define a atitude como um fenômeno mental e emocional, intrínseco ao comportamento físico, mas que não pode ser observado diretamente, apenas inferindo, para que haja explicações satisfatórias e consistentes relacionadas ao indivíduo, grupo social ou etc.

Portanto, as atitudes são observadas como posturas que os indivíduos assumem frente a algo, que pode derivar em uma reação com juízo de valor, isto é, em julgamentos favoráveis ou desfavoráveis em relação a um objeto, situação, grupo social, etc. Diante disso, observou-se a necessidade de mensurar as possíveis reações por meio de escalas que possibilitem a medição das atitudes (FROSI, *et al*., 2010).

Dentre as formas de mensurar os efeitos e do julgamento de atitudes encontradas na comunicação humana, o estudioso Thurstone, em 1928, buscou por medidas objetivas e, em 1957, Dr. Osgood e seus colegas aperfeiçoaram e adaptaram a sua criação: A Escala de Diferencial Semântico. Esta é uma ferramenta precursora em que se baseia na premissa de que qualquer conceito (uma pessoa, palavra, pintura etc.), pode ser definido ou descrito linguisticamente em termos de direção e magnitude através do uso de pares adjetivos bipolares como, por exemplo, se o objeto é “bom ou mau”, “suave ou duro” ou “forte ou fraco” (OSGOOD, 1957).

Foi demonstrado que a escala é uma medida confiável e válida de atitude, isto é, fácil de administrar e de marcar (HEISE, 1970). Esta técnica revela informações sobre três dimensões básicas: a avaliação, potência (ou seja, força) e atividade. A dimensão da avaliação consiste em verificar se uma pessoa pensa de forma positiva ou negativamente sobre o tema da atitude, por exemplo, atributos como sujo-limpo e feio-bonito. A potência relaciona-se com características de força, isto é, o quão poderoso o sujeito é para a pessoa. Por fim, a dimensão de atividade consiste no fato de o indivíduo pesquisado ser considerado ativo ou passivo mediante à situações do dia a dia, com atributos relacionados à segurança e insegurança, por exemplo (OSGOOD, 1957).

A voz transmite diversas informações sobre os falantes e os ouvintes são capazes de identificar atitudes, com base nas características vocais do interlocutor (IRANI, 2014). É um marcador de informação linguística e faz parte da socialização humana, e como está inserida na comunicação oral torna-se uma das principais formas de interação entre as pessoas, desta forma, uma voz saudável terá maior eficiência interpessoal (AMIR; YUNDOF, 2013).

3.1.2 O Julgamento de atitudes em indivíduos disfônicos

A voz do ser humano pode se constituir como um som produzido a partir das pregas vocais que se abrem e se fecham em um movimento ondulatório, por meio do ar que vem dos pulmões, direcionada para o filtro, isto é, o trato vocal (cavidades oral, nasal e faríngea) alterando assim a sua frequência e intensidade (MCDERMOTT, 2012). A disfonia é caracterizada por qualquer dificuldade que o indivíduo possa apresentar na emissão deste som, impedindo-o de realizar a transmissão verbal e

emocional num contexto comunicativo (BEHLAU, 2005). Portanto, a disfonia pode ter um impacto psicossocial no falante, produzindo a sensação de que sua voz não pode expressar adequadamente suas emoções, competência e traços de personalidade (ISETTI D. *et al*., 2014; VAN HOUTTE, *et al*., 2011; ANGELILLO M. *et al*., 2009).

Ao analisar a relação existente entre a voz do falante e o julgamento de atitudes por parte dos ouvintes, uma atitude geral pode ser observada pelo conteúdo da informação transmitida. Contudo, o ouvinte pode perceber diversas características como a idade, gênero e intenções do falante por meio de alguns segundos de amostra de fala (LATINUS, 2011; BESTELMEYER *et al*., 2010), devido à estereótipos pré-estabelecidos de características vocais e influenciado pelas próprias expectativas, experiência e cultura.

Por meio do julgamento de atitudes através da voz, é possível investigar diversos aspectos do falante relacionados à simpatia, honestidade, ansiedade, autoridade e emoção (SOROKOWSKI, 2019; ZUCKERMAN, 1993; APPLE, 1979). Os indivíduos com disfonia podem ser julgados com atributos negativos, isto é, a fraqueza, incompetência e insegurança ou transmitir sentimentos relacionados ao medo, tristeza e raiva (ALTENBERG, FERRAND, 2022; MORSOMME, MINEL E VERDUYCKT 2011). Os falantes disfônicos são considerados menos atraentes, mais autoritários e menos confiáveis do que indivíduos vocalmente saudáveis (MORSOMME, MINEL, VERDUYCKT, 2011; ROGERSON, 2005).

Pesquisadores examinaram a percepção de adolescentes e adultos com relação a indivíduos disfônicos. Por meio da Escala de Diferencial Semântico foram observados julgamentos relacionados à credibilidade e ansiedade. Os ouvintes também relataram que manteriam maior distância social dos mesmos (MCKINNON, HESS & LANDRY, 2006).

Lallh e Putmam (2000) observaram que mulheres disfônicas são avaliadas pelos ouvintes como menos atraentes, fracas e desagradáveis em comparação com mulheres vocalmente saudáveis. Resultados semelhantes foram encontrados em adolescentes (LASS *et al*., 1991).

Há estudos baseados em um único falante (WILSON E TIM, 2006) em atores que retratam a disfonia, porém não são disfônicos (LALLH, 2000), amostras vocais que incluem apenas o sexo feminino (EADIE *et al*., 2017; BLOOD *et al*., 1979) ou que utilizaram apenas a qualidade vocal predominante (vozes rugosas ou soprosas) no

julgamento de atitudes (EADIE *et al*., 2017; ALTENBERG, FERRAND, 2006; GOBL, 2003). Todos os estudos, com diversos métodos e amostras, evidenciam a relação entre o julgamento negativo e a presença de disfonia.

Amir e Yundof (2013) realizaram um estudo com o objetivo de diminuir algumas lacunas existentes na literatura, a partir da utilização de seis amostras vocais de vozes disfônicas masculinas e femininas apresentadas a 74 ouvintes de ambos os sexos. Através da Escala de Diferencial Semântico, os pesquisadores observaram que os falantes disfônicos foram classificados como mais negativos, doentes e tensos, enquanto falantes não disfônicos foram classificados como mais bem sucedidos, sexy, sociáveis e inteligentes. Além disso, as mulheres disfônicas foram avaliadas mais negativamente do que os homens disfônicos.

Diante dos estudos descritos, nota-se que há interesse do meio científico em compreender a relação ao julgamento de atitudes em vozes disfônicas, porém os achados atuais se limitam a língua nativa do falante e do ouvinte, que pode influenciar as diferenças nas características vocais desejáveis num contexto comunicativo (WAARAMAA *et al*., 2021; IRANI, ABDALLA, HUGHES, 2014; KREIMAN J., 2010; ALTENBERG, FERRAND *et al*., 2006).

Evangelista e colaboradores (2022) desenvolveram o primeiro estudo sobre as atitudes associadas a vozes disfônicas entre falantes e ouvintes nativos do português brasileiro, com o objetivo de verificar possíveis diferenças no julgamento de atitudes sobre os falantes disfônicos e não disfônicos, bem como identificar os fatores preditivos das atitudes dos ouvintes em relação às vozes disfônicas desta população.

Dentre os principais resultados do estudo, foi possível observar que os sujeitos disfônicos, de ambos os sexos, foram julgados mais negativamente pelos ouvintes em comparação com indivíduos não disfônicos, nas três dimensões do questionário de diferencial semântico (avaliação, potência e atributos de atividade). Os disfônicos foram julgados como mais desagradáveis, fracos, frágeis, antipáticos, introvertidos, doentes, submissos, inseguros, incompetentes e dependentes (EVANGELISTA *et al*., 2022).

Destaca-se que o estudo utilizou apenas os parâmetros relacionados à qualidade vocal, considerada a característica mais estável do falante e torna-se uma fonte essencial de informações este indivíduo no meio social do falante, sendo uma das principais manifestações da disfonia (MA E YU, 2013; MAIDMENT, 1983). No

estudo em questão, avaliou-se a presença ou ausência do desvio, a intensidade classificada como leve, moderada ou intensa, e por fim, o predomínio vocal que se manifesta através dos parâmetros: rugosidade, soprosidade ou tensão.

A realização de pesquisas que busquem entender o impacto da disfonia no meio social, a partir de análises com medidas perceptivas e acústicas, repercutem positivamente para a compreensão das atitudes com relação aos indivíduos disfônicos e não disfônicos. Ao abordar apenas a natureza e o impacto do distúrbio de voz, há lacunas para compreensão da formação de atitudes negativas (PICKENS, 2005). É possível que outros fatores estejam influenciando a percepção social e uma melhor compreensão poderá ajudar a direcionar as intervenções destinadas aos disfônicos.

Desta forma, a voz é uma ferramenta comunicativa e de convívio social, a partir de uma visão multidimensional, precisa ser continuamente estudada em busca de parâmetros clínicos que sejam cada vez mais consistentes nas indicações de normalidade ou alteração.

3.2 AS MEDIDAS ACÚSTICAS E O JULGAMENTO DE ATITUDES

3.2.1 Medidas acústicas e o julgamento de atitudes

Diversas características do falante podem ser estimadas de forma válida e confiável com base em informações acústicas (KYRIAKOU; PETINOU e PHINIKETTOS, 2018). Estudos demonstram essa relação, sobretudo os internacionais, e ressaltam sua importância na clínica e social. Por isso é importante observar e descrever o que os artigos abordam sobre o tema ao longo do tempo, os processos metodológicos utilizados para coleta dos julgamentos e extração das medidas, e o que a literatura expõe de modo geral.

Dessa forma, buscou-se responder à seguinte pergunta norteadora: Quais as relações entre o julgamento de atitudes e as medidas acústicas de amostras vocais? Por meio de uma revisão integrativa da literatura, realizada nas bases de dados SciELO, LILACS e PubMed, desenvolvida em novembro de 2022.

Utilizou-se para a busca os descritores em Ciências da Saúde (DeCS) em português e seu correspondente em inglês: voz/*voice*; julgamento/ *Judgment*; acústica/*acoustics*, e os termos: julgamento de atitude/*judgment attitude*;

atratividade/*attractiveness*, selecionados de acordo com as palavras-chave encontradas em artigos sobre o tema, incluídas a fim de tornar a busca mais abrangente. Os termos foram conectados por meio do operador boleano *AND*.

O termo atratividade foi utilizado nesta busca devido ao consenso internacional de que este estereótipo pode resultar em gostos e preferências sociais por meio da voz. A atratividade relaciona-se às impressões positivas geradas nas pessoas em suas relações interpessoais (PUTS, 2003 2007 2011 2013). Desta forma, para revelar a existência do estereótipo da atratividade vocal, é preciso que haja concordância entre o julgamento da voz realizado pelo ouvinte e os diversos parâmetros vocais encontrados em uma emissão (ZUCKERMAN; MIYAKE,1993).

A busca contou com um total de 1768 artigos. A seleção dos artigos foi realizada por meio da leitura de títulos, resumos e dos artigos completos. Também foram considerados artigos, relacionados com o tema, que foram citados nos artigos encontrados nas bases. Eles foram selecionados seguindo os seguintes critérios de elegibilidade: a) presença dos termos utilizados na busca citados em seu título, resumo ou palavras chaves; b) ter objetivo de relacionar julgamento com acústica da voz; c) estudos originais com amostras vocais de seres humanos; d) descrever o método realizado para coleta dos julgamentos em relação às amostras vocais apresentadas; e)abranger amostras vocais de indivíduos de qualquer gênero ou idade do ciclo vital; f) amostras com indivíduos disfônicos ou vocalmente saudáveis; g) estudos publicados nos últimos 12 anos. A pesquisa não foi restritiva quanto à língua. Os artigos presentes em mais de uma base de dados e/ou pesquisa por palavras chaves foram analisados apenas uma vez.

Inicialmente, foi realizada leitura e avaliação dos títulos e resumos dos artigos encontrados de acordo com os descritores, em que foi observada a relação com o tema proposto e enquadramento nos critérios de seleção. Em seguida, houve a avaliação do texto completo. Os 25 manuscritos selecionados passaram por análise e categorização dos dados a partir dos seguintes aspectos: a) ano de publicação; b) objetivo; c) local de realização do estudo e língua; d) método para julgamento do ouvinte; e) número de estímulos auditivos; f) gênero da população das amostras vocais e de juízes; h) descrição dos resultados; i) presença de relação entre julgamento e análise acústica. Os dados foram apresentados no quadro 2, seguindo

a ordem cronológica de publicação do artigo e com destaque aos aspectos preestabelecidos.

QUADRO 02: Caracterização dos estudos que relacionam julgamento de atitudes e medidas acústicas dos artigos selecionados.

TÍTULO ANO IDIOMA	OBJETIVO	NÚMERO DE ESTÍMULOS AUDITIVOS E MÉTODO PARA O JULGAMENTO DO OUVINTE	NÚMERO E GÊNERO DOS JUIZES	RESULTADOS E CORRELACÕES ACÚSTICAS
Correlated preferences for men's facial and vocal masculinity. 2008 **Idioma do estudo: Inglês**	Investigar as preferências para a masculinidade vocal e facial dos homens.	- 06 falantes do sexo masculino. - A amostra vocal foi coletada através da contagem de números. - O julgamento de atitude foi realizado através de perguntas fechadas.	- 1.213 ouvintes do sexo feminino (experimento realizado de forma online)	- A fo reduzida culmina em vozes masculinas mais atraentes.
The Sound of Symmetry Revisited: Subjective and Objective Analyses of Voice. 2008. **Idioma do estudo: Inglês**	Investigar se o julgamento perceptivo das vozes se relaciona com a simetria facial.	- 31 falantes do sexo feminino. - 40 falantes do sexo masculino. - A amostra vocal foi coletada através da contagem de números - O julgamento de atitude foi realizado através Escala *Likert* através dos atributos: Acessível, dominante, saudável, honesto, inteligente, probabilidade de conseguir encontros, maduro, sensual e cordial.	- 50 ouvintes do sexo feminino. - 51 ouvintes do sexo masculino.	- A f_0 reduzida dos homens foi avaliada como mais atrativa. - Correlações positivas foram observadas entre o julgamento dos atributos honestidade e inteligência para vozes masculinas nas medidas de *Jitter* local, *jitter* rap e *jitter* ddp. - O estudo não apresentou correlações entre os atributos e as medidas acústicas para as amostras vocais femininas.

Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. 2008. Idioma do estudo: Inglês	Investigar a relação da fo com a atratividade para vozes masculinas.	- 04 falantes do sexo masculino. - A amostra vocal foi coletada através de Pergunta padrão dirigida ao falante. - O julgamento de atitude foi realizado através Escala *Likert* através dos atributos: Escala *Likert.*	- 123 ouvintes do sexo feminino.	A f_o reduzida relaciona-se com a maior atratividade para falantes do sexo masculino.
Face and voice attractiveness judgments change during adolescence. 2009. Idioma do estudo: Inglês.	Analisar a relação entre o julgamento e pistas multimodais de rosto, corpo e fala.	- 06 falantes do sexo feminino. - 06 falantes do sexo masculino. - 11- 13 anos. - 06 falantes do sexo feminino não disfônicas. - 06 falantes do sexo masculino. - 13-15 anos não disfônicos. - 06 falantes do sexo masculino. - 13 -15 anos não disfônicos. - O julgamento de atitude foi mensurado através da Escala *Likert* por meio de do atributo: Atratividade.	- 148 ouvintes do sexo feminino -177 ouvintes do sexo masculino	- A f_o feminina elevada relaciona-se com a atratividade (ouvintes masc. 11-13 anos). - A f_o feminina reduzida relaciona-se com a atratividade (ouvintes fem. 13-15 anos).
A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. 2010.	Investigar as preferências masculinas e femininas, a partir da fo da voz	- 06 falantes do sexo feminino . - 06 falantes do sexo masculino. - A amostra vocal foi coletada através de sentenças espontâneas.	- 200 ouvintes do sexo feminino. - 200 ouvintes do sexo masculino	- A f_o feminina elevada relaciona-se com atratividade dos ouvintes do sexo masculino.

Idioma do estudo: **Inglês.**		- O julgamento de atitude foi coletado através da Escala *Likert* por meio dos atributos: Atratividade e Dominância.		- A fo masculina reduzida relaciona-se com o atributo de dominância para ambos os ouvintes.
Different Vocal Parameters Predict Perceptions of Dominance and Attractiveness. 2010. **Idioma do estudo:** **Inglês.**	Investigar as relações entre os parâmetros vocais e o julgamento de dominância e atratividade de mulheres para vozes masculinas.	- 111 falantes do sexo masculino. - A amostra vocal foi coletada através da fala espontânea. - O julgamento de atitude foi mensurado através da Escala *Likert* por meio de dois atributos: Dominância e Atratividade.	- 142 ouvintes do sexo feminino.	- A f_o reduzida foi preditora para julgamentos de maior dominância e atratividade. - A menor dispersão dos formantes relacionou-se com a atratividade para vozes masculinas, apenas na fase fértil das ouvintes pesquisadas.
Correlated Male Preferences for Femininity in Female Faces and Voices. 2011. **Idioma do estudo:** **Inglês.**	Verificar a relação entre a atratividade vocal e visual dos homens em mulheres.	- 06 falantes do sexo feminino. - A amostra vocal foi coletada através da Vogais isoladas. - O julgamento perceptivo auditivo foi mensurado através de uma pergunta padrão dirigida ao falante: A voz e a imagem são atraentes?	- 178 ouvintes do gênero asculino.	- A fo feminina elevada e a menor dispersão dos formantes repercutem em maior atratividade para os ouvintes do sexo masculino.
Intrasexual competition among women: Vocal femininity affects perceptions of	Analisar os efeitos de medidas acústicas (média e desvio padrão de fo) na atratividade vocal masculina e feminina.	- 72 falantes do sexo feminino. - A amostra vocal foi coletada através da leitura de sentenças.	- 63 ouvintes do sexo masculino. - 46 ouvintes do sexo masculino.	- O Desvio Padrão da f_o apresentou correlação forte com o julgamento para relacionamento de longo

attractiveness and flirtatiousness. 2011. **Idioma do estudo:** **Inglês.**		- O julgamento de atitude foi mensurado através da Escala Likert e através de perguntas relacionadas a relacionamento afetivo de curto à longo prazo.		prazo (ouvintes do sexo masculino e feminino)
Female voice frequency in the context of dominance and attractiveness perception. 2011. **Idioma do estudo:** **Polonês.**	Investigar a relacão da f_o e as percepcões de atratividade e dominância.	- 58 falantes do sexo feminino. - A amostra vocal foi coletada através de vogais isoladas. - O julgamento de atitude foi mensurado através da Escala Likert dos atributos: Atratividade e dominância	- 144 ouvintes do sexo masculino avaliaram a atratividade. - 140 ouvintes do sexo masculino avaliaram a dominância. - 189 ouvintes do sexo feminino avaliaram a dominância	- Vozes femininas com f_o reduzida foram julgadas como mais atraentes (ouvintes do sexo masculino) e dominantes (ambos os sexos dos ouvintes).
What makes a female voice attractive? 2011. **Idioma do estudo:** **Inglês.**	Investigar a relação entre a atratividade vocal feminina, qualidade vocal e emoções.	- 1 falante do sexo feminino. - A amostra vocal foi coletada através de pergunta padrão dirigida ao falante. - Julgamento de atitude através da Escala *Likert* por meio do atributo de atratividade.	- 10 ouvintes do sexo masculino.	- A fo elevada, a qualidade vocal soprosa e o menor comprimento do trato vocal repercutem em maior atratividade para os ouvintes pesquisados.
Low Pitched Voices Are Perceived as Masculine and	Investigar as preferências vocais femininas, a partir de características de fertilidade dos homens.	- 54 falantes do sexo masculino. - A amostra vocal foi coletada através de vogais isoladas.	- 15 ouvintes do gênero feminino.	- A fo reduzida está relacionada ao julgamento de maior atratividade nos ouvintes pesquisados.

Attractive but Do They Predict Semen Quality in Men? 2011. **Idioma do estudo:** **Inglês**		- Julgamento de atitude através da Escala *Likert* por meio de dois atributos: atratividade e masculinidade.		- Não houve correlação significativa entre a masculinidade e a qualidade do sêmen do falante.
Dimension esthétique des voix normales et dysphoniques: Approches perceptive et acoustic. **2012.** **Idioma do estudo:** **Francês**	Investigar a relação da disfonia com a atratividade vocal e as medidas acústicas em falantes do sexo masculino.	- 31 falantes disfônicos do sexo masculino. - 31 falantes não disfônicos do sexo masculino - A amostra vocal foi coletada através Leitura de texto + vogal isolada. - Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade.	- 92 ouvintes do sexo feminino.	- Vozes rugosas e soprosas foram julgadas com maior atratividade. O estudo não apresentou correlações significativas entre a atratividade e a fo.
Preferences for Very Low and Very High Voice Pitch in Humans. **2012.** **Idioma do estudo: Inglês**	Investigar a relação da frequência vocal e a atratividade em vozes masculinas e femininas.	- 01 amostra vocal masculina criada em laboratório. - 01 amostra vocal feminina criada em laboratório. - As amostras vocais (vogais) foram criadas a partir de uma média de extrações. - Julgamento de atitude através de perguntas fechadas/binária.	- 09 ouvintes do sexo feminino. - 10 ouvintes do sexo masculino.	- A fo elevada repercute em maior atratividade para os ouvintes do sexo masculino. - A fo reduzida relaciona-se com maior atratividade para as ouvintes do sexo feminino.

Faking it: Deliberately altered voice pitch and vocal attractiveness 2013. **Idioma do estudo: Inglês**	Investigar a relação das alterações na fo e a atratividade em homens e mulheres.	- 04 falantes do sexo masculino - 04 falantes do sexo feminino - A amostra vocal foi coletada através de vogais isoladas em três etapas: tom habitual de voz, mais aguda e mais grave. - Julgamento de atitude através de perguntas fechadas/binária de dois atributos: Atratividade e domínio.	- 104 ouvintes do sexo feminino - 110 ouvintes do sexo masculino	- A manipulação através do aumento e diminuição da fo para vozes masculinas e femininas não afetou a atratividade dos ouvintes. - A manipulação com repercussão na fo reduzida, repercute em julgamentos de maior dominância para ambos os sexos dos falantes.
Men's preferences for women's femininity in dynamic cross-modal stimuli. 2013. **Idioma do estudo: Inglês.**	Pesquisar a relação de preferências masculinas e femininas, a partir de estímulos visuais e vocais.	- 05 falantes do sexo masculino. - 05 falantes do sexo feminino. - A amostra vocal foi coletada através de palavras controladas em experimento. - Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade.	- 128 ouvintes do sexo masculino.	- A fo elevada e traços femininos repercutem na maior atratividade vocal dos ouvintes pesquisados.
Human Vocal Attractiveness as Signaled by Body Size Projection. 2013. **Idioma do estudo: Inglês**	Investigar a relação entre os parâmetros acústicos e dimensões corporais dos falantes na atratividade vocal dos ouvintes.	- 01 falante do sexo feminino. - 01 falante do sexo masculino. - A amostra vocal foi coletada através de frases controladas. -Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade	- 16 ouvintes do sexo masculino. - 16 ouvintes do sexo feminino.	- A fo elevada das mulheres relaciona-se com maior atratividade para os homens. - A fo reduzida relaciona-se com maior atratividade para as mulheres. - A soprosidade relaciona-se com a atratividade para homens e mulheres.
Towards a More Nuanced View of Vocal Attractiveness. 2014.	Analisar Diversas medidas acústicas como preditores do julgamento de ouvintes, a partir da manipulação vocal dos falantes.	- 30 falantes do sexo masculino. - 30 falantes do sexo feminino	-30 ouvintes do sexo masculino - 30 ouvintes do sexo feminino.	- Formantes mais dispersos ou tratos vocais aparentemente mais curtos

Idioma do estudo: **Inglês**		- A amostra vocal foi coletada através de frases controladas. - Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade		foram preditores para maior atratividade dos ouvintes. - As medidas de fo e da posição dos formantes não foram preditivas para o julgamento de atratividade de vozes masculinas. - Vozes soprosas femininas foram julgadas com maior atratividade pelos ouvintes do sexo masculino. - VOT curto em vozes masculinas repercute em maior atratividade.
The Perception and Parameters of Intentional Voice Manipulation. 2014. **Idioma do estudo: Inglês**	Investigar a percepção de ouvintes perante intenções comunicativas por meio da voz.	- 20 falantes do sexo masculino. - 20 falantes do sexo feminino. - A amostra vocal foi coletada através da leitura de palavras em que os falantes realizaram duas gravações: A leitura normal e com mudança intencionalmente de fala para se tornarem mais atraentes, confiantes, dominantes e inteligentes. - Julgamento de atitude dos atributos sexy, confiável, dominante e inteligente, através de uma Escala *Likert*.	- 20 ouvintes do sexo masculino - 20 ouvintes do sexo feminino.	- Em geral as vozes manipuladas das mulheres foram julgadas como mais atraentes. - A Frequência Fundamental reduzida e velocidade de fala lentificada das vozes femininas repercute em maior atratividade para ambos os ouvintes (sexo masculino e feminino). - Não foram encontradas correlações entre o julgamento de atitudes e as medidas de *Jitter, Shimmer* e Relação Harmônico Ruído.

Body height, immunity, facial and vocal attractiveness in young men. 2015. **Idioma do estudo: Letã**	Investigar a relação de parâmetros acústicos e dados imunológicos e hormonais na atratividade vocal de homens.	- 60 falantes do sexo masculino. - A amostra vocal foi coletada através de vogais isoladas. - Julgamento de atitude através da Escala *Likert* para o julgamento do atributo: atratividade.	- 29 ouvintes do sexo feminino.	- A fo masculina reduzida relaciona-se com maior atratividade dos ouvintes. - Valores baixos do segundo formante gera maior atratividade vocal masculina.
Low Vocal Pitch Preference Drives First Impressions Irrespective of Context in Male Voices but Not in Female Voices. 2017. **Idioma do estudo: Inglês**	Investigar a influência da fo nos julgamentos de confiabilidade e dominância.	- 10 falantes do sexo feminino. - 09 falantes do sexo masculino - A amostra vocal foi coletada através de uma palavra padrão. - Julgamento de atitude através de dois atributos: Dominância e Confiança. "Qual voz você percebeu como mais confiável/dominante?"	- 183 ouvintes do sexo feminino - 57 ouvintes do sexo masculino	- A fo masculina reduzida relaciona-se com maior confiança para ambos os sexos e mais dominante apenas para as vozes masculinas. - Não foram observadas correlações da medida acústica pesquisada e a atratividade para as vozes femininas.
Voices of Africa: Acoustic predictors of human male vocal attractiveness. 2017. **Idioma do estudo: Estudo realizado de forma Interlinguística**	Identificar os preditores acústicos de atratividade em amostras vocais masculinas africanas.	- 45 falantes do sexo masculino (falantes de Camarões) - 48 falantes do sexo masculino (Falantes de Namíbia). - A amostra vocal foi coletada através de sentenças espontâneas. - Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade.	- 62 ouvintes do sexo feminino thecos	- A fo reduzida relaciona-se com maior atratividade em falantes de camarões. - A posição dos formantes e a relação harmônico ruído (HNR) relacionam-se com maior atratividade em falantes da Namíbia.
Filipino Women's Preferences for Male Voice Pitch: Intra-Individual, Life History, and Hormonal Predictors.	Verificar a atratividade vocal de mulheres lactantes e nulíparas.	- 06 falantes do sexo masculino - A amostra vocal foi coletada através de sentenças espontâneas de sentenças (frases) controladas.	- 63 ouvintes Lactantes. - 65 ouvintes multíparas	- A fo elevada repercute em maior atratividade para ambos os grupos de ouvintes.

2018. **Idioma do estudo: Estudo realizado de forma Interlinguística**		- Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade.		
Human vocal behavior within competitive Human vocal behavior within competitive and courtship contexts and its relation to mating success. 2018. **Idioma do estudo: Francês.**	Investigar a relação entre atratividade sexual com as diferenças acústicas	- 01 falante do sexo masculino. - 01 falante do sexo feminino. - A amostra vocal foi coletada através de uma sentença com traços de competitividade. - Julgamento de atitude através fechadas/binária.	- 68 ouvintes do sexo feminino. - 56 ouvintes do sexo masculino.	- A fo reduzida e o desvio padrão da mesma relaciona-se com maior atratividade sexual masculina. - A medida relação harmônico ruído elevada (HNR) relaciona-se com maior atratividade sexual feminina.
Male vocal quality and its relation to females' preferences. 2019. **Idioma do estudo: Francês**	Investigar as relações entre os parâmetros acústicos e a atratividade vocal feminina.	- 58 falantes do sexo masculino. - A amostra vocal foi coletada através da leitura de frases. - Julgamento de atitude através da Escala *Likert* por meio do atributo*:* atratividade.	- 135 ouvintes do sexo feminino	- Vozes masculinas francesas atrativas apresentaram a fo reduzida, o desvio padrão elevado. - As medidas de *Jitter* e HNR não apresentou correlações com o julgamento de atratividade.
Vocal attractiveness and voluntarily pitch-shifted voices. 2020. **Idioma do estudo: Chinês.**	Analisar a atratividade vocal de falantes, a partir de manipulações na fo.	- 80 falantes do sexo feminino. - 35 falantes do sexo masculino - A amostra vocal foi coletada através da leitura de frases. - Julgamento de atitude através fechadas/binária.	- 88 ouvintes do sexo feminino - 79 ouvintes do sexo masculino	- As amostras vocais femininas com a fo elevada apresentaram maior atratividade para ambos os ouvintes. - A fo reduzida em falantes masculinos foi relacionada à maior atratividade em ouvintes do sexo feminino.

De modo geral, a maioria das pesquisas utilizou amostras de falantes da língua inglesa e a f_0 foi a medida que mais se destacou.

No ponto de vista metodológico, a busca bibliográfica responde ao objetivo, contudo, não há padronização dos estímulos vocais e a quantidade usada nos experimentos para o julgamento de atitudes, utilizou-se fala espontânea, palavras ou vogais, leitura, entre outros. Da mesma forma, o número de ouvintes é variável. Para a análise das medidas acústicas as abordagens consistem em estudos correlacionais, que visam basicamente relacionar as medidas acústicas e a atratividade vocal por meio do julgamento dos ouvintes através de escalas tipo *Likert*. Há uma tendência, observada a partir dos estudos selecionados, de investigar a atitude dos ouvintes através de um atributo, isto é, a atratividade.

Os *insighs*, de modo geral, evidenciam que os estudos analisados apresentam um enfoque nas medidas de f_0 e o desvio padrão, a dispersão dos formantes e as suas relações com o julgamento de atitudes.

A fo é definida pelo número de vibrações por segundo produzidas pelas pregas vocais e as medidas de perturbação de frequência e amplitude reflete a eficiência do sistema fonatório, a biomecânica laríngea e interação com a aerodinâmica (BrOCKMANN-BAUSER E DRINNAN, 2011).

A dispersão de formantes é caracterizada pela diferença média entre frequências de formantes sucessivos e relaciona-se ao comprimento do trato vocal e ao tamanho do corpo (FEINBERG, 2008). Os tratos vocais mais longos produzem formantes mais baixos e mais espaçados, diferentemente de indivíduos com pregas vocais mais encurtadas e tratos vocais supralaríngeos mais finos que repercutem em formantes mais altos.

Há uma tendência regular de que as mulheres em seus ambientes linguísticos e/ou origens culturais, são predominantemente atraídas por homens exibindo vozes graves (GAULIN e PUTS, 2010; HUGHES, FARLEY, RHODES, 2010; JONES *et al*., 2010; PISANSKI e RENDALL, 2011; VUKOVIC *et al.,* 2008; XU *et al*., 2013; SUIRE *et al*., 2019 2018), sobretudo para falantes da língua inglesa, considerando o grande volume de artigos encontrados nesta língua (há variações a partir do inglês canadense, australiano, americano e britânico). Observando o panorama internacional considerou-se dois estudos interlinguísticos, em que os pesquisadores (SHIRAZI, PUTS, ESCASA-DORNE, 2018) observaram que mulheres filipinas apresentam maior atratividade para vozes masculinas com a fo elevada. A relação

harmônico ruído (HNR) pode estar relacionada com a atratividade vocal (SEBESTA *et al*., 2018 2017).

Quanto à atratividade vocal feminina, a grande maioria dos estudos apresentam que os homens são constantemente atraídos por mulheres com a fo elevada (BORKOWSKA e PAWLOWSKI, 2011; JONES *et al*., 2010; PUTS, *et al.,* 2011; RE *et al*., 2012).

Foi observado ainda que o desvio padrão da fo relaciona-se com maior atratividade em ambos os sexos (LEONGÓMEZ *et al*., 2014). Ao observar a dispersão dos formantes (HODGES-SIMEOn *et al*., 2010, PISANSKI e RENDALL 2011, SEBESTA *et al.,* 2017), mostraram que quanto menor a dispersão de formantes, as vozes masculinas são mais atraentes. Por outro lado, Skrinda *et al*. (2014) e Xu *et al*. (2013) não encontraram correlação entre a dispersão dos formantes e o julgamento de atratividade.

Destacam-se dois estudos que relacionaram as medidas acústicas de *Jitter*, *Shimmer*, relação harmônico ruído (HNR)*,* H1H2 e o julgamento de ouvintes para amostras vocais (BABEL *et al*., 2014; HUGHES *et al*., 2014; HUGHES *et al.,* 2008). A medida H1H2 relaciona-se com a inclinação de curta distância da amplitude do primeiro harmônico menos o segundo harmônico, isto é, mensura indiretamente o comprimento relativo da fase aberta das PPVV. Em vozes soprosas, a amplitude do primeiro harmônico é relativamente alta em comparação com os harmônicos seguintes relativamente (HILLENBRAND e HOUDE, 1996).

O *jitter* e o *shimmer* relacionam-se com as variações que ocorrem na fo e amplitude. O *Jitter* evidencia a variabilidade da fo, o *shimmer* refere-se a perturbação relacionada a amplitude da onda sonora ou intensidade do desvio vocal. O *jitter* altera-se principalmente com a falta de controle de vibração de pregas vocais e o *shimmer* com a redução da resistência glótica e lesões de massa nas pregas vocais, correlacionada com a presença de ruído à emissão e com a soprosidade (GOLDINO *et al.,* 2010). A medida de relação harmônico ruído (HNR) avalia a presença do ruído no sinal da voz. A HNR é uma medida que relaciona o componente periódico vibração das pregas vocais) com o componente não periódico (ruído glótico) dando indicação da eficiência do processo da fonação (FREITAS, 2012).

Foi possível verificar que apenas um estudo (DEFRADAS *et al.,* 2012) descreveu a presença de disfonia dentre os estímulos vocais utilizados. Hughes e colaboradores (2014) destacaram no manuscrito a dificuldade de correlacionar as

medidas acústicas pesquisadas com o julgamento dos ouvintes, no entanto, eram necessárias investigações futuras com a ampliação destas medidas e associações de fatores comportamentais (percepção dos ouvintes) e parâmetros clínicos da voz. Destaca-se que não foram encontrados estudos nacionais acerca da relação entre as medidas acústicas e o julgamento de ouvintes.

3.3 *Machine Learning*

O termo “Aprendizado de Máquina” (*Machine Learning - ML*), refere-se ao funcionamento de sistemas computacionais, como uma subdivisão da IA, capazes de aprender ou modificar o seu comportamento através de estímulos externos (ALPAYDIN, 2010 2014). De modo geral, a ML é um conjunto de abordagens que identificam de forma automática padrões dos dados existentes e, posteriormente, empregam esses padrões descobertos para antecipar eventos futuros ou facilitar a tomada de decisões (MALIK *et al.,* 2019; SHI; IYENGAR, 2020; XING; YANG, 2016).

A ML é a capacidade de aprimorar o desempenho na execução de tarefas por meio da experiência (MITCHELL, 1997). Esta fermenta desempenha um papel importante ao aprimorar habilidades, facilitar na elucidação de diagnósticos clínicos, contribuindo para a redução de falhas médicas, pois amplia o entendimento sobre uma determinada enfermidade, embasando efetivamente o processo de tomada de decisões clínicas (IYENGAR, 2020).

Na área de saúde, a ML tem sido aplicada em diversas esferas para solucionar desafios, como direcionar abordagens terapêuticas e reduzir atrasos nos diagnósticos de doenças. Em fonoaudiologia, tem sido utilizada nos processos de identificação e classificação de vozes disfônicas e não disfônicas, bem como no reconhecimento de padrões de fala e expressões faciais (FACELI, 2015).

Observa-se então uma integração de conceitos provenientes de várias áreas, como neurociências e biologia, juntamente com princípios estatísticos, matemáticos e físicos, a fim de concentra-se na identificação de mecanismos pelos quais os computadores podem aprender tarefas específicas (MARLSLAND, 2014; SHOLKOPF, 2008). Em outras palavras, a ênfase reside na inferência indutiva e na habilidade de generalização.

Dessa forma, informações relevantes relacionadas ao objeto de estudo são coletadas e elencadas em um conjunto de dados, que passarão pelo processo de seleção, aplicação e avaliação, como sugere Marsland (2014), descrito a seguir:

- Conjunto de dados: Os dados que serão utilizados para o aprendizado ou avaliação do modelo. A quantidade de dados, nesse caso, precisa ser considerada, pois algoritmos de aprendizados de máquina requerem quantidades significativas de dados.
- Treinamento: A etapa de treinamento consiste no uso de recursos computacionais afim de construir uma função *f*, ou um classificador, para prever as saídas em novos dados.
- Dados de validação: Uma porção do conjunto de dados, afim de realizar uma análise imparcial do modelo.
- Dados de teste: Uma porção de conjunto de dados empregada para uma avaliação conclusiva do modelo, com previsões reais que o modelo executará, possibilitando a verificação do desempenho do modelo.
- Hiperparâmetros: São parâmetros que definem a arquitetura de um método de aprendizado, sendo o processo de busca pela arquitetura ideal chamado de ajuste de hiperparâmetros. Enquanto os parâmetros do modelo especificam como transformar os objetos de entrada na saída desejada, os hiperparâmetros definem como o modelo é realmente estruturado.

Os métodos mais comumente utilizados na análise em questão de ML são classificadas em: supervisionados e não supervisionados (FACELI, 2015).

- Supervisionados: Nesse método, é disponibilizado um conjunto de exemplos contendo respostas corretas. Com base nesse conjunto, o algoritmo selecionado generaliza para responder corretamente à todas as possíveis futuras entradas. Esse conjunto é chamado de treinamento, pois é a partir desses que o algoritmo é treinado para realizar as generalizações. São exemplos de aprendizado supervisionado: análise de regressão linear, análise

de regressão logística, bem como abordagens mais modernas como *support vector machine* e *boosting* (ALBON, 2018).

- Não Supervisionados: Neste método, observar-se que respostas precisas para cada um dos itens não são previamente oferecidas. Dessa forma, os algoritmos buscam reconhecer semelhanças entre os itens de entrada, de modo que aqueles que compartilham características comuns são agrupados. A exploração das relações entre variáveis ou entre os itens de entrada é uma abordagem possível. É dito não supervisionado devido a falta de um vetor de respostas para orientar a análise. Uma ferramenta muito utilizada no aprendizado não supervisionado é a análise de cluster (JO, 2021).

Ao se comparar o desempenho de algoritmos de aprendizado não supervisionado com aqueles de aprendizado supervisionado, nota-se que os primeiros são mais eficientes/capazes na execução de tarefas complexas, especialmente em situações em que não há disponibilidade de dados já classificados ou rotulados. Entretanto, os modelos supervisionados geram resultados mais precisos, uma vez que um especialista instrui explicitamente o sistema sobre o que procurar nos dados fornecidos, mas é importante destacar que ainda assim os resultados podem ser imprevisíveis (JO, 2021).

Na área da voz, a maioria das estratégias que utilizam o ML, concentram-se principalmente na técnica de aprendizado supervisionado (HEGDE *et al*., 2019), devido à disponibilidade, ainda que limitada, de conjuntos de dados previamente rotulados, especialmente para as classificações de vozes disfônicas e não disfônicas (AL-NASHERI *et al*.,2018). Destacam-se estudos com o método supervisionado para reconhecimento de emoções de fala (Chen *et al*., 2020) e detecção e classificação do desvio vocal (WU H. *et al*., 2018; HOSSAIN M. S., MOHAMMAD G. 2016).

Para a utilização das técnicas de aprendizado supervisionado ou não supervisionado, há diversos modelos que auxiliam na classificação e otimização dos dados:

QUADRO 03: Modelos de machine learning.

Modelo	Descrição
Regressão Logística Penalizada (Elasticnet)	Técnica de Aprendizado de Máquinas que combina as penalidades L1 (Lasso) e L2 (Ridge) durante o treinamento de modelos de regressão. Útil em conjunto de dados com muitas características e multicolinearidade.
SVM Linear	Conjunto de algoritmos que podem ser aplicados tanto em problemas de classificação como também de regressão. Ferramenta valiosa na classificação binária e na possibilidade de os dados serem separados linearmente. Dentre as aplicações, tal modelo pode ser usado na classificação de gênero, identificação de falantes, detecção de emoções e patologias vocais.
SVM polinomial	Permite a modelagem de relações não lineares entre as características dos dados. Mostra-se viável para modelar relações não lineares das características acústicas da voz.
SVM KERNEL	A função principal do modelo é mapear os dados originais para um espaço de características de dimensão superior, proporcionando uma vantagem computacional significativa. Contribui para modelagem e interpretação da informação acústica da voz.
K-Vizinhos mais próximos (K-NN)	Técnica utilizada para classificação e regressão. Realiza previsões com base na proximidade entre os exemplos de treinamento. Útil para conjunto de dados pequenos a moderados e em cenários onde a interpretação é uma prioridade.
Árvore de classificação	Representação gráfica de decisões que tomam base em características dos dados. É importante gerenciar o *overfittin*g e ajustar os hiperparâmetros para obter um equilíbrio adequado entre a complexidade e desempenho.
Floresta Aleatória (Random Forest)	Torna-se uma extensão das árvores de decisão e combinam previsões de várias árvores para melhorar a previsão e reduzir o *overfitting.* Apresenta vantagens ao lidar com diversas características acústicas, reconhecimento de falantes e emoções por meio da voz.
Redes Neurais MLP (Multilayer Perceptron)	Pertence a categoria de redes neurais artificias. Utilizadas em uma variedade de tarefas como classificação, regressão, reconhecimento de padrões e processamento de imagens. Tal método oferece a capacidade de aprender representações complexas e identificar padrões sutis nas informações vocais.
Redes Neurais MLP combinadas através de Bagging (Bagged MLP)	Abordagem que visa melhorar a robustez do modelo, isto é, treina múltiplas instâncias do mesmo modelo em subconjuntos diferentes dos dados de treinamento e, em seguida, combina as previsões. Torna-se versátil e pode ser aplicada na análise da voz humana, onde padrões complexos podem ser aprendidos de maneira mais robusta por meio da combinação de múltiplas instâncias do modelo.
Naive Bayes	Se baseia no Teorema de Bayes para realizar classificação e torna-se eficaz para tarefas de classificação de textos e mineração de dados, tomada de decisões, entre outros.
Análise Discriminante Linear (LDA)	Técnica estatística utilizada em tarefas que visam classificação e redução de dimensionalidade. Na voz humana, a LDA pode ser explorada para extrair características discriminativas ou realizar tarefas de classificação baseadas em padrões vocais.
Análise Discriminante Regularizada (RDA).	Torna-se uma extensão da LDA e incorpora termos de regularização. É projetada para lidar com problemas como situações em que o número excessivo de características é maior do que as observações. Sua capacidade em equilibrar víeis e variância através da regularização, torna-o útil na área da voz humana, pois os conjuntos de dados acústicos podem ser complexos e multidimensionais.

FONTE: MORAIS R. M. *et al*., 2020; CHEN L. *et al*., 2019; MARTINEZ L., 2015; MORAIS R. M. *et al*., 2009; OLSON D. L.; DELEN D., 2008; HAYKIN, 2001.

Assim, análises a partir da ML permitem que clínicos e pesquisadores reconheçam os aspectos vocais que mais se destacam diante da identificação e diagnóstico de uma disfonia, possibilitando sua predição de forma mais rápida, acessível, objetiva e efetiva. Os resultados obtidos por meio destes métodos são capazes de embasar a tomada de decisão em esferas primárias e secundárias, como rastreios e triagens de problemas vocais, bem como no diagnóstico da alteração vocal instalada.

4 METODOLOGIA

4.1 DESENHO DO ESTUDO

Trata-se de uma pesquisa retrospectiva, observacional, descritiva, transversal e quantitativa. O projeto foi aprovado pelo Comitê de Ética em Pesquisa com Seres Humanos do Centro de Ciências da Saúde/UFPB, sob o número de parecer 5.259/734.

4.2 LOCAL DA PESQUISA

Esta pesquisa foi realizada no Campus I da Universidade Federal da Paraíba (UFPB), no Laboratório Integrado de Estudos da Voz (LIEV) do Departamento de Fonoaudiologia e vinculada ao Programa de Pós-Graduação em Linguística da UFPB.

4.3 AMOSTRA

Participaram do estudo 152 ouvintes nativos do português brasileiro, graduandos em cursos da área de saúde de uma Instituição de Ensino Superior (IES). A amostra foi selecionada por conveniência, sendo composta por 132 mulheres e 20 homens, com idade média de 21,9±5,6 anos.

Para o recrutamento dos ouvintes, foram considerados os seguintes critérios de elegibilidade: não apresentar história pregressa ou atual de queixa auditiva que impossibilitasse a escuta e o julgamento das tarefas de fala apresentadas; e não ter experiência anterior ou treinamento para JPA da qualidade vocal. Vale ressaltar que o gênero dos ouvintes não foi considerado variável de análise e que por isso não foi realizado pareamento da quantidade de homens e mulheres.

4.4 MATERIAIS E MÉTODOS PARA COLETA DE DADOS

O estudo contou com duas etapas: seleção das amostras vocais e julgamento dos ouvintes.

Inicialmente, foram selecionadas vozes saudáveis e disfônicas, de acordo com critérios de elegibilidade pré-definidos. Após a seleção das amostras vocais, foi

realizada abordagem dos participantes/ouvintes por meio da explanação dos objetivos do estudo e assinatura do assinaram o termo de Consentimento Livre e Esclarecido (TCLE) (Anexo B), autorizando o início da coleta de dados por meio da apresentação das vozes selecionas para realização do julgamento de atitudes, como descrito a seguir.

4.4.1 Seleção de amostras vocais

Para viabilizar o julgamento de atitudes pelos ouvintes, foi realizada construção de um *corpus* de vozes saudáveis e desviadas, a partir do banco de dados constituído no LIEV, após autorização dos coordenadores do laboratório (ANEXO C) Todos os indivíduos avaliados no laboratório autorizaram sua participação em pesquisas por meio do TCLE.

As amostras vocais foram coletadas durante avaliação inicial, atentes do início da terapia vocal, de acordo com os procedimentos rotineiros do LIEV, que incluem também: anamnese, aplicação de questionários de autoavaliação, JPA e acústica da voz, bem como realização do exame visual laríngeo, a videoestroboscopia, para verificar a presença de alterações estruturais ou funcionais. Os dados coletados são codificados em uma planilha no software Excel. Destaca-se que para o presente estudo, utilizou-se apenas dados pessoais, gênero e idade, e o laudo com o resultado do diagnóstico laríngeo.

Para a gravação vocal, foi utilizado: *software Fonoview* (Pato Branco, Paraná, Brasil), versão 4.5, da CTS Informática; *desktop* Dell *all-in-one*; microfone cardioide unidirecional, da marca Senheiser (Hanover, Alemanha), modelo E-835, localizado em um pedestal e acoplado a um pré-amplificador Behringer (U-Phoria UMC 204, Willich, Alemanha), modelo U-Phoria UMC 204.

As vozes foram coletadas em cabine de gravação com tratamento acústico e ruído inferior a 20 dB NPS (SPL) (R8050 REED *Instruments Sound Level Meter*), com taxa de amostragem de 44000 Hz, com 16 bits por amostra. Após receber instruções, os indivíduos realizaram as tarefas vocais em pé, com o pedestal a sua frente, com distância de 10 cm e um ângulo de 45º entre os lábios e o microfone. A distância entre a boca do locutor e o microfone foi medida usando uma régua de 10 cm antes de registrar cada tarefa.

As gravações das tarefas de voz e fala selecionadas no presente estudo foram a vogal [e] sustentada e as seis frases do CAPE-V, traduzidas e adaptadas para o português brasileiro (BELHAU, 2003): “Érica tomou suco de pêra e amora”, “Sônia sabe sambar sozinha”, “Olha lá o avião azul”, “Agora é hora de acabar”, “Minha mãe namorou um anjo” e “Papai trouxe pipoca quente”. Ambas as tarefas foram gravadas em frequência e intensidade habitual e confortável.

Inicialmente, foi incluída uma amostra de 656 vozes de sujeitos, sendo 511 mulheres e 145 homens. Desse grupo, 587 se enquadraram nos seguintes critérios de elegibilidade: adultos, com idade entre 18 e 65 anos, com queixa de voz e com avaliação laringológica prévia, incluindo laudo otorrinolaringológico escrito. Foram excluídos pacientes que não tivessem as duas gravações completas das tarefas de fala, usuários profissionais da voz, indivíduos que receberam terapia fonoaudiológica formal anteriormente ou aqueles que foram submetidos a cirurgia na região de cabeça ou pescoço.

Na base de dados, as amostras vocais estavam classificadas quanto à presença, intensidade e tipo de desvio vocal, a partir do JPA, que foi realizado de forma independente por três fonoaudiólogos especialistas em voz, com mais de dez anos de experiência no atendimento a indivíduos com distúrbios vocais.

Por ocasião da análise das amostras, os juízes foram treinados com 16 estímulos âncora (vogais sustentadas e frases) contendo quatro amostras de indivíduos com vozes saudáveis, quatro amostras de indivíduos com desvio vocal leve a moderado, quatro amostras de indivíduos com desvio vocal moderado e quatro amostras de indivíduos com desvio vocal grave. Os juízes foram instruídos a ouvir os estímulos âncora imediatamente antes de analisar as amostras vocais.

A Escala de Desvio Vocal – EDV (YAMASAKI *et al*., 2017) foi utilizada para avaliar a intensidade do desvio vocal (GG - grau geral, classificado em leve, moderado e intenso) e os graus de rugosidade (GR), soprosidade (GB) e tensão (GS). Para avaliação, os juízes escutavam as duas amost*ras (vogal sustentada [e] e as seis sentenças do CAPE-V), podendo ser repetidas por até três vezes.*

Após cada apresentação, os juízes realizaram as seguintes etapas no julgamento perceptivoauditivo (JPA): 1) Categorização das vozes em saudável ou disfônica, respondendo à pergunta “Você considera esta voz saudável ou disfônica?”; 2) Avaliação geral do GG, GR, GB e GS pelo CAPE-V; 3) Se a voz for considerada

disfônica, o avaliador deve indicar o tipo de desvio vocal predominante, incluindo rugosidade, soprosidade ou tensão.

Sendo assim, os valores da EDV foram utilizados para a classificação do GG. Os valores de 0 a 35,5 mm representavam indivíduos vocalmente saudáveis, valores de 35,6 a 50,5 mm representavam indivíduos com desvio leve, 50,6 a 90,5 mm representavam desvio moderado e valores > 90,5 mm representavam um desvio intenso. Ao final da sessão de avaliação perceptiva, 20% (50 sinais) das amostras foram repetidas aleatoriamente para analisar a confiabilidade da avaliação intrajuiz por meio do coeficiente kappa de Cohen. O juiz com o maior coeficiente (0,80) foi selecionado, o que indicou excelente confiabilidade/ concordância.

Dessa forma, após o JPA da qualidade vocal, foram selecionadas 44 amostras vocais dos sujeitos, equilibrando-as de acordo com o tipo de desvio, considerando os critérios de elegibilidade. A distribuição das vozes pode ser observada no Quadro 6. A amostra foi composta por oito vozes saudáveis (quatro amostras vocais de ambos os sexos) e 36 vozes desviantes (18 masculinas e 18 femininas) com diferentes graus e tipos de desvio vocal.

QUADRO 04: Distribuição quantitativa de gravações para as vozes saudáveis e disfônicas.

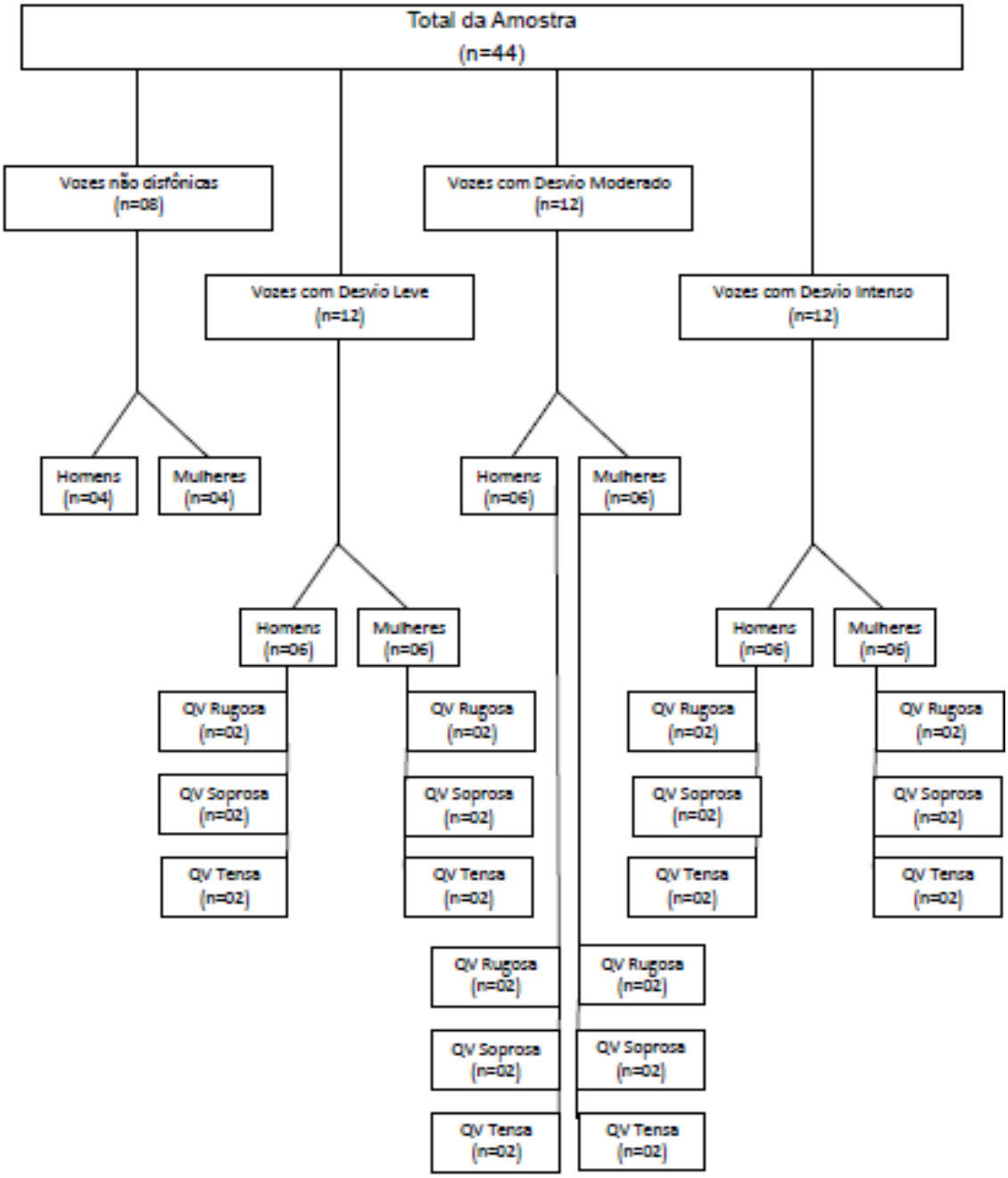
Sexo	**Vozes não disfônicas (VNQV)**	**Vozes rugosas (graus leve, moderado e intenso)**	**Vozes soprosas (graus leve, moderado e intenso)**	**Vozes tensas (graus leve, moderado e intenso)**
Homens	04	06	06	06
Mulheres	04	06	06	06
TOTAL: 44 gravações				

VNQV: Variabilidade normal da qualidade vocal.

Os sujeitos com vozes saudáveis apresentaram valores abaixo do ponto de corte na EDV, e ausência de alteração estrutural ou funcional na laringe. Os sujeitos disfônicos apresentaram valores acima do ponto de corte da EDV, e presença de lesões benignas das pregas vocais.

Segue o fluxograma construído, para melhor entendimento da distribuição das vozes:

FIGURA 02: Distribuição das vozes selecionadas.



FONTE: Dados da pesquisa.

Posteriormente, as amostras vocais selecionadas foram editadas a partir do software Sound Forge (versão 10.0; Sony, Tóquio, Japão). Foram inseridas as as seis

sentenças CAPE-V de cada locutor em um único arquivo .wav. Os sinais foram normalizados para obter uma padronização na saída de áudio entre -6 e 6 dB.

A extração das medidas acústicas foi realizada no software Praat, versão 5.3.77h, por meio do *script* – VoxMore (ABREU *et al*., 2023). O software gera um relatório com informações e imagens referentes aos valores de medidas acústicas relacionadas a f0, medidas de período, perturbação e ruído. As instruções sobre instalação do VoxMore podem ser encontradas em: https://github.com/abreusamuel/VoxMore. Foram extraídas 47 medidas acústicas, descritas no Quadro 4, a seguir, no tópico “definição das variáveis”.

4.4.2 Julgamento de atitudes

4.4.2.1 Instrumento para Julgamento de atitudes

O julgamento de atitudes associado a vozes disfônicas e não disfônicas, foi realizado a partir da escala diferencial semântica denominada “Escala de Julgamento de Atitudes Associadas a Vozes Disfônicas” (ANEXO A). A escala foi elaborada com base no conhecimento teórico dos métodos de mensuração da atitude linguística por meio da técnica do diferencial semântico de Osgood *et al*., 1957, a partir da tradução e adaptação cultural para o português brasileiro (EVANGELISTA *et al.,* 2022), realizadas a partir da seleção e tradução de 56 atributos/atitudes para julgamento (OSGOOD *et al*., 1957).

Na adaptação cultural, apresentou-se a lista de atributos a 40 graduandos em Fonoaudiologia, que cursavam as disciplinas correspondentes à área de voz, e solicitou-se que eles escrevessem pelo menos um antônimo de cada atributo em sua língua nativa. Este procedimento é uma forma de viabilizar a adaptação linguística e adaptação de termos em uma determinada língua (HEISE, 1970), possibilitando a etapa de seleção dos atributos que possuíam equivalência linguística ao nativo. Após análise estatística, foram selecionados 12 atributos/atitudes equivalentes, sendo eles: Simpático – Antipático; Agradável – Desagradável; Saudável – Doente; Extrovertido-Introvertido; Poderoso-Fraco; Autoritário-Submisso; Forte – Fraco; Competente – Incompetente; Resistente-Frágil; Calmo – Agitado; Seguro – Inseguro; e Independente - Dependente.

Esses 12 atributos e seus respectivos antônimos foram usados para compor Escala de Julgamento de Atitudes associadas a vozes disfônicas, a partir de uma chave de resposta escalar de seis pontos, em que o ouvinte faz a marcação de acordo com sua percepção, quanto mais positiva a valência mais próxima de seis é a marcação. Ela tem como base as três dimensões básicas das atitudes: (1) *avaliação*, que verifica se uma pessoa pensa positivamente ou negativamente sobre o tema da atitude; (2) *potência* que se relaciona com a força, medindo o quão poderoso é o sujeito; (3) *atividade,* que diz respeito ao fato do indivíduo ser considerado ativo ou passivo diante de situações da vida (OSGOOD *et al*., 1957), como demonstra o quadro 5.

QUADRO 05: Distribuição das atitudes por dimensões.

AVALIACÃO	Simpático-Antipático; Agradável- Desagradável; Saudável-Doente; Extrovertido – Introvertido;
POTÊNCIA	Poderoso – Fraco; Autoritário – Submisso; Forte - Fraco; Competente – Incompetente;
ATIVIDADE	Resistente – Frágil; Calmo – Agitado; Seguro – Inseguro; Independente – Dependente;

FONTE: Dados da pesquisa.

4.4.2.2 Procedimento para Julgamento de atitudes

A apresentação dos sinais vocais aos ouvintes foi realizada em sala de aula com nível de ruído inferior a 50 dB NPS. A sala acomodava 45 alunos. Os 152 juízes recrutados para participar deste estudo foram divididos em quatro grupos: três grupos de 40 juízes e um grupo de 32 juízes. Cada grupo passou por uma sessão para julgar

atitudes associadas a vozes disfônicas e saudáveis. A duração média da sessão de julgamento de atitudes foi de 50 minutos.

Inicialmente, foram apresentadas as seis sentenças CAPE-V de um falante saudável (cuja amostra não foi selecionada para esta pesquisa) usando um laptop e caixas e som portáteis. O objetivo desse procedimento foi verificar se a intensidade da apresentação dos estímulos era confortável e suficiente para os ouvintes. Este procedimento foi repetido duas vezes.

Após ouvir as seis frases CAPE-V emitidas por cada um dos 44 falantes, os ouvintes preencheram a Escala de Julgamento de Atitudes Associadas a Vozes Disfônicas e identificaram o quanto a voz de cada falante transmitia as impressões descritas. Para cada impressão (atributo), os números variavam em uma escala Likert de 1 (mais negativa) a 6 (mais positiva). Quanto mais negativa a impressão, mais próxima ela estaria do número “1”. Quanto mais positiva a impressão, mais próxima ela estaria do número “6”. Caso houvesse solicitação de um dos juízes, a amostra do locutor era repetida para escuta.

Ao final da apresentação dos 44 estímulos de fala, 20% da amostra (n=9) foram repetidas aleatoriamente utilizando o coeficiente de correlação intraclasse (CCI) para avaliar a confiabilidade teste-reteste dos ouvintes. De modo geral, os juízes obtiveram confiabilidade moderada no CCI (de 0,75 a 0,90), o que indica confiabilidade teste-reteste moderada.

4.5 DEFINIÇÃO DAS VARIÁVEIS:

4.5.1 Variáveis dependentes

- Julgamento do atributo/atitude do falante realizado pelos ouvintes: Simpático – Antipático; Agradável – Desagradável; Saudável – Doente; Extrovertido-Introvertido; Poderoso-Fraco; Autoritário-Submisso; Forte – Fraco; Competente – Incompetente; Resistente-Frágil; Calmo – Agitado; Seguro – Inseguro; e Independente - Dependente. Variável escalar;
- Valência do julgamento dos atributos do falante, positivo ou negativo. Variável nominal dicotômica.

4.5.2 Variáveis independentes

- JPA realizado pelos fonoaudiólogos especialistas: Grau geral, rugosidade, soprosidade e tensão. Variável numérica contínua;
- Medidas acústicas relacionadas à f_o, medidas de período, perturbação e ruído, e medidas cepstrais, descritas no Quadro 03. Variáveis numéricas contínuas.

QUADRO 06: Descrição das medidas acústicas extraídas na pesquisa.

MEDIDA ACÚSTICA	DESCRIÇÃO DA MEDIDA
f_o **média**	Média da f_o
f_o **SD**	Desvio Padrão da f_o.
f_o **mediana**	Mediana da f_o.
f_o **1ªquartil**	1º Quartil da f_o.
f_o **3ªquartil**	3º Quartil da f_o.
f_o **Min**	Mínima da f_o.
f_o **Max**	Máxima da f_o.
f_o **CV**	Coeficiente de variação da f_o.
Período Média	Média do período. Tornou-se um parâmetro importante na avaliação da soprosidade vocal
PSD	Densidade Espectral de potência de ruído.
LNPSD	Logaritmo do desvio padrão do período.
Jitter percep	Variação de ciclo na fo.
Jitter ABS	Perturbação média relativa, calculada como a diferença média absoluta entre um período.
Jitter RAP	Diferença média absoluta entre um período e a média dele e de seus dois vizinhos, dividida pelo período médio.
Jitter ppq5	Diferença média absoluta entre um período e a média dele e seus quatro vizinhos mais próximos, dividida pelo período médio.
Jitter DDP	Diferença média absoluta entre diferenças consecutivas entre períodos consecutivos, dividida pelo período médio.
Shimmer local	Diferença média absoluta entre as amplitudes de períodos consecutivos, dividida pela amplitude média.
Shimmer dB	Mudança da amplitude pico a pico. A unidade de medida em db. Corresponde ao logaritmo médio absoluto de base 10 da diferença entre as amplitudes de períodos consecutivos, multiplicado por 20.
Shimmer Apq3	A diferença absoluta média entre a amplitude do período e a média das amplitudes de seus vizinhos, dividida pela amplitude média.
Shimmer Apq5	Diferença média absoluta entre a amplitude de um período e a média das amplitudes dele e de seus quatro vizinhos mais próximos, dividida pela amplitude média.
Shimmer apq11	diferença média absoluta entre a amplitude de um período e a média da amplitude do intradorso e seus vizinhos mais próximos, dividida pela amplitude média".
Shimmer DDA	Diferença média absoluta entre diferenças consecutivas entre as amplitudes de períodos consecutivos

CPP	Proeminência do pico cepstral.
CPPS	Proeminência do pico cepstral suavizada
Declinio	Inclinação geral do espectograma
Tilt	Inclinação da linha de regressão do espectrograma.
GNE1	Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz.
GNE 2	Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz
GNE3	Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz.
H1H2	Razão da amplitude do primeiro e segundo harmônico. Caracteriza-se pela diferença entre as amplitudes do primeiro e segundo harmônicos no espectro.
H2A	Amplitude do segundo harmônico e apresenta-se diminuída com o aumento da rugosidade.
Hfno	Nível de energia relativa do ruído de alta frequência.
HNR media	Média da relação harmônico-ruído.
HNRD	Desvio Padrão da Relação Harmônico Ruído.
H1A1	Amplitude do primeiro harmômico
H1A3	Amplitude do primeiro harmônico em relação ao terceiro harmônico
H1H2	Razão da amplitude do primeiro e segundo harmônico.
SNL1	100–2600 Hz (nível de ruído espectral = 100–2600 Hz).
SNL2	100–3000 Hz (nível de ruído espectral = 100–3000 Hz).
SNL3	100–5100 Hz (nível de ruído espectral = 100–5100 Hz).
SNL4	100–8000 Hz (nível de ruído espectral = 100–8000 Hz).
SNL5	2600–5100 Hz (nível de ruído espectral = 2600–5100 Hz)
SNL6	5100–8000 Hz (nível de ruído espectral = 5100–8000 Hz).
AVI	Índice de variabilidade da amplitude. Torna-se importante medida e com alta sensibilidade na predição da rugosidade.
Correlação media	Média de correlação.
PA	Amplitude da frequência.
SFR	Planicidade espectral do sinal residual.

A caracterização das medidas acústicas e do JPA dos sinais utilizados no estudo para as vozes disfônicas e não disfônicas podem ser observados na Tabela 2. No entanto, considerando os objetivos do presente estudo, tais valores não serão discutidos e servirão de base para a validação dos dados e caracterização dos sinais.

TABELA 01: Diferença entre valores de parâmetros acústicos em vozes normais e desviadas.

Medidas Acústicas	**VALORES GRUPO**				**p-valor**
	NÃO DISFÔNICOS		**DISFÔNICOS**		
	MÉDIA	**DP**	**MÉDIA**	**DP**	
FALA_ f_o CPP	175,1900	38,38115	171,3857	46,17190	0,830

FALA_CPP	29,2900	2,56406	23,7617	5,68012	0,011*
FALA_CPPS	14,8663	1,23057	12,4583	2,98429	0,032*
Declinio	-21,8312	2,53951	-22,4400	4,25966	0,701
Tilt	-10,3450	,72036	-9,3734	1,61269	0,106
H1A1	-2,7850	1,74847	-3,8329	2,81328	0,322
H1A3	24,1813	3,50130	22,6553	7,12045	0,561
H1H2	5,0413	1,83463	5,6356	3,81993	0,673
f_o **media**	172,8250	42,77358	177,5954	43,15969	0,779
f_o **DP**	26,1588	7,07400	34,0866	19,05268	0,257
f_o **mediana**	172,2387	43,01256	180,0623	46,21322	0,664
f_o **q1**	159,0200	41,93070	163,1617	43,23980	0,807
f_o **q3**	187,9225	46,01148	196,2823	53,17408	0,684
f_o **CV**	15,4275	3,74223	18,7503	8,78229	0,304
f_o **min**	81,5237	6,58642	81,9303	16,48929	0,946
f_o **max**	375,8513	152,83816	365,6340	142,43350	0,857
PERIODO	0,00615	0,00163	0,00600	0,00152	0,803
PSD	0,00093	0,00029	0,00112	0,00049	0,289
LNPSD	-7,02688	0,35859	-6,85780	0,3700496	0,248
Jitter Loc	1,8913	0,40329	2,1446	0,82641	0,407
Jitter ABS	0,00012	0,00005	00,00013	0,000078	0,685
Jitter RAP	0,00738	0,001302	0,01006	0,005450	0,178
Jitter PPQ5	0,9325	0,23825	1,1089	0,47803	0,319
Jitter ddp	2,3038	0,48074	2,9626	1,42012	0,032*
Shimmer Loc	6,1388	1,23634	7,5974	2,35562	0,023*
Shimmer dB	0,6450	0,11161	0,7851	0,21235	0,016*
Shimmer APQ3	1,8488	0,43653	2,8854	1,30837	<0,001*
Shimmer APQ5	2,8138	0,74233	3,8566	1,47179	0,009*
ShimmerAPQ11	6,0987	1,72572	6,8094	1,75384	0,306
Shimmer dda	5,5450	1,31157	8,6566	3,92430	0,034*
AVI	2,5188	0,08774	2,5617	0,14066	0,415
Correlacao	0,9375	0,01832	0,9151	0,04661	0,193
GNE1	0,9438	0,02264	0,8694	0,09390	0,033*
GNE2	0,9375	0,03196	0,8517	0,10291	0,026*
GNE3	0,9175	0,04268	0,8223	0,10516	0,017*
Hfno	2,1350	0,15520	2,0477	0,28916	0,416
HNR media	17,4075	2,77413	15,8980	3,29703	0,238
HNR dp	7,1163	0,65733	6,9046	1,03109	0,584
HNRD	22,1925	4,39666	20,9869	4,05836	0,459
PA	0,7850	0,03703	0,7594	0,05368	0,210
SFR	-17,2325	0,52082	-17,2283	1,09209	0,992
SNL1_3Hz	19,3550	4,26410	19,8411	5,67058	0,821
SNL2_3Hz	13,6088	3,61371	15,2189	5,36087	0,426
SNL3_3Hz	26,0275	4,14415	26,0377	5,36735	0,996
SNL4_3Hz	24,7075	4,16777	24,8423	5,51628	0,949
SNL5_3Hz	12,6863	4,54442	13,6466	6,38093	0,690
SNL6_3Hz	4,0787	3,21468	7,5754	6,31539	0,138

Teste t-Student independente; significância p<0,05*.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma; **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico; **H1H2:** Razão da amplitude do primeiro e segundo harmônico; **fo média**: Média da Frequência Fundamental; **fo DP**: Desvio Padrão da Frequência Fundamental; **fo mediana**: Mediana da Frequência Fundamental;**fo q1:** 1º Quartil da Frequência Fundamental. **fo q3:** 3º Quartil da Frequência Fundamental; **fo CV:** Coeficiente de variação da Frequência Fundamental; **fo min:** Mínima da Frequência Fundamental; **F0 max**: Máxima da Frequência Fundamental; **PSD:** Densidade Espectral de potência de ruído; **LNPSD:** Logaritmo do desvio padrão do período; **Jitter LOC:** Média de perturbação relativa; **Jitter ABS:** perturbação média relativa; **Jitter RAP:** Perturbação média relativa em porcentagem; **Jitter; PPQ5:** Coeficiente médio de perturbação;.**Jitter DDP**: Diferença média absoluta entre períodos consecutivos; **Shimmer Loc:** Média de perturbação relativa;. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3; **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz; **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz; **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz;. **Hfno:** Nível de energia relativa do ruído de alta frequência; **HNR média:** Média da relação harmônico-ruído.; **HNRD:** Desvio Padrão da relação harmônico ruído; **PA:**

Amplitude da Frequência; **SFR:** Planicidade espectral do sinal residual; **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**); SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

Na Tabela 2 pode ser observada as médias das valências das atitudes dos ouvintes para as vozes disfônicas e não disfônicas. O objetivo dessa tabela é caracterizar o julgamento dos sinais, de modo que tais achados não serão discutidos nesse estudo.

TABELA 02: Diferença da valência média em cada atitude julgada pelos ouvintes.

ATITUDES	VALÊNCIA GRUPO				p-valor
	VOZ SAUDÁVEL		VOZ DESVIADA		
	MÉDIA	DP	MÉDIA	DP	
Valência Geral	4,5507	0,480672	2,9913	0,926197	<0,001*
Valência med	4,5171	0,385692	3,0198	0,841280	<0,001*
Agradável	4,6110	0,55074	2,5910	1,1251	<0,001*
Poderoso	4,4013	0,5133	2,6679	1,1409	<0,001*
Simpático	4,4292	0,5647	3,0861	,6359	<0,001*
Forte	4,5641	0,6105	2,7809	1,1588	<0,001*
Resistente	4,4851	0,6991	2,8278	1,1688	<0,001*
Extrovertido	4,4025	0,8126	2,8636	0,7779	<0,001*
Saudável	5,0085	0,6002	2,8261	1,4806	<0,001*
Autoritário	4,1797	0,5005	3,0403	0,8773	0,001*
Calmo	3,87116	0,5706	4,1257	0,62937	0,299
Seguro	4,6003	0,5810	3,10380	0,9292	<0,001*
Competente	4,7617	0,34308	3,49463	0,75748	<0,001*
Independente	4,8906	0,122068	2,829892	0,832746	<0,001*

Legenda: Teste t-Student independente; significância p<0,05*.

Com a finalidade de atender aos objetivos do estudo, as vozes foram alocadas em grupo, de acordo com o julgamento dos ouvintes: valência positiva e valência negativa.

4.6 ANÁLISE DOS DADOS

Os dados foram alocados em planilha digital e foi realizada análise estatística descritiva, por meio de medidas de frequência e tendência central, como média e desvio padrão, e análise inferencial. Considerou-se o intervalo de confiança de 95% ($p < 0, 05$) e utilizou-se o software SPPS para as análises.

Inicialmente, verificou-se a distribuição do conjunto de dados por meio do teste Komolgorov-Smirnov, que confirmou a normalidade dos dados e a possibilidade da realização de testes paramétricos.

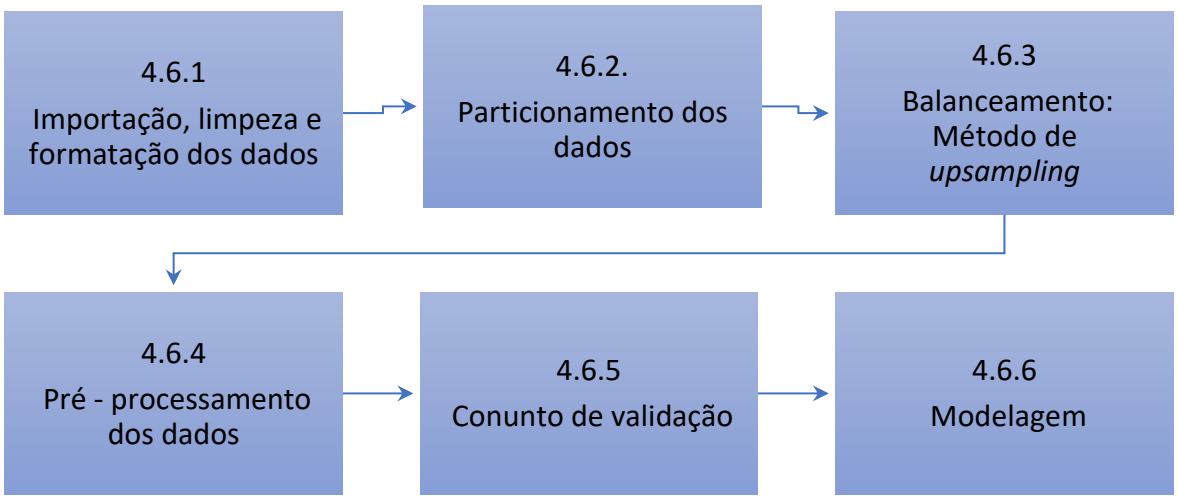
Utilizou-se o teste t-Student para amostras independentes com a finalidade de verificar se existiam diferenças nas medidas do JPA e nas medidas acústicas em vozes com valência positiva e negativa. O teste de correlação de Pearson foi utilizado para verificar se existe correlação entre o julgamento de atitudes dos ouvintes e as medidas do JPA e medidas acústicas.

A força das correlações foi classificada de acordo com o preconizado por Daniel (2009): 0.9 para mais ou para menos indica uma correlação muito forte; 0.7 a 0.9 positivo ou negativo indica uma correlação forte; 0.5 a 0.7 positivo ou negativo indica uma correlação moderada; 0.3 a 0.5 positivo ou negativo indica uma correlação fraca. 0 a 0.3 positivo ou negativo indica uma correlação desprezível.

Na sequência, foi desenvolvido um modelo de regressão linear para analisar a capacidade das medidas do JPA e das medidas acústicas em estimar a variabilidade no julgamento de atitudes dos ouvintes. Foram considerados os seguintes pré-requisitos para verificação da adequabilidade do modelo produzido: distribuição normal da variável dependente “julgamento de atitudes” para as amostras utilizadas; ausência de outliers (estatística do resíduo variando entre -2.351 e -4.185); ausência de multicolinearidade (tolerância das medidas selecionadas para o modelo variando entre 1.231 e 1.890); independência dos resíduos, com teste de Durbin-Watson de 1.406. O desempenho do modelo para explicar a variabilidade na percepção do GR será avaliado a partir do R^2. O R^2 é uma medida estatística que explica o quanto as variáveis do JPA e acústicas são capazes de explicar a variabilidade no julgamento de atitudes dos ouvintes. Adotou-se significância de 5% para interpretação dos modelos.

Foi desenvolvido ainda o modelo de predição das atitudes dos ouvintes baseado em ML, que foi capaz de destacar as medidas acústicas capazes de predizer a valência das atitudes dos ouvintes em relação às vozes disfônicas. As medidas acústicas extraídas formaram o conjunto de dados que embasaram a comparação de modelos de MA. O fluxograma para desenvolvimento do modelo é apresentado na Figura 03 e descrito a seguir.

FIGURA 03: Etapas para a análise dos dados.



FONTE: Dados da análise realizada na pesquisa.

4.6.1 Importação, limpeza e formatação dos dados

Foram realizadas etapas de importação, limpeza e formatação dos dados, por meio dos sistemas de codificações descritos neste tópico.

A limpeza dos dados foi realizada a fim de remover todas as medidas que apresentaram forte correlação entre si, com o objetivo de não repetir os dados, e a importação foi feita a partir do script a seguir:

QUADRO 07: Importação dos dados.

```
ℹ Using "','" as decimal and "'.'" as grouping mark. Use `read_delim()` for
more control.
Rows: 43 Columns: 82
── Column specification ────────────────────────────────────────────────────
Delimiter: ";"
chr (1): FALANTE
dbl (80): sexo, idade, laudo, GRAU, presença desvio, Tipo, I, J, K, L, M,
N,...
lgl (1): EAV

ℹ Use `spec()` to retrieve the full column specification for this data.
ℹ Specify the column types or set `show_col_types = FALSE` to quiet this
message.

# A tibble: 6 × 82
 FALANTE   sexo idade laudo EAV  GRAU `presença desvio` Tipo   I   J
 <chr>   <dbl> <dbl> <dbl> <lgl> <dbl>       <dbl> <dbl> <dbl> <dbl>
1 Falante 5   1   3   9 NA    3         2   3   1   1
2 Falante 41   1   2   5 NA    4         2   3   1   1
3 Falante 36   2   2   4 NA    4         2   2   1   1
4 Falante 28   1   2   5 NA    4         2   3   1   1
5 Falante 43   1   3   1 NA    4         2   4   1   1
6 Falante 4   1   2   7 NA    3         2   2   1   1
# ℹ 72 more variables: K <dbl>, L <dbl>, M <dbl>, N <dbl>, O <dbl>, P
<dbl>,
#   Q <dbl>, R <dbl>, S <dbl>, T <dbl>, U <dbl>, V <dbl>, W <dbl>, X <dbl>,
#   Y <dbl>, Z <dbl>, AA <dbl>, AB <dbl>, AC <dbl>, AD <dbl>, AE <dbl>,
#   AF <dbl>, AG <dbl>, AH <dbl>, AI <dbl>, AJ <dbl>, AK <dbl>, AL <dbl>,
#   AM <dbl>, AN <dbl>, AO <dbl>, AP <dbl>, AQ <dbl>, AR <dbl>, AS <dbl>,
#   AT <dbl>, AU <dbl>, AV <dbl>, AW <dbl>, AX <dbl>, AY <dbl>, AZ <dbl>,
#   BA <dbl>, BB <dbl>, BC <dbl>, BD <dbl>, BE <dbl>, BF <dbl>, BG <dbl>, …
```

FONTE: Dados da análise realizada na pesquisa.

Após importação, os dados foram formatados e analisados, observando-se os melhores ajustes do modelo, para seleção dos resultados mais robustos, seguindo-se as etapas sugeridas no quadro 03.

A formatação refere-se ao processo de preparar os dados antes de alimentá-los no modelo de Aprendizado de Máquina, dessa forma, foi necessária conversão das variáveis para o formato numérico, através do tipo fator R (script 2).

QUADRO 08: Importação dos dados.

```
df1$I = factor(df$I, levels = 1:2, labels = c("Não", "Sim"))

#df1$J = factor(df$J, levels = 1:2, labels = c("Não", "Sim"))

df1$sexo = factor(df$sexo, levels = 1:2, labels = c("M", "F"))
```

FONTE: Dados da análise realizada na pesquisa.

4.6.2 Particionamento dos dados

O conjunto de dados foi particionado em um conjunto de treinamento com 33 observações, utilizado para o ajuste dos modelos, e um conjunto de teste com 10 observações, que foi utilizado para avaliar os modelos.

A utilização de uma amostragem estratificada garantiu a mesma proporção de exemplos nas duas classes da variável alvo, para os conjuntos de treinamento e de teste. Segue o sistema de codificação:

QUADRO 09: Sistema de codificação - particionamento.

```
set.seed(1313)
df_split <- initial_split(df1, prop = 34/43, strata = I)
train_data <- training(df_split)
test_data <- testing(df_split)
```

FONTE: Dados da análise realizada na pesquisa.

4.6.3 Balanceamento dos dados

Foi utilizado o método de *upsampling* para o balanceamento das classes da variável alvo, que é conceituado como um método de processamento digital de sinais que aumenta artificialmente a taxa de amostragem em N vezes, inserindo um número N -1 de zeros entre as amostras originais do sinal e submetendo o conjunto obtido por um filtro de reconstrução, que nada mais é que um filtro do tipo passa-baixa (BRUCE, 2019).

Neste método, exemplos da classe minoritária são replicados aleatoriamente até que as classes da variável alvo contenham o mesmo número de exemplos. Desta forma, cada classe da variável alvo no conjunto de treinamento continha 22 exemplos após o balanceamento.

QUADRO 10: Resumo dos dados na etapa de balanceamento.

NOME	NEW_TRAIN_DATA
NÚMERO DE LINHAS	44
NÚMERO DE COLUNAS	48
FREQUÊNCIA DO TIPO DE COLUNA	
FATOR	2
NÚMERICO	46
VARIÁVEL DOS GRUPOS	None

FONTE: Dados da pesquisa.

4.6.4 Pré-processamento dos dados

O pré-processamento dos dados consistiu na imputação dos valores ausentes utilizando o método dos vizinhos mais próximos (K-NN), considerando K=3. Em seguida, as variáveis numéricas foram padronizadas de modo a terem média igual a zero e desvio-padrão igual a um. Diante do balanceamento dos dados, o K-NN estimou os valores ausentes com base na informação de instâncias semelhantes, através da inferência. Dessa forma, para cada instância com o valor ausente, o algoritmo K-NN calcula a distância entre as instâncias no conjunto de dados, por meio da métrica de distância euclidiana ao quadrado (IMANDOUS T.; BOLANDRAFTAR, 2013).

A variável sexo foi codificada como uma variável *dummy*, sendo o sexo feminino codificado como o valor “1” (um) e o sexo masculino codificado como o valor “0” (zero). A *dummy* também é conhecida como variável indicadora ou binária e torna-se uma abordagem comum para representar variáveis categóricas (SCIKITLEARN, 2021).

Por fim, preditores altamente correlacionados (r >0,85) com algum outro preditor/variável foram removidos, a fim de se evitar multicolinearidade. Segue o sistema de codificação:

QUADRO 11: Sistema de codificação – pré-processamento.

```
df_rec <- recipe(I ~ ., data = new_train_data) %>%
 step_impute_knn(all_predictors(), neighbors = 3) %>% # imputa valores
ausentes por K-NN
 # step_naomit(everything(), skip = TRUE) %>% # remove linhas que contém
NA ou NaN
 #                step_interact(terms                =                ~
all_numeric_predictors():all_numeric_predictors()) %>%
 # step_log( # Transformação log: y = log(x)
 #  idade,
 #  q05
 #  ) %>%
 # step_YeoJohnson( # Transformação Yeo-Johnson
 #  idade,
 #  q05
 #  ) %>%
 # step_poly(all_numeric_predictors(), degree = 3) %>%
 step_normalize(all_numeric(),   -all_outcomes())   %>%   #   normaliza
variáveis numéricas para terem média 0 e variância 1
 step_dummy(all_nominal(), -all_outcomes()) %>% # converte variáveis
qualitativas em variáveis dummy
 #  step_nzv(all_numeric(),  -all_outcomes())  %>%  #  remove  variáveis
numéricas que têm variância 0
 step_corr(all_predictors(), threshold = 0.85, method = "spearman") #
remove preditores que tenham alta correlação com algum outro preditor
```

FONTE: Dados da análise realizada na pesquisa.

4.6.5 Conjunto de validação

Foi utilizado o método K-*fold cross-validation* para construir um conjunto de validação com k *folds*. Consideramos um procedimento com K=11 *folds* repetido 30 vezes. Os dados são particionados em 11 partes utilizando amostragem estratificada. Em cada iteração, os modelos são ajustados em um conjunto de treinamento com composto por 10 dessas partes (40 observações) e avaliado em um conjunto de teste composto por 1 dessas partes (4 observações). Esse processo foi repetido 30 vezes de modo a abranger um amplo espaço de possibilidades de particionamento, dado que dispomos de poucas observações. Esse procedimento foi utilizado para obter os valores ótimos dos hiperparâmetros dos modelos. Segue o sistema de codificação:

QUADRO 12: Sistema de codificação – validação.

```
set.seed(1313)
cv_folds <- vfold_cv(new_train_data,
           v = 11,
           strata = I,
           repeats = 30)
```

FONTE: Dados da análise realizada na pesquisa.

4.6.6 Técnica de Modelagem

Foram considerados 12 modelos de ML mais utilizados para a classificação de dados, a saber: Regressão Logística penalizada (Elasticnet), SVM Linear, SVM Polinomial, SVM Kernel, K-Vizinhos mais próximos (K-NN), Árvore de Classificação, Floresta Aleatória (Random Forest), Redes Neurais MLP (Multilayer Perceptron), Redes Neurais MLP combinadas através de Bagging (Bagged MLP), Naive Bayes, Análise Discriminante Linear (LDA) e Análise Discriminante Regularizada (RDA) (MELLEY L. E. & SATALOFF R. T., 2022; BERTELSEN C. *et al*., 2018; BAINBRIDGE K. E. *et al*., 2017; RITCHINGS R. T. *et al*., 2002).

Os hiperparâmetros dos modelos foram otimizados no processo de validação cruzada repetida. A busca pelos valores ótimos dos hiperparâmetros se deu através de um processo de busca em um *grid* aleatório de valores definido através de um esquema de hipercubo latino, visando preencher adequadamente o espaço de valores dos hiperparâmetros. Segue o sistema de codificação:

QUADRO 13: Sistema de codificação – modelagem.

```
enet_spec <- logistic_reg(penalty = tune(),
             mixture = tune()) %>% # Elastic net
 set_engine(engine = "glmnet", standardize = FALSE) %>%
 set_mode("classification")

svm_spec1 <- svm_linear(cost = tune(),
             margin = tune()) %>% # Linear SVM
 set_engine(engine = "kernlab") %>%
 set_mode("classification")

svm_spec2 <- svm_rbf(cost = tune(),
           rbf_sigma = tune(),
           margin = tune()) %>% # Gaussian kernel SVM
 set_engine(engine = "kernlab") %>%
 set_mode("classification")

svm_spec3 <- svm_poly(cost = tune(),
           margin = tune()) %>% # Polynomial kernel SVM
 set_engine("kernlab") %>%
 set_mode("classification")

knn_spec <- nearest_neighbor(neighbors = tune(),
               weight_func = tune(),
               dist_power = tune()) %>% # K-NN
 set_mode("classification") %>%
 set_engine("kknn")

dt_spec <- decision_tree(cost_complexity = tune(),
             min_n = tune(),
             tree_depth = tune()) %>% # Decision tree
 set_engine(engine = "rpart") %>%
 set_mode("classification")

rf_spec <- rand_forest(mtry = tune(),
            min_n = tune(),
            trees = tune()) %>% # Random forest
 set_engine(engine = "ranger", importance = "impurity") %>%
 set_mode("classification")

mlp_spec <- mlp(hidden_units = tune(),
        penalty = tune(),
        epochs = tune()) %>% # MLP
 set_engine("nnet") %>%
 set_mode("classification")

bmlp_spec <- bag_mlp(hidden_units = tune(),
           penalty = tune(),
           epochs = tune()) %>% # Bagged MLP
 set_engine("nnet") %>%
 set_mode("classification")

xgb_spec <- boost_tree(tree_depth = tune(),
            learn_rate = tune(),
            loss_reduction = tune(),
            min_n = tune(),
            sample_size = tune(),
            trees = tune(),
```

```
                  mtry = tune()) %>% # Boosting trees
       set_engine(engine = "xgboost") %>%
       set_mode("classification")

      nbayes_spec <- naive_Bayes(smoothness = tune(),
                    Laplace = tune()) %>% # Naive Bayes
       set_engine("naivebayes") %>%
       set_mode("classification")

      lda_spec <- discrim_linear() %>% # Linear discriminant analysis
       set_engine("MASS") %>%
       set_mode("classification")

      rda_spec <- discrim_regularized(frac_common_cov = tune(),
                      frac_identity = tune()) %>% # Reg. discriminant
analysis
       set_engine(engine = "klaR") %>%
       set_mode("classification")
```

FONTE: Dados da análise realizada na pesquisa.

Segue o sistema de codificação, a partir da otimização dos hiperparâmetros:

QUADRO 14: Sistema de codificação – validação hiperparâmetros.

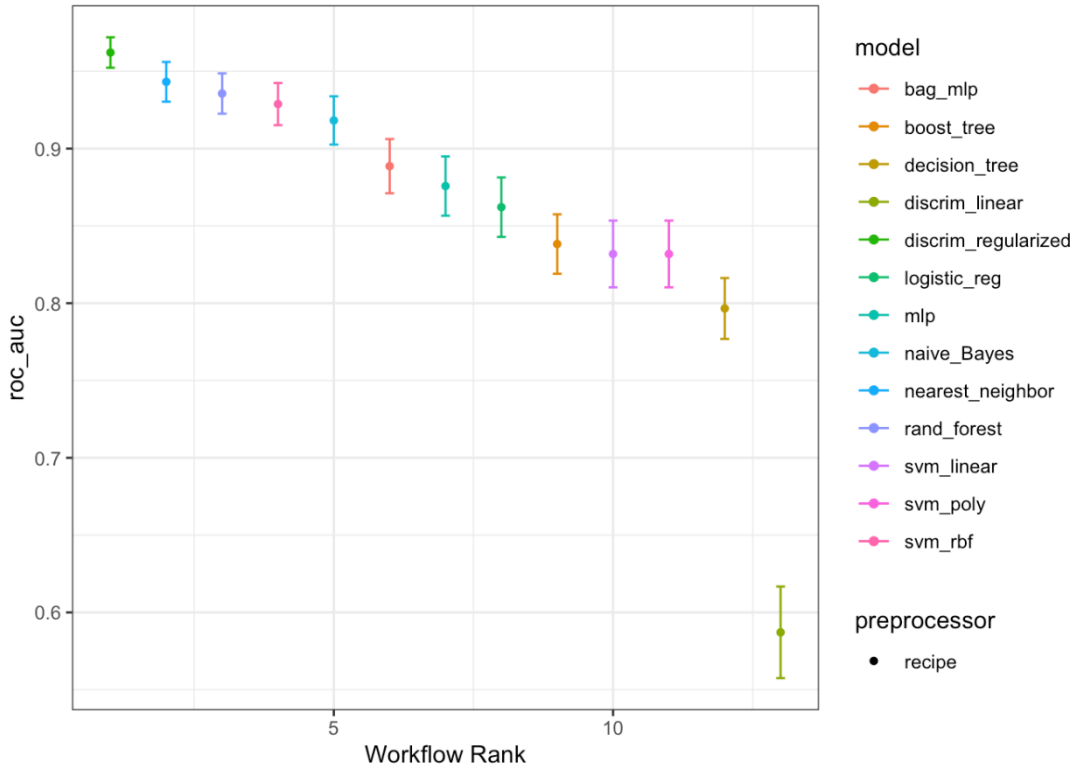
```
      cores <- parallel::detectCores()
      cl <- makePSOCKcluster(cores)
      registerDoParallel(cl)

      race_ctrl = control_race(
       save_pred = TRUE,
       parallel_over = "resamples",
       save_workflow = TRUE
      )
      race_results = wf %>%
       workflow_map(
        "tune_race_anova",
        # "tune_race_win_loss",
        seed = 1313,
        resamples = cv_folds,
        grid = 30,
        control = race_ctrl )
```

FONTE: Dados da análise realizada na pesquisa.

Para a análise dos dados, foi extraída a área sob a curva ROC no conjunto de validação de acordo com o melhor conjunto de hiperparâmetros obtidos para cada modelo. Em seguida, foi extraída a acurácia no conjunto de validação de acordo com o melhor conjunto de hiperparâmetros obtido para cada modelo. Na Figura 5, pode

ser observado o resultado da área sob a curva ROC para cada modelo de ML no conjunto de validação, utilizando o melhor conjunto de hiperparâmetros.

FIGURA 04: Área sob a curva ROC dos modelos de machine learning considerados na pesquisa.



Legenda: REG_DISC: Análise Discriminante Regularizada; **K_NN:** K- Vizinhos mais próximos; **RAND_FORES:** Floresta Aleatória; **MPL:** Redes Neurais MPL; **BAG_MPL:** Redes Neurais combinadas através de Bagging; **LIN_SVM:** SVM Linear; **POLY_SVM:**SVM Polimonial; **EL_NET:** Regressão Logística Penalizada; **LIN_DISCR:** Análise Discriminante Regularizada.

Dessa forma, os modelos de ML foram ordenados de acordo com o resultado da AUC (Quadro 07).

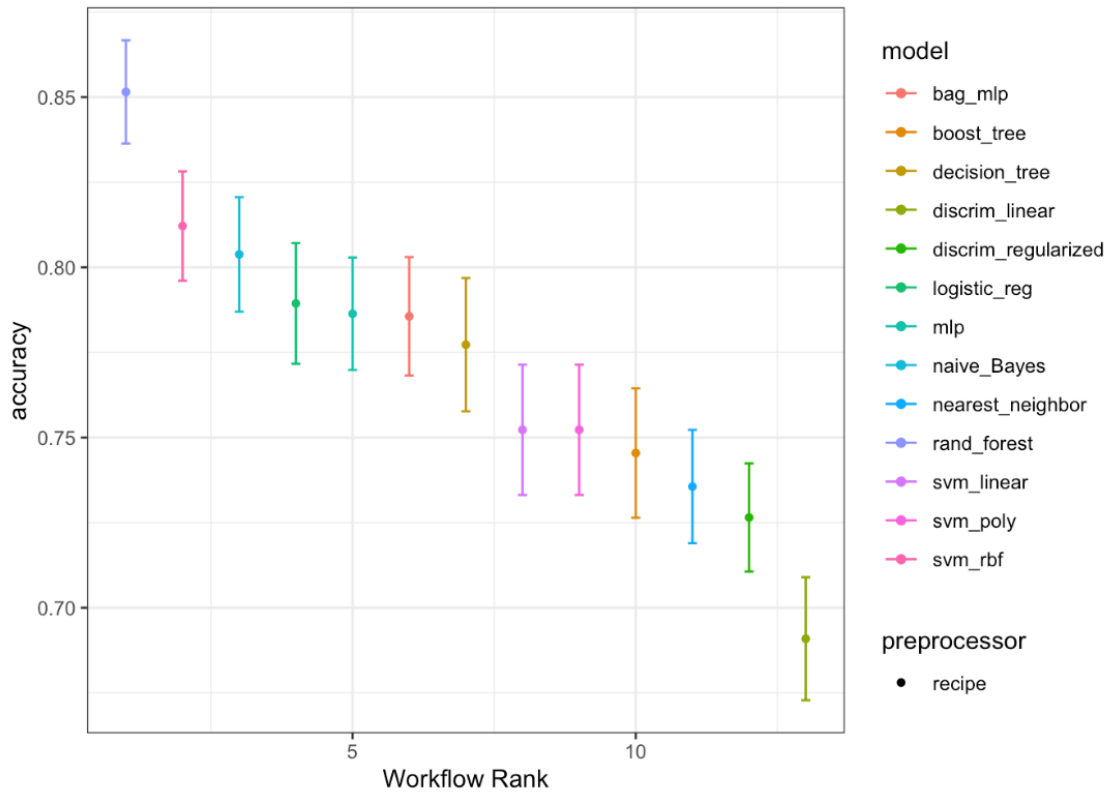
QUADRO 15: Ranqueamento dos 12 modelos de machine learning investigados na etapa de validação, considerando-se os valores de área sob a curva ROC.

RANKING	MODELO	AUC	DP
1	Análise Discriminante Regularizada (RDA)	0.962	0.005
2	K-vizinhos mais Próximos (K-NN)	0.943	0.007
3	Floresta Aleatória	0.936	0.007
4	SVM Kermel	0.929	0.008
5	Naïve Bayes	0.918	0.009
6	Redes Neurais MLP combinadas através de Bagging	0.889	0.010
7	Redes Neurais MLP	0.876	0.011
8	Regressão Logística	0.862	0.011
9	SVM Linear	0.832	0.013
10	SVM Polimonial	0.832	0.013
11	Árvore de Classificação	0.797	0.012
12	Análise Discriminante Linear (LDA)	0.797	0.018

Legenda: **AUC**: área under curve; **DP**: desvio-padrão.

Em seguida, foi extraída a acurácia no conjunto de validação de acordo com o melhor conjunto de hiperparâmetros obtidos para cada modelo.

FIGURA 05: Acurácia com os modelos de aprendizagem considerados na pesquisa.



Legenda: REG_DISC: Análise Discriminante Regularizada; **K_NN:** K- Vizinhos mais próximos; **RAND_FORES:** Floresta Aleatória; **MPL:** Redes Neurais MPL; **BAG_MPL:** Redes Neurais combinadas através de Bagging; **LIN_SVM:** SVM Linear; **POLY_SVM:**SVM Polimonial; **EL_NET:** Regressão Logística Penalizada; **LIN_DISCR:** Análise Discriminante Regularizada.

A partir da curva ROC, foi realizada seleção do melhor conjunto de hiperparâmetros para cada modelo, que foram utilizados para reajustar os modelos do conjunto de treinamento completo para, em seguida, se obter predições das classes da variável alvo (valência positiva ou negativa das atitudes dos ouvintes) no conjunto de teste. Foram calculadas as seguintes medidas no conjunto de teste: acurácia, acurácia balanceada, área sob a curva ROC, F-measure, precision, recall, especificidade e Kappa.

Na Tabela 3 são encontrados os valores das medidas de desempenho para os 12 modelos de ML investigados nesta pesquisa.

TABELA 03: Valores dos modelos ajustados com os conjuntos de hiperparâmetros, selecionados de acordo com a área sob a curva ROC.

MÉTODO	ACURÁCIA	ACURÁCIA BALANCEADA	ÁREA SOB CURVA ROC	*F-MEASURE*	PRECISÃO	RECALL	ESPECIFICIDADE	KAPPA
ANÁLISE DISCRIMINANTE REGULARIZADA (RDA)	0.8	0.8571	0.8333	1	0.7143	1	0.6	1
SVM KERNEL	0.9	0.9286	0.9231	1	0.8571	1	0.7826	0.9524
NAIVE BAYES	0.9	0.9286	0.9231	1	0.8571	1	0.7826	0.9524
K-VIZINHOS MAIS PRÓXIMOS (K-NN)	0.7	0.5952	0.8	0.75	0.8571	0.3333	0.2105	0.8095
FLORESTA ALEATÓRIA	0.8	0.7619	0.8571	0.8571	0.8571	0.6667	0,5238	0.8571
REDES NEURAIS MLP	0.7	0.7857	0.7253	1	0.5714	1	0.4444	0.7143
REDES NEURAIS MLP COMBINADAS ATRAVÉS DE BAGGING	0.6	0.619	0.6667	0.8	0.5714	0.6667	0.2	0.6667
XGBOOST	0.7	0.5	0.8235	0.7	1	0	0	0.5
SVM LINEAR	0.6	0.619	0.6667	0.8	0.5714	0.6667	0.2	0.8571
SVM POLINOMIAL	0.6	0.619	0.6667	0.8	0.5714	0.6667	0.2	0.8571
REGRESSÃO LOGÍSTICA PENALIZADA	0.7	0.5	0.8235	0.7	1	0	0	0.7143
ÁRVORE DE CLASSIFICAÇÃO	0.8	0.7619	0.8571	0.8571	0.8571	0.6667	0.5238	0.8571
ANÁLISE DISCRIMINANTE LINEAR (LDA)	0.6	0.7143	0.6	1	0.4286	1	0.3103	0.619

Foram utilizados os seguintes valores de referência para a categorização das métricas de acurácia, recall, especificidade e f1-Score e Precisão na avaliação dos modelos: excelente (> 0,90), bom (0,80 -0,90), aceitável (0,70-0,80), ruim (0,60 -0,70) e sem capacidade de discriminação aceitável (<0,60). Para o kappa, usamos a (79) seguinte interpretação para os valores: falta de concordância (<0), concordância ruim (0-0,19); concordância leve (0,20-0,39); concordância moderada (0,40-0,59); concordância substancial (0,60-0,79); e concordância quase perfeita (0,80-1,00) (HEMMERLING *et al*., 2016; BEHROOZMAND R. *et al*., 2007; LANDIS & KOCH, 1977).

Sendo assim, foram analisados os resultados de desempenho dos modelos de ML utilizados na classificação das vozes avaliadas positiva ou negativamente pelos ouvintes. Na análise dos modelos, os resultados de acurácia, sensibilidade, especificidade e f1-Score menores que 0,70 foram usados como critérios de exclusão para este estudo. Valores inferiores a 0,70 são considerados modelos ruins e sem capacidade de discriminação aceitável para detectar vozes disfônicas e não disfônicas (YER *et al*., 1991). Semelhante ao coeficiente Kappa, os modelos que obtiverem resultados abaixo de 0,40 serão descartados. Dessa forma, apenas o Kernel SVM e o Naïve Bayes cumpriram os pré-requisitos e serão, portanto, analisados na seção de resultados.

Feita a seleção do modelo de predição a ser utilizado, as medidas acústicas extraídas das vozes disfônicas e não disfônicas, bem como a valência, foram selecionadas pelos pesquisadores para observação da predição.

5 RESULTADOS

Os resultados serão apresentados em subseções, divididas didaticamente, para facilitar o fluxo de leitura e compreensão dos leitores. Dessa forma, serão consideradas as seguintes subseções: diferenças nas medidas acústicas e do JPA entre vozes com valência positiva e negativa; diferenças nas medidas acústicas e do JPA entre vozes com valência positiva e negativa nas dimensões avaliação, potência e atividade; correlação entre atitudes dos ouvintes e as medidas acústicas e do JPA; modelos de predição da variabilidade no julgamento de atitudes dos ouvintes a partir das medidas acústicas e do JPA; modelo baseado em ML para predição da valência das atitudes dos ouvintes em relação às vozes disfônicas e não disfônicas de falantes nativos do português brasileiro, a partir das medidas acústicas.

5.1 DIFERENÇAS NAS MEDIDAS ACÚSTICAS E DO JPA ENTRE VOZES COM VALÊNCIA POSITIVA E NEGATIVA

Houve diferença nos valores da EAV-GG, EAV-R, EAV-S e EAV-T entre as vozes avaliadas com valência positiva e negativa (Tabela 4). Vozes avaliadas negativamente tinham maiores valores em todas as medidas do JPA.

TABELA 04: Comparação das médias da EAV em função da valência das atitudes das vozes julgadas.

Variáveis	VALÊNCIA				p-valor
	NEGATIVA (1)		POSITIVA (2)		
	Média	DP	Média	DP	
EAV - GG	68,519	21,6758	32,269	12,1099	<0,001*
EAV-R	58,258	22,8024	26,923	13,8937	<0,001*
EAV-S	47,016	30,9958	11,885	14,5733	<0,001*
EAV-T	43,261	26,7051	29,269	10,6255	0,038*

*Valores significativos (p<0,005) – Teste T Student para amostras independentes
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; GR:** grau de rugosidade**; GS:** grau de soprosidade**; GT:** grau de tensão**.**

Observou-se que as medidas acústicas apresentaram diferença estatística significante em vozes com valência positiva e negativa, sendo que os valores de CPP-CS, CPPS-CS, Spectral decline-CS, Spectral tilt-CS, foram maiores para

vozes com valência positiva do que negativa, e as medidas JitterRAP-CS, Jitter PPQ5-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerrdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, Shimmerdda-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS tiveram valores maiores para vozes avaliadas negativamente pelos ouvintes (Tabela 05).

TABELA 05: Comparação das médias dos parâmetros acústicos em função da valência das atitudes das vozes julgadas.

Variáveis	VALÊNCIA				p-valor
	NEGATIVA (1)		POSITIVA (2)		
	Média	DP	Média	DP	
CPP-CS	23,2783	5,81264	28,2792	3,35853	0,001*
CPPS-CS	11,9783	2,95344	15,0477	1,03087	<0,001*
Spectral decline-CS	-23,0383	4,24190	-20,6846	2,77510	0,038*
Spectral tilt-CS	-9,2013	1,67742	-10,3685	0,59526	0,002*
H1A1-CS	-3,7386	3,08230	-3,3985	1,37352	0,706
H1A3-CS	23,4452	7,49677	21,8323	3,84085	0,469
H1H2-CS	5,9941	4,03103	4,4700	1,62145	0,198
fo mean-CS	181,5810	43,36348	165,4623	40,20911	0,260
F_0SD-CS	35,1837	20,11969	26,6762	7,67025	0,149
fo median-CS	184,4100	46,38053	165,2146	41,06530	0,205
F_0q1-CS	166,6853	43,99148	152,4815	38,76861	0,320
F_0q3-CS	201,3653	53,50682	179,4077	44,73029	0,203
F_0CV-CS	18,8417	9,14646	16,4946	5,13371	0,392
F_0min-CS	79,8430	11,64598	86,4969	20,84554	0,187
F_0max-CS	362,3630	140,28737	379,4700	152,90914	0,723
PER-CS	0,0058	0,00149	0,00642	0,0015	0,277
PSD-CS	0,001149	0,00052	0,00096	0,00026	0,226
LNPSD-CS	-6,8481	0,3960	-6,9840	0,2922	0,274
JitterLoc-CS	2,2037	0,85846	1,8523	0,43645	0,171
JitterABS-CS	0,00013	0,000081	0,000124	0,000056	0,703
JitterRAP-CS	0,01060	0,005648	0,00715	0,001725	0,004*
JitterPPQ5-CS	1,1460	0,49688	0,9146	0,24558	0,048*
Jiterddp-CS	3,1087	1,46657	2,2200	0,53825	0,006*
ShimmerLoc-CS	7,7833	2,44050	6,2708	1,28270	0,011*
ShimmerdB-CS	0,8003	,22186	0,6638	0,11244	0,011*
ShimmerAPQ3-CS	3,0273	1,35065	1,9200	0,45731	<0,001*
ShimmerAPQ5-CS	3,9667	1,54455	2,9608	0,72039	0,006*
ShimmerAPQ11-CS	6,7673	1,72962	6,4692	1,85194	0,614
Shimmerdda-CS	9,0830	4,05081	5,7577	1,37044	<0,001*
AVI-CS	2,5773	0,14215	2,4992	0,09004	0,076
Autocorrelation-CS	0,9130	0,04928	0,9338	0,02063	0,151
GNE1000Hz-CS	0,8613	0,09871	0,9338	0,02755	<0,001*
GNE2000Hz-CS	0,8420	0,10772	0,9269	0,03326	<0,001*
GNE3000Hz-CS	0,8110	0,10787	0,9069	0,04733	<0,001*
Hfno-CS	2,0517	0,31131	2,0923	0,14013	0,656
HNRmean-CS	15,9903	3,46130	16,6138	2,69784	0,567
HNRDP-CS	6,9103	1,07892	7,0215	,68275	0,734
HNRD-CS	20,8677	3,93472	22,0038	4,51000	0,410
PA-CS	0,7550	0,05425	0,7854	0,03886	0,046*
SFR-CS	-17,1837	1,16289	-17,3338	0,50227	0,658
SNL1-CS	18,8860	5,57213	21,7462	4,54143	0,111
SNL2-CS	14,5437	5,46943	15,7862	4,12199	0,468

SNL3-CS	25,1717	5,32363	28,0300	4,12023	0,093
SNL4-CS	23,9503	5,46373	26,8177	4,24284	0,100
SNL5-CS	12,6027	6,28902	15,4646	5,11234	0,156
SNL6-CS	7,3683	6,65162	5,9015	4,13749	0,468

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.
Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma; **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico; **H1H2:** Razão da amplitude do primeiro e segundo harmônico; **fo média**: Média da Frequência Fundamental; **fo DP**: Desvio Padrão da Frequência Fundamental; **fo mediana**: Mediana da Frequência Fundamental;**fo q1:** 1º Quartil da Frequência Fundamental. **fo q3:** 3º Quartil da Frequência Fundamental; **fo CV:** Coeficiente de variação da Frequência Fundamental; **fo min:** Mínima da Frequência Fundamental; **F0 max**: Máxima da Frequência Fundamental; **PSD:** Densidade Espectral de potência de ruído; **LNPSD:** Logaritmo do desvio padrão do período; **Jitter LOC:** Média de perturbação relativa; **Jitter ABS:** perturbação média relativa;. **Jitter RAP:** Perturbação média relativa em porcentagem; **Jitter; PPQ5:** Coeficiente médio de perturbação; **Jitter DDP**: Diferença média absoluta entre períodos consecutivos; **Shimmer Loc:** Média de perturbação relativa;. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3; **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz; **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz; **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz;. **Hfno:** Nível de energia relativa do ruído de alta frequência; **HNR média:** Média da relação harmônico-ruído.; **HNRD:** Desvio Padrão da relação harmônico ruído; **PA:** Amplitude da Frequência; **SFR:** Planicidade espectral do sinal residual; **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**); SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

5.2 DIFERENÇAS NAS MEDIDAS ACÚSTICAS E DO JPA ENTRE VOZES COM VALÊNCIA POSITIVA E NEGATIVA NAS DIMENSÕES AVALIAÇÃO, POTÊNCIA E ATIVIDADE.

Nós observamos diferenças em todas as medidas do JPA entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão avaliação: agradabilidade, simpatia, saúde e extroversão (Tabela 06). Indivíduos com vozes avaliadas como mais agradáveis, mais simpáticas, mais saudáveis e mais extrovertidas apresentaram menores valores na EAV-GG, EAV-R, EAV-S e EAV-T em relação às vozes avaliadas com valência negativa nessas atitudes.

TABELA 06: Comparação das médias referente ao julgamento perceptivo auditivo das vozes em relação às atitudes da dimensão avaliação: agradabilidade, simpatia, saúde e extroversão.

Variáveis	AGRADABILIDADE				p-valor
	AGRADÁVEL		DESAGRADÁVEL		
	Média	DP	Média	DP	
EAV - GG	32,958	13,4966	67,128	22,5025	<0,001*
EAV-R	27,375	17,1969	57,109	22,6750	<0,001*
EAV-S	15,500	15,7826	44,563	32,4106	0,005*
EAV-T	25,417	10,9416	44,269	25,4788	0,018*
	SIMPATIA				p-valor
	SIMPÁTICO		ANTIPÁTICO		
	Média	DP	Média	DP	
EAV - GG	30,650	12,3537	65,797	22,6779	<0,001*
EAV-R	25,100	14,2766	56,029	23,1474	<0,001*
EAV-S	8,900	10,5956	44,794	30,9981	<0,001*
EAV-T	30,700	7,8429	41,606	26,4404	0,041*
	SAÚDE				p-valor
	SAUDÁVEL		DOENTE		
	Média	DP	Média	DP	
EAV - GG	37,763	12,5868	73,044	21,9862	<0,001*
EAV-R	30,763	14,9558	62,860	22,1362	<0,001*
EAV-S	19,947	18,9553	49,320	33,4751	0,001*
EAV-T	27,211	16,0540	48,184	25,1288	0,003*
	EXTROVERSÃO				p-valor
	EXTROVERTIDO		INTROVERTIDO		
	Média	DP	Média	DP	
EAV - GG	31,591	12,1857	66,548	22,5859	<0,001*
EAV-R	25,773	13,2540	56,742	23,2100	<0,001*
EAV-S	10,182	14,8497	45,455	30,7234	<0,001*
EAV-T	28,500	11,0815	42,670	26,0159	0,016*

*Valores significativos (p<0,005) – Teste T Student para amostras independentes
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; GR:** grau de rugosidade**; GS:** grau de soprosidade**; GT:** grau de tensão**.**

Nós observamos diferenças em todas as medidas do JPA entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão potência: força e autoridade (Tabela 07). Indivíduos avaliados positivamente quanto às atitudes de força e autoridade apresentaram menores valores na EAV-GG, EAV-R, EAV-S e EAV-T em relação às vozes avaliadas com valência negativa nessas atitudes. Por sua vez, houve diferença apenas nas medidas EAV-GG, EAV-R, EAV-S entre os indivíduos avaliados com valência positiva e negativa nas atitudes poder e competência. Sujeitos julgados positivamente quanto às atitudes de poder e competência apresentaram menores valores na EAV-GG, EAV-R e EAV-S, em relação aos sujeitos avaliados negativamente nessas atitudes.

TABELA 07: Comparação das médias referente ao julgamento perceptivo-auditivo das vozes em relação às atitudes da dimensão potência: poder, força, autoridade e competência.

Variáveis	PODER				p-valor
	PODEROSO (1)		FRACO (2)		
	Média	DP	Média	DP	
EAV - GG	32,269	12,1099	68,519	21,6758	<0,001*
EAV-R	26,923	13,8937	58,258	22,8024	<0,001*
EAV-S	11,885	14,5733	47,016	30,9958	<0,001*
EAV-T	29,269	10,6255	43,261	26,7051	0,076
	FORÇA				p-valor
	FORTE (1)		FRACO (2)		
	Média	DP	Média	DP	
EAV - GG	32,269	12,1099	68,519	21,6758	<0,001*
EAV-R	26,923	13,8937	58,258	22,8024	<0,001*
EAV-S	11,885	14,5733	47,016	30,9958	<0,001*
EAV-T	29,269	10,6255	43,261	26,7051	0,017*
	AUTORIDADE				p-valor
	AUTORITÁRIO (1)		SUBMISSO (2)		
	Média	DP	Média	DP	
EAV - GG	32,875	12,4410	67,159	22,6686	<0,001*
EAV-R	27,667	14,2388	57,000	23,5334	<0,001*
EAV-S	11,792	15,2173	45,953	31,0791	<0,001*
EAV-T	29,458	11,0751	42,753	26,4277	0,024*
	COMPETÊNCIA				p-valor
	COMPETENTE (1)		INCOMPETENTE (2)		
	Média	DP	Média	DP	
EAV - GG	38,583	18,6005	71,119	20,7360	<0,001*
EAV-R	32,028	19,8787	60,750	21,3688	<0,001*
EAV-S	14,333	15,7377	52,077	30,5368	<0,001*
EAV-T	32,222	20,6242	43,908	25,1868	0,112

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; GR:** grau de rugosidade**; GS:** grau de soprosidade**; GT:** grau de tensão**.**

Nós observamos diferenças em todas as medidas do JPA entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão atividade: resistência, segurança e independência (Tabela 08). Indivíduos com vozes avaliadas como mais resistentes, seguros e independentes apresentaram menores valores na EAV-GG, EAV-R, EAV-S e EAV-T em relação às vozes avaliadas com valência negativa nessas atitudes. Por sua vez, não houve diferença em nenhuma das medidas do JPA para a atitude calma.

TABELA 08: Comparação das médias referente ao julgamento perceptivo-auditivo das vozes em relação aos atributos da dimensão atividade: resistência, segurança, calma e independência.

Variáveis	RESISTÊNCIA				p-valor
	RESISTENTE (1)		FRÁGIL (2)		
	Média	DP	Média	DP	
EAV - GG	35,033	13,3810	69,590	22,0213	<0,001*
EAV-R	29,900	15,2024	58,879	23,4291	<0,001*
EAV-S	11,733	13,7372	49,517	30,4251	<0,001*
EAV-T	29,467	13,7158	44,124	26,5822	0,020*
	SEGURANÇA				p-valor
	SEGURO (1)		INSEGURO (2)		
	Média	DP	Média	DP	
EAV - GG	35,033	13,3810	69,590	22,0213	<0,001*
EAV-R	29,900	15,2024	58,879	23,4291	<0,001*
EAV-S	11,733	13,7372	49,517	30,4251	<0,001*
EAV-T	29,467	13,7158	44,124	26,5822	0,020*
	CALMA				
	CALMO (1)		AGITADO (2)		
	Média	DP	Média	DP	p-valor
EAV - GG	59,004	23,3724	55,719	29,4748	0,686
EAV-R	49,839	23,5844	47,531	28,0900	0,772
EAV-S	44,196	29,4808	23,406	31,4199	0,034*
EAV-T	35,950	21,8866	44,688	26,8725	0,248
	INDEPENDÊNCIA				p-valor
	INDEPENDENTE (1)		DEPENDENTE (2)		
	Média	DP	Média	DP	
EAV - GG	33,571	13,1980	69,120	21,6266	<0,001*
EAV-R	25,714	14,7384	59,867	21,1322	<0,001*
EAV-S	18,750	18,2090	44,983	33,1120	0,008*
EAV-T	28,786	13,2849	43,953	26,2957	0,049*

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; GR:** grau de rugosidade**; GS:** grau de soprosidade**; GT:** grau de tensão**.**

Nós observamos diferenças em medidas acústicas entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão **avaliação**: agradabilidade, simpatia, saúde e extroversão (Tabela 9). Indivíduos com vozes avaliadas como mais agradáveis, simpáticos, saudáveis e extrovertidos apresentaram valores maiores, em relação às vozes com valência negativa, nas medidas: CPP e CPPS, e menores em: PSD; LNPSD; jitterLoc; jitterABS; jitterRAP; jitterPPQ5; jiterddp; shimmerLoc; shimmerdB; shimmerAPQ3; shimmerAPQ5; shimmerAPQ11; shimmerdda; AVI; Correlacao; GNE1; GNE2; HNRmedia; f0dp; f0max; SNL1_3Hz; SNL2_3Hz; SNL3_3Hz; SNL4_3Hz; SNL5_3Hz; SNL6_3Hz.

As medidas acústicas f0CPP, H1A1, PA e SFR apresentaram diferença apenas para o atributo saúde.

TABELA 09: Comparação das médias dos parâmetros acústicos das vozes em relação às atitudes da dimensão avaliação: agradabilidade, simpatia, saúde e extroversão.

Variável	VALÊNCIA (1)Positiva (2)Negativa	AGRADABILIDADE			SIMPATIA			SAÚDE			EXTOVERSÃO		
		Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor
f_0CPP	1,0	160,9992	42,92394	0,211	168,9110	46,19051	0,616	158,1032	42,82237	0,030*	164,1855	48,52387	0,359
	2,0	181,0147	47,76923		177,5103	47,58687		188,8200	46,23713		179,3461	46,45577	
CPP	1,0	26,7600	6,35634	0,036*	28,5650	6,48282	0,011*	26,8889	5,58320	0,007*	29,0127	4,13050	0,003*
	2,0	22,5641	6,97334		22,2800	6,56450		21,2912	7,08922		21,9403	6,89717	
CPPS	1,0	17,7558	2,67035	0,005*	18,7970	1,87469	0,001*	17,2842	2,87006	0,001*	18,4200	2,41901	0,001*
	2,0	13,3325	4,94253		13,2865	4,73609		12,4524	5,02568		13,2452	4,76655	
declinio	1,0	-22,3717	4,89538	0,735	-20,5640	3,80050	0,180	-24,0900	6,30385	0,274	-21,3055	5,00666	0,332
	2,0	-23,0981	6,71080		-23,5871	6,65989		-21,9956	6,12934		-23,4315	6,55825	
Tilt	1,0	-7,7792	1,44015	0,151	-7,5050	1,82647	0,368	-7,5437	2,06264	0,123	-7,6045	1,93575	0,271
	2,0	-6,4663	2,96568	,	-6,6241	2,88408		-6,2776	3,00434		-6,5642	2,87289	
H1A1	1,0	-1,3667	6,29462	0,298	-1,9590	6,15578	0,611	-0,8632	5,05130	0,037*	-2,3264	5,77698	0,785
	2,0	-3,1978	4,65164		-2,9159	4,88585		-4,0932	4,84753		-2,8224	5,00035	
H1A3	1,0	13,5367	7,60817	0,582	12,2560	7,84089	0,299	16,3300	8,66016	0,256	12,3136	7,51110	0,280
	2,0	15,1106	8,63276		15,3947	8,41926		13,4284	7,97848		15,4706	8,51876	
H1H2	1,0	-,3800	3,31842	0,283	-,8370	3,38472	0,194	0,7989	3,74896	0,910	-0,8418	3,22028	0,165
	2,0	1,3706	5,16413		1,4021	5,02342		0,9648	5,48222		1,4715	5,08288	
$f_0$media	1,0	167,9333	43,32411	0,486	166,5010	45,98631	0,479	162,3505	43,70227	0,115	163,3100	48,00538	0,332
	2,0	181,3303	60,30113		180,9635	58,84937		189,3244	62,15019		182,4655	58,32661	
$f_0$dp	1,0	1,2758	0,38156	0,022*	1,2760	0,42466	0,043*	1,3205	0,40809	0,001*	1,3600	0,40581	0,032*
	2,0	17,1113	22,82794		16,1797	22,44631		21,5112	24,10110		16,6033	22,65641	
$f_0$mediana	1,0	167,9750	43,38209	0,487	166,5240	46,03405	0,479	162,3689	43,75150	0,115	163,3318	48,05376	0,332
	2,0	181,3725	60,39043		181,0112	58,93799		189,3844	62,24658		182,5142	58,41616	
f_0q1	1,0	167,0692	43,07682	0,637	165,6220	45,69210	0,610	161,4753	43,55993	0,208	162,3900	47,76167	0,444
	2,0	176,0075	59,41535		175,9074	57,98049		182,7616	61,74503		177,2964	57,52711	
f_0q3	1,0	168,7583	43,48232	0,335	167,3530	46,20842	0,350	163,2242	43,79730	0,056	164,2009	48,16953	0,234
	2,0	189,2009	67,24115		188,4118	65,58911		199,1308	69,57055		190,1006	65,16096	
f_0CV	1,0	,7750	0,21944	0,021*	,7810	0,23435	0,042*	0,8568	0,34562	0,001*	0,8755	0,32104	0,033*
	2,0	8,6434	11,24957		8,1788	11,06537		10,7844	11,89867		8,3715	11,17948	
$f_0$min	1,0	164,5325	42,59108	0,275	163,0640	45,09771	0,382	158,5811	43,13660	0,334	159,7209	47,34314	0,477
	2,0	142,8516	62,50917		144,5588	61,35720		141,3040	67,22768		145,1124	61,49475	
$f_0$max	1,0	171,8292	43,35484	0,036*	170,3110	45,95038	0,044*	166,1526	43,60742	0,011*	167,3155	47,99169	0,027
	2,0	220,5700	105,99519		218,1494	103,42714		238,5316	111,60017		220,5976	103,56760	
PERIODO	1,0	,006380	0,001875	0,677	0,0064928	0,002026	0,564	0,0066225	0,0018815	0,186	0,0067067	0,002244	0,299
	2,0	,006105	0,001956		0,0060888	0,001905		0,0058447	0,0019127		0,0060052	0,001799	
PSD	1,0	,000053	0,000030	0,021*	0,0000544	0,000033	0,044*	0,0000627	0,0000410	0,001*	0,0000648	0,000048	0,037*
	2,0	,000539	0,000699		0,0005107	0,000687		0,0006686	0,0007428		0,0005211	0,000691	
LNPSD	1,0	-9,93564	0,4171428	0,001*	-9,928106	0,4455156	0,003*	-9,827425	0,5274385	0,0001*	-9,80901	0,556232	0,005*
	2,0	-8,41186	1,403370		-8,503717	1,4117124		-8,067454	1,366023		-8,50025	1,436203	

jitterLoc	1,0	0,3375	0,11112	0,002*	0,3100	0,11944	0,001*	0,3868	0,15265	0,007*	0,3618	0,17429	0,065
	2,0	1,0663	1,17691		1,0315	1,14889		1,2328	1,28394		1,0361	1,16799	
jitterABS	1,0	0,000022	0,0000131	0,001*	0,0000216	0,0000149	0,001*	0,0000275	0,000018	0,011*	0,0000270	0,000021	0,093
	2,0	0,0000615	0,0000600		0,0000596	0,0000587		0,0000687	0,000065		0,0000589	0,000059	
jitterRAP	1,0	0,001742	0,0007633	0,003*	0,001490	0,0007156	0,067	0,002153	0,0011247	0,018*	0,001900	0,0012288	0,090
	2,0	0,005750	0,0070023		0,005588	0,0068141		0,006560	0,0077302		0,005576	0,0069374	
jitterPPQ5	1,0	0,2008	0,06947	0,006*	0,1850	0,07634	0,105	0,2253	0,08269	0,020*	0,2127	0,09850	0,114
	2,0	0,6300	0,82362		0,6094	0,80236		0,7316	0,90826		0,6130	0,81529	
jiterddp	1,0	0,5817	0,19940	0,002*	0,5240	0,20414	0,001*	0,6742	0,28338	0,008*	0,6236	0,31951	0,068
	2,0	1,7978	1,98226		1,7432	1,93292		2,0680	2,16735		1,7470	1,96631	
shimmerLoc	1,0	1,8883	1,01168	0,029*	1,8120	1,07907	0,046*	1,8663	,90916	0,001*	1,8773	1,04595	0,039*
	2,0	4,7000	4,23423		4,5571	4,14632		5,5040	4,46425		4,6185	4,19490	
shimmerdB	1,0	0,1633	0,08927	0,019*	0,1570	0,09592	0,033*	0,1616	0,08126	0,0001*	0,1627	0,09296	0,027*
	2,0	0,4525	0,40420		0,4374	0,39665		0,5348	0,42128		0,4439	0,40090	
shimmerAPQ3	1,0	1,0475	0,54443	0,034*	0,9990	0,56785	0,002*	1,0289	0,49722	0,001*	1,0336	0,55050	0,043*
	2,0	2,4294	2,14134		2,3624	2,09477		2,8304	2,26005		2,3921	2,11996	
shimmerAPQ5	1,0	1,1508	0,63079	0,002*	1,1040	0,67722	0,002*	1,1232	0,56269	0,002*	1,1491	0,65954	0,052
	2,0	2,8241	2,67219		2,7394	2,61334		3,3136	2,83338		2,7739	2,64598	
shimmerAPQ11	1,0	1,4175	,86179	0,041*	1,3620	0,94494	0,062	1,4042	0,76275	0,003*	1,4209	0,91885	0,045*
	2,0	3,5694	3,47421		3,4591	3,39668		4,1820	3,70239		3,5030	3,43942	
shimmerdda	1,0	3,1450	1,63355	0,034*	2,9970	1,70372	0,049*	3,0905	1,49163	0,001*	3,1018	1,65227	0,043*
	2,0	7,2916	6,42147		7,0912	6,28141		8,4940	6,77758		7,1803	6,35702	
AVI	1,0	1,0042	0,44086	0,007*	,9460	0,46388	0,006*	1,0200	0,41753	0,0001*	0,9782	0,45349	0,007*
	2,0	1,5334	0,58904		1,5194	0,57370		1,6696	0,56441		1,5261	0,58110	
Correlacao	1,0	0,9917	0,00577	0,047*	0,9930	0,00675	0,069	0,9932	0,00671	0,002*	0,9918	0,00751	0,061
	2,0	0,9456	0,07746		0,9479	0,07563		0,9316	0,08245		0,9470	0,07659	
GNE1	1,0	0,9425	0,08192	0,031*	0,9760	0,01506	0,008*	0,9358	0,08181	0,003*	0,9573	0,07030	0,018*
	2,0	0,8134	0,19233		0,8112	0,18700		0,7824	0,20325		0,8124	0,18903	
GNE2	1,0	0,9358	0,07573	0,008*	0,9660	0,01897	0,003*	0,9126	0,10614	0,002*	0,9473	0,06798	0,007*
	2,0	0,7622	0,21074		0,7635	0,20472		0,7312	0,21841		0,7636	0,20756	
GNE3	1,0	0,8783	0,13842	0,008*	0,9250	0,09846	0,001*	0,8379	0,17466	0,008*	0,8991	0,13932	0,004*
	2,0	0,6888	0,21960		0,6862	0,21259		0,6664	0,21916		0,6876	0,21332	
HNRmedia	1,0	24,1175	3,43656	0,012*	24,4190	3,75910	0,018*	24,1726	3,32902	0,0001*	23,9955	3,94925	0,022*
	2,0	18,1994	7,48096		18,4588	7,32121		16,5004	7,45974		18,4194	7,41199	
HNRdp	1,0	2,8550	0,82607	0,249	2,7880	,81094	0,258	2,8421	1,09911	0,091	2,7500	,88710	0,203
	2,0	3,7181	2,48824		3,6871	2,42322		3,9696	2,66514		3,7270	2,43796	
HNRD	1,0	25,9808	7,23423	0,207	26,6930	7,77963	0,142	24,9026	7,30742	0,342	25,2882	8,52187	0,408
	2,0	22,8166	7,31599		22,7932	7,09262		22,7500	7,39115		23,1433	6,97791	
PA	1,0	0,8983	0,03664	0,105	0,8970	0,04029	0,165	0,8963	0,03467	0,025*	0,8936	0,04225	0,173
	2,0	0,8369	0,12548		0,8409	0,12271		0,8212	0,13743	,	0,8403	0,12431	
SFR	1,0	-18,3900	1,17284	0,083	-18,3640	1,22664	0,145	-18,3337	1,00688	0,024*	-18,4145	1,18627	0,090
	2,0	-17,5675	1,43413		-17,6235	1,42504		-17,3800	1,53740		-17,5842	1,42559	
SNL1_3Hz	1,0	13,1233	3,94289	0,004*	13,6790	3,80570	0,034*	13,8258	4,13166	0,001*	14,2300	4,30710	0,033*
	2,0	18,2828	5,33822		17,8159	5,57085		19,1936	5,27733		17,7576	5,58613	
SNL2_3Hz	1,0	9,1792	6,09463	0,004*	9,6180	6,12689	0,022*	9,6368	5,35600	0,0001*	10,0018	5,94294	0,028*
	2,0	15,2916	5,75403		14,8029	6,06890		16,6552	5,45020		14,8321	6,16278	

SNL3_3Hz	1,0	14,8750	4,79876	0,002*	14,6130	3,99117	0,005*	15,9100	4,98044	0,001*	15,5727	4,68473	0,018*
	2,0	20,9409	5,75443		20,6612	5,96687		21,8528	5,67599		20,5245	6,07039	
SNL4_3Hz	1,0	15,1575	4,60293	0,003*	14,9650	3,93634	0,006*	16,1016	4,80246	0,001	15,8873	4,53065	0,023*
	2,0	20,8978	5,59951		20,6168	5,79012		21,7876	5,50927		20,4806	5,89833	
SNL5_3Hz	1,0	11,3708	4,37268	0,032*	12,7450	4,05777	0,302	11,7411	4,36090	0,006*	12,8882	4,58736	0,314
	2,0	15,6250	6,06531		14,9706	6,32990		16,5348	6,18569		14,9903	6,27944	
SNL6_3Hz	1,0	2,6017	10,89974	0,014*	2,8300	11,06973	0,039*	2,6163	9,16861	0,0001*	2,9345	10,75260	0,031*
	2,0	10,3009	8,02902		9,7809	8,43229		12,4456	7,26426		9,9567	8,40315	

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma; **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico; **H1H2:** Razão da amplitude do primeiro e segundo harmônico; **fo média**: Média da Frequência Fundamental; **f_o DP**: Desvio Padrão da Frequência Fundamental; **fo mediana**: Mediana da Frequência Fundamental;**fo q1:** 1º Quartil da Frequência Fundamental. **fo q3:** 3º Quartil da Frequência Fundamental; **fo CV:** Coeficiente de variação da Frequência Fundamental; **fo min:** Mínima da Frequência Fundamental; **F0 max**: Máxima da Frequência Fundamental; **PSD:** Densidade Espectral de potência de ruído; **LNPSD:** Logaritmo do desvio padrão do período; **Jitter LOC:** Média de perturbação relativa; **Jitter ABS:** perturbação média relativa; **Jitter RAP:** Perturbação média relativa em porcentagem; **Jitter; PPQ5:** Coeficiente médio de perturbação;**Jitter DDP**: Diferença média absoluta entre períodos consecutivos; **Shimmer Loc:** Média de perturbação relativa;. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3; **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz; **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz; **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz;. **Hfno:** Nível de energia relativa do ruído de alta frequência; **HNR média:** Média da relação harmônico-ruído.; **HNRD:** Desvio Padrão da relação harmônico ruído; **PA:** Amplitude da Frequência; **SFR:** Planicidade espectral do sinal residual; **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**); SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

Nós observamos diferenças em todas as medidas acústicas entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão **potência**: poder, força, autoridade e competência (Tabela 10). Indivíduos com vozes avaliadas como mais poderosos, fortes, autoritários e competentes apresentaram valores maiores, em relação às vozes com valência negativa, nas medidas: CPP e CPPS, e menores em: shimmerLoc; shimmerdB; shimmerAPQ3; shimmerAPQ5; shimmerAPQ11; shimmerdda; AVI; Correlacao; GNE1; GNE2; GNE3; HNRmedia; f0dp; f0CV; f0max; PSD; LNPSD; jitterLoc; jitterABS; jitterRAP; SNL1_3Hz; SNL2_3Hz; SNL4_3Hz; SNL5_3Hz; SNL6_3Hz.

As medidas acústicas tilt e jitterPPQ5 apresentaram diferença apenas para a atitude competência.

TABELA 10: Comparação das médias dos parâmetros acústicos das vozes em relação às atitudes da dimensão potência: resistência, segurança, calma, independência.

Variável	VALÊNCIA (1)Positiva (2)Negativa	RESISTÊNCIA			SEGURANÇA			CALMA			INDEPENDÊNCIA		
		Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor
f0CPP	1,0	158,941333	42,9929474	0,091	158,941333	42,9929474	0,091	174,2936	46,80650	0,816	171,170714	44,2860679	0,677
	2,0	184,149655	47,1859601		184,149655	47,1859601		177,7650	48,46078		177,602333	48,6405281	
CPP	1,0	27,503333	6,1274647	0,008*	27,503333	6,1274647	0,008*	22,7311	6,70586	0,224	27,413571	5,5696618	0,015*
	2,0	21,745517	6,6872472		21,745517	6,6872472	0	25,4188	7,38394		21,979333	6,9984663	
CPPS	1,0	18,266667	2,4552558	0,000*	18,266667	2,4552558	0,000*	13,5257	5,07818	0,049*	17,484286	2,7596396	0,004*
	2,0	12,610690	4,6515051		12,610690	4,6515051		16,3119	3,92574		13,164333	5,0149547	
Declinio	1,0	-21,535333	4,5157421	0,301	-21,535333	4,5157421	0,301	-24,3936	6,31721	0,033	-23,319286	5,7747773	0,764
	2,0	-23,605862	6,9086589		-23,605862	6,9086589		-20,2863	5,26016		-22,704333	6,5050753	
Tilt	1,0	-7,954667	1,8009397	0,043*	-7,954667	1,8009397	0,043*	-6,8082	2,56981	0,959	-7,078571	1,6727447	0,673
	2,0	-6,239655	2,9033558		-6,239655	2,9033558		-6,8525	2,96855		-6,705667	3,0671156	
H1A1	1,0	-1,972667	5,1325327	0,507	-1,972667	5,1325327	0,507	-1,9761	4,69167	0,222	-1,997143	6,2161532	0,543
	2,0	-3,073793	5,1930272		-3,073793	5,1930272		-3,9625	5,78148		-3,025667	4,6366357	
H1A3	1,0	12,668667	6,6187244	0,253	12,668667	6,6187244	0,253	17,0554	7,97248	0,010*	16,149286	9,5670633	0,430
	2,0	15,722414	8,9902926		15,722414	8,9902926		10,5269	7,38701		13,996333	7,7336950	
H1H2	1,0	-,556000	2,9985897	0,149	-,556000	2,9985897	0,149	1,8136	4,67835	0,090	-0,387143	3,8844313	0,227
	2,0	1,642759	5,3482860		1,642759	5,3482860		-,07175	4,60738		1,490667	5,0677506	
$f_0$media	1,0	156,800667	42,7023802	0,075	156,800667	42,7023802	0,075	180,9604	60,28356	0,613	176,366429	42,7431687	0,917
	2,0	188,474483	59,6085293		188,474483	59,6085293		171,9300	48,97955		178,288000	61,9093308	
$f_0$dp	1,0	1,357333	0,4370758	0,007*	1,357333	0,4370758	0,007*	13,6586	21,21318	0,718	3,267857	7,7324821	0,035*
	2,0	18,707241	23,4261461		18,707241	23,4261461		11,2769	20,22686		17,237333	23,2709493	
$f_0$mediana	1,0	156,812000	42,7508311	0,075	156,812000	42,7508311	0,075	182,7796	59,26688	0,435	176,963571	43,5286379	0,952
	2,0	188,532414	59,6990338		188,532414	59,6990338		168,8619	50,67162		178,071000	61,7788912	
f_0q1	1,0	155,895333	42,5182458	0,127	155,895333	42,5182458	0,127	177,2164	58,39178	0,568	176,060714	43,2014915	0,840
	2,0	182,711724	59,1988230		182,711724	59,1988230		167,1881	49,97437		172,407333	60,5020675	
f_0q3	1,0	157,680667	42,8328061	0,044	157,680667	42,8328061	0,044	187,6154	68,81067	0,578	177,790714	43,7416177	0,674
	2,0	197,045517	66,4407913		197,045517	66,4407913		176,6438	48,67535		186,348667	69,2409949	
f_0CV	1,0	0,908000	,03664930	0,007*	0,908000	0,3664930	0,007*	6,6054	10,04664	0,927	1,608571	3,4166490	0,028*
	2,0	9,388621	11,5744129		9,388621	11,5744129		6,3088	10,76369		8,779000	11,4793044	
$f_0$min	1,0	153,118000	42,2646499	0,725	153,118000	42,2646499	0,725	147,4425	61,27441	0,844	164,395000	43,5793935	0,227
	2,0	146,512759	65,3960803		146,512759	65,3960803		151,0781	53,96181		141,470333	63,1036424	
$f_0$max	1,0	160,840000	42,6005159	0,018*	160,840000	42,6005159	0,018*	204,0907	83,34854	0,773	180,750000	45,0405158	0,210

	2,0	231,296207	106,101310		231,296207	106,101310		212,8531	115,85721		219,656333	109,681700	
PERIODO	1,0	0,006856	0,0019476	0,093	0,006856	0,0019476	0,093	0,0060825	,00185989	0,659	0,006004	0,0014971	0,682
	2,0	0,005831	0,0018379		0,005831	0,0018379		0,0063523	0,00206423		0,006263	0,0021032	
PSD	1,0	,000068	0,0000441	0,009*	0,000068	00,0000441	0,009*	0,0004181	0,00062184	0,880	0,000101	0,0002102	0,027*
	2,0	,000582	0,0007219		0,000582	,0007219		0,0003876	0,00067259		0,000550	0,0007128	
LNPSD	1,0	-9,748941	0,5362573	0,001*	-9,748941	0,5362573	0,001*	-8,7788252	1,41989158	0,763	-9,845034	0,8389014	0,000*
	2,0	-8,350805	1,4623746		-8,350805	1,4623746		-8,9125221	1,38116748		-8,352566	1,3495834	
JitterLoc	1,0	,370000	0,1693264	0,022*	0,370000	0,1693264	0,022*	1,0018	1,26020	0,268	0,321429	0,0999890	0,017*
	2,0	1,124828	1,2199930		1,124828	1,2199930		,6325	,47899		1,122333	1,1956262	
JitterABS	1,0	0,000028	0,0000199	0,039*	0,000028	0,0000199	0,039*	0,0000560	,00006264	0,415	0,000020	0,0000087	0,007*
	2,0	0,000063	0,0000624		0,000063	0,0000624		0,0000420	,00003559		0,000066	0,0000604	
JitterRAP	1,0	0,001927	0,0012285	0,035*	0,001927	0,0012285	0,035*	0,005425	,0074827	0,284	0,001707	0,0008325	0,030*
	2,0	0,006069	0,0072601		0,006069	0,0072601	*	0,003313	,0027011		0,006033	0,0071413	
jitterPPQ5	1,0	0,216000	0,0910102	0,050	0,216000	0,0910102	0,050*	0,6093	,88016	0,249	0,187857	0,0589887	0,041*
	2,0	0,666552	0,8570392		0,666552	0,8570392		0,3444	,26566		0,664667	0,8402370	
Jiterddp	1,0	0,634667	0,3142080	0,023*	0,634667	0,3142080	0,023*	1,7136	2,11843	0,225	0,552857	0,1931691	0,018*
	2,0	1,896207	2,0532142		1,896207	2,0532142		1,0331	,78620		1,892333	2,0132658	
ShimmerLoc	1,0	1,846667	0,9385069	0,008*	1,846667	0,9385069	0,008*	4,3657	4,33030	0,330	1,705000	0,6763562	0,007*
	2,0	5,012414	4,3285990		5,012414	4,3285990		3,1763	2,77304		4,973000	4,2675626	
ShimmerdB	1,0	0,160667	0,0836205	0,005*	0,160667	0,0836205	0,005*	0,4207	,41784	0,269	0,148571	0,0609990	0,004*
	2,0	0,483793	0,4119589		0,483793	0,4119589		0,2913	,25786		0,478667	0,4065645	
shimmerAPQ3	1,0	1,004000	0,4926720	0,008*	1,004000	0,4926720	0,008*	2,2371	2,14277	0,410	0,940714	0,3874863	0,008*
	2,0	2,594828	2,1847991		2,594828	2,1847991		1,7294	1,53789		2,571333	2,1554994	
shimmerAPQ5	1,0	1,132667	0,5940956	0,013*	1,132667	0,5940956	0,013*	2,6071	2,71511	0,390	1,042143	0,3908999	0,011*
	2,0	3,006552	2,7419509		3,006552	2,7419509		1,9488	1,76478		2,986333	2,7044146	
shimmerAPQ11	1,0	1,429333	0,8361602	0,016*	1,429333	,8361602	0,016*	3,3450	3,60550	0,316	1,292857	0,5296941	0,013*
	2,0	3,785862	3,5760408		3,785862	3,5760408		2,3481	2,02726		3,771000	3,5257059	
Shimmerdda	1,0	3,015333	1,4782174	0,008*	3,015333	1,4782174	0,008*	6,7150	6,42497	0,410	2,827143	1,1630691	0,008*
	2,0	7,787586	6,5518692		7,787586	6,5518692		5,1906	4,61503		7,716333	6,4646031	
AVI	1,0	0,982000	0,4350238	0,001*	0,982000	0,4350238	0,001*	1,4282	0,63723	0,572	0,885000	0,3363092	0,000*
	2,0	1,599655	0,5643484		1,599655	0,5643484		1,3206	0,53275		1,624333	0,5461696	
Correlacao	1,0	0,9913	0,00640	0,020*	0,9913	0,00640	0,020*	0,9518	0,07940	0,423	0,9936	0,00633	0,018*
	2,0	0,9410	0,08002		0,9410	0,08002		0,9694	0,04582		0,9417	0,07848	
GNE1	1,0	0,950000	0,0746420	0,005*	0,950000	0,0746420	0,005*	0,8279	0,18974	0,312	0,956429	,0608412	0,005*
	2,0	0,796207	0,1940253		0,796207	0,1940253		0,8850	0,15505		0,798333	,1929103	
GNE2	1,0	0,9433	0,06935	0,001*	0,9433	0,06935	0,001*	0,7846	0,20823	0,277	0,9286	0,09960	0,005*
	2,0	0,7403	0,20945		0,7403	0,20945		0,8531	0,17973		0,7540	0,21035	
GNE3	1,0	0,895333	0,1274960	0,000*	0,895333	0,1274960	0,0001*	0,7057	0,20707	0,162	0,840714	0,1771772	0,034*
	2,0	0,660345	0,2109415		0,660345	0,2109415		0,8013	0,22639		0,693667	0,2201016	
HNRmedia	1,0	23,5180	3,44199	0,011*	23,5180	3,44199	0,011*	19,5457	7,83809	0,745	024,8614	3,05751	0,001*
	2,0	17,8972	7,77423		17,8972	7,77423		20,2819	5,80920		17,4577	7,25333	
HNRdp	1,0	2,918000	,9830724	0,222	2,918000	,9830724	0,222	3,7807	2,54060	0,236	2,846429	1,0340853	0,191
	2,0	3,774828	2,5714915		3,774828	2,5714915		2,9613	1,28652		3,779667	2,5170131	
HNRD	1,0	24,295333	7,4734710	0,694	24,295333	7,4734710	0,694	23,3718	7,05709	0,718	28,078571	6,5774228	0,005*
	2,0	23,361034	7,3962483		23,361034	7,3962483		24,2181	8,04294		21,626667	6,8544792	
PA	1,0	0,892667	0,0359497	0,096	0,892667	0,0359497	0,096	0,8404	0,13443	0,302	0,903571	0,0297702	0,041*

	2,0	0,833448	0,1313468		0,833448	0,1313468		0,8769	0,04785		0,830333	0,1277250	
SFR	1,0	-18,264667	1,1139432	0,109	-18,264667	1,1139432	0,109	-17,7614	1,31611	0,852	-17,882857	0,9869879	0,773
	2,0	-17,547241	1,4920358		-17,547241	1,4920358		-17,8450	1,59069		-17,749333	1,5752327	
SNL1_3Hz	1,0	14,768000	4,0491184	0,065	14,768000	4,0491184	0,065	16,5443	5,71715	0,601	13,500714	3,8393599	0,004*
	2,0	17,965862	5,8434704		17,965862	5,8434704		17,4556	5,13651		18,450667	5,4495061	
SNL2_3Hz	1,0	10,782000	5,3439595	0,032*	10,782000	5,3439595	0,032*	13,3525	6,31675	0,714	9,448571	5,8768462	0,002*
	2,0	15,094828	6,4843937		15,094828	6,4843937		14,1006	6,72307		15,573333	5,7354520	
SNL3_3Hz	1,0	16,808667	4,9422174	0,052*	16,808667	4,9422174	0,052*	19,1850	6,12971	0,886	14,497857	3,9019545	0,0001*
	2,0	20,568276	6,3224459		20,568276	6,3224459		19,4644	6,24861		21,521333	5,6694966	
SNL4_3Hz	1,0	17,040667	4,7669449	0,063*	17,040667	4,7669449	0,063*	19,1714	5,97154	0,814	14,813571	3,8540141	0,0001*
	2,0	20,517586	6,1382080		20,517586	6,1382080		19,6138	5,93200		21,441000	5,5094391	
SNL5_3Hz	1,0	12,728000	4,2900952	0,164*	12,728000	4,2900952	0,164*	13,9036	6,30070	0,412	12,500714	4,2334846	0,135*
	2,0	15,363103	6,4982993		15,363103	6,4982993		15,4469	5,24455		15,381333	6,4232862	
SNL6_3Hz	1,0	4,130667	9,3966829	0,038*	4,130667	9,3966829	0,038*	7,9764	8,52543	0,837	2,595000	10,2969546	0,006*
	2,0	10,306552	8,8828902		10,306552	8,8828902		8,5944	11,12439		10,817333	7,8760159	

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma. **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico. **H1H2:** Razão da amplitude do primeiro e segundo harmônico. **fo média**: Média da Frequência Fundamental. **fo DP**: Desvio Padrão da Frequência Fundamental. **fo mediana**: Mediana da Frequência Fundamental. **fo q1:** 1º Quartil da Frequência Fundamental. **fo q3:** 3º Quartil da Frequência Fundamental. **fo CV:** Coeficiente de variação da Frequência Fundamental. **fo min:** Mínima da Frequência Fundamental. **fo max**: Máxima da Frequência Fundamental. **PSD:** Densidade Espectral de potência de ruído. **LNPSD:** Logaritmo do desvio padrão do período. **Jitter LOC:** Média de perturbação relativa. **Jitter ABS:** perturbação média relativa. **Jitter RAP:** Perturbação média relativa em porcentagem. **Jitter PPQ5:** Coeficiente médio de perturbação.**Jitter DDP**: Diferença média absoluta entre períodos consecutivos. **Shimmer Loc:** Média de perturbação relativa. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3. **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz. **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz. **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz. **Hfno:** Nível de energia relativa do ruído de alta frequência. **HNR media:** Média da relação harmônico-ruído. **HNRD:** Desvio Padrão da relação harmônico ruído. **PA:** Amplitude da Frequência. **SFR:** Planicidade espectral do sinal residual. **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**). SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

Nós observamos diferenças em todas as medidas acústicas em vozes avaliadas com valência positiva e negativa pelos ouvintes nas seguintes atitudes da dimensão **atividade:** resistência, segurança e independência (Tabela 11). Indivíduos com vozes avaliadas como mais resistentes, seguros e independentes apresentaram maiores valores para as medidas: CPP e CPPS, e menores

em relação as avaliadas negativamente para as medidas: shimmerLoc; shimmerdB; shimmerAPQ3; shimmerAPQ5; shimmerAPQ11; shimmerdda; AVI; Correlacao; GNE1; GNE2; GNE3; HNRmedia; f0dp; f0CV; f0max; PSD; LNPSD; jitterLoc; jitterABS; jitterRAP; SNL2_3Hz; SNL4_3Hz; SNL3_3Hz; SNL5_3Hz; SNL6_3Hz.

Por sua vez, para a atitude calma, só houve diferença para as medidas CPPS e H1A3. As medidas acústicas HNRD; PA e SNL1_3Hz apresentaram diferença apenas para a atitude independência.

TABELA 11: Comparação das médias dos parâmetros acústicos das vozes em relação aos atributos da dimensão atividade: resistência, segurança calma e independência.

Variável	VALÊNCIA **(1)**Positiva **(2)**Negativa	RESISTÊNCIA			SEGURANÇA			CALMA			INDEPENDÊNCIA		
		Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor	Media	Desvio padrão	p-valor
f0CPP	1,0	158,941333	42,9929474	0,091	158,941333	42,9929474	0,091	174,2936	46,80650	0,816	171,170714	44,2860679	0,677
	2,0	184,149655	47,1859601		184,149655	47,1859601		177,7650	48,46078		177,602333	48,6405281	
CPP	1,0	27,503333	6,1274647	0,008*	27,503333	6,1274647	0,008*	22,7311	6,70586	0,224	27,413571	5,5696618	0,015*
	2,0	21,745517	6,6872472		21,745517	6,6872472		25,4188	7,38394		21,979333	6,9984663	
CPPS	1,0	18,266667	2,4552558	0,000*	18,266667	2,4552558	0,000*	13,5257	5,07818	0,049*	17,484286	2,7596396	0,004*
	2,0	12,610690	4,6515051		12,610690	4,6515051		16,3119	3,92574		13,164333	5,0149547	
Declinio	1,0	-21,535333	4,5157421	0,301	-21,535333	4,5157421	0,301	-24,3936	6,31721	0,033	-23,319286	5,7747773	0,764
	2,0	-23,605862	6,9086589		-23,605862	6,9086589		-20,2863	5,26016		-22,704333	6,5050753	
Tilt	1,0	-7,954667	1,8009397	0,043*	-7,954667	1,8009397	0,043*	-6,8082	2,56981	0,959	-7,078571	1,6727447	0,673
	2,0	-6,239655	2,9033558		-6,239655	2,9033558		-6,8525	2,96855		-6,705667	3,0671156	
H1A1	1,0	-1,972667	5,1325327	0,507	-1,972667	5,1325327	0,507	-1,9761	4,69167	0,222	-1,997143	6,2161532	0,543
	2,0	-3,073793	5,1930272		-3,073793	5,1930272		-3,9625	5,78148		-3,025667	4,6366357	
H1A3	1,0	12,668667	6,6187244	0,253	12,668667	6,6187244	0,253	17,0554	7,97248	0,010*	16,149286	9,5670633	0,430
	2,0	15,722414	8,9902926		15,722414	8,9902926		10,5269	7,38701		13,996333	7,7336950	
H1H2	1,0	-,556000	2,9985897	0,149	-,556000	2,9985897	0,149	1,8136	4,67835	0,090	-0,387143	3,8844313	0,227
	2,0	1,642759	5,3482860		1,642759	5,3482860		-0,07175	4,60738		1,490667	5,0677506	
$f_0$media	1,0	156,800667	42,7023802	0,075	156,800667	42,7023802	0,075	180,9604	60,28356	0,613	176,366429	42,7431687	0,917
	2,0	188,474483	59,6085293		188,474483	59,6085293		171,9300	48,97955		178,288000	61,9093308	
$f_0$dp	1,0	1,357333	0,4370758	0,007*	1,357333	0,4370758	0,007*	13,6586	21,21318	0,718	3,267857	7,7324821	0,035*
	2,0	18,707241	23,4261461		18,707241	23,4261461		11,2769	20,22686		17,237333	23,2709493	
$f_0$mediana	1,0	156,812000	42,7508311	0,075	156,812000	42,7508311	0,075	182,7796	59,26688	0,435	176,963571	43,5286379	0,952
	2,0	188,532414	59,6990338		188,532414	59,6990338		168,8619	50,67162		178,071000	61,7788912	
f_0q1	1,0	155,895333	42,5182458	0,127	155,895333	42,5182458	0,127	177,2164	58,39178	0,568	176,060714	43,2014915	0,840
	2,0	182,711724	59,1988230		182,711724	59,1988230		167,1881	49,97437		172,407333	60,5020675	
f_0q3	1,0	157,680667	42,8328061	0,044	157,680667	42,8328061	0,044	187,6154	68,81067	0,578	177,790714	43,7416177	0,674
	2,0	197,045517	66,4407913		197,045517	66,4407913		176,6438	48,67535		186,348667	69,2409949	
f_0CV	1,0	0,908000	,03664930	0,007*	0,908000	0,3664930	0,007*	6,6054	10,04664	0,927	1,608571	3,4166490	0,028*
	2,0	9,388621	11,5744129		9,388621	11,5744129		6,3088	10,76369		8,779000	11,4793044	
$f_0$min	1,0	153,118000	42,2646499	0,725	153,118000	42,2646499	0,725	147,4425	61,27441	0,844	164,395000	43,5793935	0,227
	2,0	146,512759	65,3960803		146,512759	65,3960803		151,0781	53,96181		141,470333	63,1036424	
$f_0$max	1,0	160,840000	42,6005159	0,018*	160,840000	42,6005159	0,018*	204,0907	83,34854	0,773	180,750000	45,0405158	0,210
	2,0	231,296207	106,101310		231,296207	106,101310		212,8531	115,85721		219,656333	109,681700	
PERIODO	1,0	0,006856	0,0019476	0,093	0,006856	0,0019476	0,093	0,0060825	0,00185989	0,659	0,006004	0,0014971	0,682
	2,0	0,005831	0,0018379		0,005831	0,0018379		0,0063523	0,00206423		0,006263	0,0021032	
PSD	1,0	0,000068	0,0000441	0,009*	0,000068	00,0000441	0,009*	0,0004181	0,00062184	0,880	0,000101	0,0002102	0,027*
	2,0	0,000582	0,0007219		0,000582	,0007219		0,0003876	0,00067259		0,000550	0,0007128	

LNPSD	1,0	-9,748941	0,5362573	0,001*	-9,748941	0,5362573	0,001*	-8,7788252	1,41989158	0,763	-9,845034	0,8389014	0,000*
	2,0	-8,350805	1,4623746		-8,350805	1,4623746		-8,9125221	1,38116748		-8,352566	1,3495834	
JitterLoc	1,0	0,370000	0,1693264	0,022*	0,370000	0,1693264	0,022*	1,0018	1,26020	0,268	0,321429	0,0999890	0,017*
	2,0	1,124828	1,2199930		1,124828	1,2199930		0,6325	0,47899		1,122333	1,1956262	
JitterABS	1,0	0,000028	0,0000199	0,039*	0,000028	0,0000199	0,039*	0,0000560	0,00006264	0,415	0,000020	0,0000087	0,007*
	2,0	0,000063	0,0000624		0,000063	0,0000624		0,0000420	0,00003559		0,000066	0,0000604	
JitterRAP	1,0	0,001927	0,0012285	0,035*	0,001927	0,0012285	0,035*	0,005425	0,0074827	0,284	0,001707	0,0008325	0,030*
	2,0	0,006069	0,0072601		0,006069	0,0072601		0,003313	0,0027011		0,006033	0,0071413	
jitterPPQ5	1,0	0,216000	0,0910102	0,050	0,216000	0,0910102	0,050*	0,6093	0,88016	0,249	0,187857	0,0589887	0,041*
	2,0	0,666552	0,8570392		0,666552	0,8570392		0,3444	0,26566		0,664667	0,8402370	
Jiterddp	1,0	0,634667	0,3142080	0,023*	0,634667	0,3142080	0,023*	1,7136	2,11843	0,225	0,552857	0,1931691	0,018*
	2,0	1,896207	2,0532142		1,896207	2,0532142		1,0331	0,78620		1,892333	2,0132658	
ShimmerLoc	1,0	1,846667	0,9385069	0,008*	1,846667	0,9385069	0,008*	4,3657	4,33030	0,330	1,705000	0,6763562	0,007*
	2,0	5,012414	4,3285990		5,012414	4,3285990		3,1763	2,77304		4,973000	4,2675626	
ShimmerdB	1,0	0,160667	0,0836205	0,005*	0,160667	0,0836205	0,005*	0,4207	0,41784	0,269	0,148571	0,0609990	0,004*
	2,0	0,483793	0,4119589		0,483793	0,4119589		0,2913	0,25786		0,478667	0,4065645	
shimmerAPQ3	1,0	1,004000	0,4926720	0,008*	1,004000	0,4926720	0,008*	2,2371	2,14277	0,410	0,940714	0,3874863	0,008*
	2,0	2,594828	2,1847991		2,594828	2,1847991		1,7294	1,53789		2,571333	2,1554994	
shimmerAPQ5	1,0	1,132667	0,5940956	0,013*	1,132667	0,5940956	0,013*	2,6071	2,71511	0,390	1,042143	0,3908999	0,011*
	2,0	3,006552	2,7419509		3,006552	2,7419509		1,9488	1,76478		2,986333	2,7044146	
shimmerAPQ11	1,0	1,429333	0,8361602	0,016*	1,429333	,8361602	0,016*	3,3450	3,60550	0,316	1,292857	0,5296941	0,013*
	2,0	3,785862	3,5760408		3,785862	3,5760408		2,3481	2,02726		3,771000	3,5257059	
Shimmerdda	1,0	3,015333	1,4782174	0,008*	3,015333	1,4782174	0,008*	6,7150	6,42497	0,410	2,827143	1,1630691	0,008*
	2,0	7,787586	6,5518692		7,787586	6,5518692		5,1906	4,61503		7,716333	6,4646031	
AVI	1,0	0,982000	0,4350238	0,001*	0,982000	0,4350238	0,001*	1,4282	0,63723	0,572	0,885000	0,3363092	0,000*
	2,0	1,599655	0,5643484		1,599655	0,5643484		1,3206	0,53275		1,624333	0,5461696	
Correlacao	1,0	0,9913	0,00640	0,020*	0,9913	0,00640	0,020*	0,9518	0,07940	0,423	0,9936	0,00633	0,018*
	2,0	0,9410	0,08002		0,9410	0,08002		0,9694	0,04582		0,9417	0,07848	
GNE1	1,0	0,950000	0,0746420	0,005*	0,950000	0,0746420	0,005*	0,8279	0,18974	0,312	0,956429	0,0608412	0,005*
	2,0	0,796207	0,1940253		0,796207	0,1940253		0,8850	0,15505		0,798333	0,1929103	
GNE2	1,0	0,9433	0,06935	0,001*	0,9433	0,06935	0,001*	0,7846	0,20823	0,277	0,9286	0,09960	0,005*
	2,0	0,7403	0,20945		0,7403	0,20945		0,8531	0,17973		0,7540	0,21035	
GNE3	1,0	0,895333	0,1274960	0,000*	0,895333	0,1274960	0,000*	0,7057	0,20707	0,162	0,840714	0,1771772	0,034*
	2,0	0,660345	0,2109415		0,660345	0,2109415		0,8013	0,22639		0,693667	0,2201016	
HNRmedia	1,0	23,5180	3,44199	0,011*	23,5180	3,44199	0,011*	19,5457	7,83809	0,745	24,8614	3,05751	0,001*
	2,0	17,8972	7,77423		17,8972	7,77423		20,2819	5,80920		17,4577	7,25333	
HNRdp	1,0	2,918000	,9830724	0,222	2,918000	,9830724	0,222	3,7807	2,54060	0,236	2,846429	1,0340853	0,191
	2,0	3,774828	2,5714915		3,774828	2,5714915		2,9613	1,28652		3,779667	2,5170131	
HNRD	1,0	24,295333	7,4734710	0,694	24,295333	7,4734710	0,694	23,3718	7,05709	0,718	28,078571	6,5774228	0,005*
	2,0	23,361034	7,3962483		23,361034	7,3962483		24,2181	8,04294		21,626667	6,8544792	
PA	1,0	0,892667	0,0359497	0,096	0,892667	0,0359497	0,096	0,8404	0,13443	0,302	0,903571	0,0297702	0,041*
	2,0	0,833448	0,1313468		0,833448	0,1313468		0,8769	0,04785		0,830333	0,1277250	
SFR	1,0	-18,264667	1,1139432	0,109	-18,264667	1,1139432	0,109	-17,7614	1,31611	0,852	-17,882857	0,9869879	0,773
	2,0	-17,547241	1,4920358		-17,547241	1,4920358		-17,8450	1,59069		-17,749333	1,5752327	
SNL1_3Hz	1,0	14,768000	4,0491184	0,065	14,768000	4,0491184	0,065	16,5443	5,71715	0,601	13,500714	3,8393599	0,004*
	2,0	17,965862	5,8434704		17,965862	5,8434704		17,4556	5,13651		18,450667	5,4495061	

SNL2_3Hz	1,0	10,782000	5,3439595	0,032*	10,782000	5,3439595	0,032*	13,3525	6,31675	0,714	9,448571	5,8768462	0,002*
	2,0	15,094828	6,4843937		15,094828	6,4843937		14,1006	6,72307		15,573333	5,7354520	
SNL3_3Hz	1,0	16,808667	4,9422174	0,052*	16,808667	4,9422174	0,052*	19,1850	6,12971	0,886	14,497857	3,9019545	0,000*
	2,0	20,568276	6,3224459		20,568276	6,3224459		19,4644	6,24861		21,521333	5,6694966	
SNL4_3Hz	1,0	17,040667	4,7669449	0,063*	17,040667	4,7669449	0,063*	19,1714	5,97154	0,814	14,813571	3,8540141	0,000*
	2,0	20,517586	6,1382080		20,517586	6,1382080		19,6138	5,93200		21,441000	5,5094391	
SNL5_3Hz	1,0	12,728000	4,2900952	0,164*	12,728000	4,2900952	0,164*	13,9036	6,30070	0,412	12,500714	4,2334846	0,135*
	2,0	15,363103	6,4982993		15,363103	6,4982993		15,4469	5,24455		15,381333	6,4232862	
SNL6_3Hz	1,0	4,130667	9,3966829	0,038*	4,130667	9,3966829	0,038*	7,9764	8,52543	0,837	2,595000	10,2969546	0,006*
	2,0	10,306552	8,8828902		10,306552	8,8828902		8,5944	11,12439		10,817333	7,8760159	

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma; **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico; **H1H2:** Razão da amplitude do primeiro e segundo harmônico; **fo média**: Média da Frequência Fundamental; **fo DP**: Desvio Padrão da Frequência Fundamental; **fo mediana**: Mediana da Frequência Fundamental;**fo q1:** 1º Quartil da Frequência Fundamental. **fo q3:** 3º Quartil da Frequência Fundamental; **fo CV:** Coeficiente de variação da Frequência Fundamental; **fo min:** Mínima da Frequência Fundamental; **F0 max**: Máxima da Frequência Fundamental; **PSD:** Densidade Espectral de potência de ruído; **LNPSD:** Logaritmo do desvio padrão do período; **Jitter LOC:** Média de perturbação relativa; **Jitter ABS:** perturbação média relativa;. **Jitter RAP:** Perturbação média relativa em porcentagem.; **Jitter; PPQ5:** Coeficiente médio de perturbação;.**Jitter DDP**: Diferença média absoluta entre períodos consecutivos; **Shimmer Loc:** Média de perturbação relativa;. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3; **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz; **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz; **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz;. **Hfno:** Nível de energia relativa do ruído de alta frequência; **HNR média:** Média da relação harmônico-ruído.; **HNRD:** Desvio Padrão da relação harmônico ruído; **PA:** Amplitude da Frequência; **SFR:** Planicidade espectral do sinal residual; **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**); SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

5.3 CORRELAÇÃO ENTRE ATITUDES DOS OUVINTES E AS MEDIDAS ACÚSTICAS E DO JPA

De modo geral, houve correlação negativa forte entre o julgamento das atitudes dos ouvintes e a EAV-GG e EAV-R, assim como correlação negativa moderada entre o julgamento de atitudes e a EAV-S e EAV-T (Tabela 12). A EAV-GG e EAV-R demonstrou correlação negativa forte no julgamento de atitudes relacionados à agradabilidade, poder simpatia, força, resistência, extroversão, saúde, autoridade, segurança, competência e independência. A EAV-S e EAV- T demonstrou correlação negativa moderada forte no julgamento de atitudes relacionados à agradabilidade, poder simpatia, força, resistência, extroversão, saúde, autoridade, segurança, competência e independência. Destaca-se que o atributo calmo não apresentou resultados significantes/força de correlação.

TABELA 12: Correlação entre a média do julgamento de atitudes realizado pelos dos juízes e valores obtidos na EAV.

Variáveis	EAV - GG		EAV-R		EAV-S		EAV-T	
	Coeficiente de correlação	p-valor	Coeficiente de correlação	p-valor	Coeficiente de correlação	p-valor	Coeficiente de correlação	p-valor
Média das atitudes	-0,872	<0,001*	-0,800	<0,001*	-0,648	<0,001*	-0,430	0,004*
AGRADABILIDADE	-0,891	<0,001*	-0,820	<0,001*	-0,608	<0,001*	-0,480	<0,001*
PODER	-0,853	<0,001*	-0,785	<0,001*	-0,680	<0,001*	-0,390	0,009*
SIMPATIA	-0,757	<0,001*	-0,692	<0,001*	-0,607	<0,001*	-0,337	0,025*
FORÇA	-0,827	<0,001*	-0,754	<0,001*	-0,689	<0,001*	-0,365	0,015*
RESISTÊNCIA	-0,833	<0,001*	-0,757	<0,001*	-0,668	<0,001*	-0,387	0,009*
EXTROVERSÃO	-0,706	<0,001*	-0,654	<0,001*	-0,645	<0,001*	-0,245	0,109*
SAÚDE	-0,881	<0,001*	-0,815	<0,001*	-0,644	<0,001*	-0,461	0,002*
AUTORIDADE	-0,758	<0,001*	-0,691	<0,001*	-0,637	<0,001*	-0,34	0,021*
CALMA	0,063	0,685	0,086	0,577	0,438	0,003	-0,253	0,098
SEGURANÇA	-0,791	<0,001*	-0,717	<0,001*	-0,623	<0,001*	-0,377	0,012*
COMPETÊNCIA	-0,838	<0,001*	-0,750	<0,001*	-0,601	<0,001*	-0,431	0,004*
INDEPENDÊNCIA	-0,797	<0,001*	-0,770	<0,001*	-0,478	0,001*	-0,401	0,007*

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.
Legenda: **EAV:** Escala Analógico-visual**; GG**: grau geral**; GR:** grau de rugosidade**; GS:** grau de soprosidade**; GT:** grau de tensão**.**

Na tabela 13 observa-se que houve correlação positiva moderada entre o julgamento das atitudes dos ouvintes e as medidas acústicas: CPP-CS, CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, PA-CS, correlação negativa fraca entre o julgamento das atitudes dos ouvintes e as medidas acústicas: f0média-CS, f0q3-CS, JitterLoc-CS e SNL6-CS. Correlação positiva fraca entre o julgamento das atitudes dos ouvintes e as medidas

acústicas: f0min-CS, Hfno-CS e HNRmean-CS. Correlação negativa moderada entre o julgamento das atitudes dos ouvintes e as medidas acústicas: Spectral Tilt CS, F0SD-CS, F0CV-CS, PSD-CS,LNPSD-CS, Jitter Rrap-CS,Jitter PPQ5-CS,Jitter rddp-CS, ShimmerLoc-CS, ShimmerdB-CS,ShimmerAPQ3-CS, ShimmerAPQ5-CS,Shimmerdda-CS e AVI-CS.

TABELA 13: Correlação entre média do julgamento realizado pelos dos juízes e parâmetros acústicos.

Variáveis	Média do julgamento de atitudes	
	Coeficiente de correlação	p-valor
CPP-CS	0,634	<0,001*
CPPS-CS	0,677	<0,001*
Spectral tilt-CS	-0,540	<0,001*
F0SD-CS	-0,547	<0,001*
F0 media-CS	-0,313	0,041*
F0 q3-CS	-0,359	0,018*
F0 CV-CS	-0,468	0,002*
F0 min-CS	0,340	0,026*
PSD-CS	-0,468	0,002*
LNPSD-CS	-0,487	<0,001*
Jitter Loc-CS	-0,397	0,08*
Jitter rRAP-CS	-0,497	<0,001*
Jitter PPQ5-CS	-0,438	0,003*
Jitter rddp-CS	-0,482	0,001*
ShimmerLoc-CS	-0,577	<0,001*
ShimmerdB-CS	-0,580	<0,001*
ShimmerAPQ3-CS	-0,683	<0,001*
ShimmerAPQ5-CS	-0,602	<0,001*
Shimmerdda-CS	-0,684	<0,001*
AVI-CS	-0,483	0,001*
Autocorrelation-CS	0,513	<0,001*
GNE1000Hz-CS	0,596	<0,001*
GNE2000Hz-CS	0,596	<0,001*
GNE3000Hz-CS	0,598	<0,001*
Hfno-CS	0,337	0,027*
HNRmean-CS	0,373	0,014*
PA-CS	0,445	0,003*
SNL6-CS	-0,316	0,039*

*Valores significativos (p<0,005) – Teste: Correlação de Pearson.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma. **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico. **H1H2:** Razão da amplitude do primeiro e segundo harmônico. **fo média**: Média da Frequência Fundamental. **foDP**: Desvio Padrão da Frequência Fundamental. **Fo mediana**: Mediana da Frequência Fundamental. **foq3:** 3º Quartil da Frequência Fundamental. **foCV:** Coeficiente de variação da Frequência Fundamental. **fomin:** Mínima da Frequência Fundamental. **fo max**: Máxima da Frequência Fundamental. **PSD:** Densidade Espectral de potência de ruído. **LNPSD:** Logaritmo do desvio padrão do período. **Jitter LOC:** Média de perturbação relativa. **Jitter ABS:** perturbação média relativa. **Jitter RAP:** Perturbação média relativa em porcentagem. **Jitter PPQ5:**

Coeficiente médio de perturbação.**Jitter DDP**: Diferença média absoluta entre períodos consecutivos. **Shimmer Loc:** Média de perturbação relativa. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3. **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz. **GNE2:** Taxa de excitação glotal-para-ruído com largura de banda de 2000 Hz. **GNE3:** Taxa de excitação glotal-para-ruído com largura de banda de 3000 Hz. **Hfno:** Nível de energia relativa do ruído de alta frequência. **HNR media:** Média da relação harmônico-ruído. **HNRD:** Desvio Padrão da relação harmônico ruído. **PA:** Amplitude da Frequência. **SNL6:** Nível de ruído espectral = 5100–8000 Hz.

Na tabela 14, observa-se correlação positiva moderada entre o julgamento de atitudes relacionado à **agradabilidade** e as medidas CPP-CS,CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, Ffno-CS,HNRmean-CS, PA-CS, correlação negativa forte entre o julgamento de atitudes relacionado à agradabilidade e as medidas acústicas Shimmer APQ3-CS e Shimmer dda-CS, correlação negativa moderada entre o julgamento de atitudes relacionado à agradabilidade e as medidas Spectral Tilt-CS, F0 SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterLoc-CS, Jitter-RAP-CS, Jitter PPQ5-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS, ShimmerAPQ5-CS e AVI-CS, correlação positiva fraca entre o julgamento de atitudes relacionado à agradabilidade e as medidas H1A1-CS e F0 min, por fim, observa-se correlação fraca negativa entre o julgamento de atitudes relacionado à agradabilidade e as medidas F0q3-CS, ShimmerAPQ11-CS, SFR-CS e a medida SNL6-CS.

Com relação ao julgamento de atitudes relacionado ao atributo **poder**, observa-se correlação positiva forte com a medida CPPS-CS, correlação moderada com as medidas CPP-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, PA-CS, correlação moderada negativa com as medidas Spectral Tilt-CS, F0SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterLoc-CS, JitterRAP-CS, JitterPPQ5-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS , ShimmerAPQ5-CS, Shimmer dda-CS e AVI-CS, correlação positiva fraca com as medidas F0 min, PER-CS, Ffno-CS,HNRmean-CS, por fim, correlação negativa fraca com as medidas F0 mean-CS, F0 median-CS, F0q3-CS e a medida SFR-CS (Tabela 14).

Observa-se também correlação positiva moderada entre o julgamento de atitudes relacionados ao atributo **simpatia** e as medidas CPP-CS,CPPS-CS, GNE1000Hz-CS, GNE2000Hz-CS e GNE3000Hz-CS, correlação negativa moderada entre o julgamento de atitudes relacionados ao atributo simpatia e as medidas Spectral Tilt-CS, JitterRAP-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS , ShimmerAPQ5-CS e a medida Shimmer dda-CS, correlação positiva fraca com as medidas F0 min, Autocorrelation-CS e a medida PA-CS, por fim, correlação negativa fraca com as medidas F0SD-CS, PSD-CS, LNPSD-CS, JitterLoc-CS, JitterPPQ5-CS, AVI-CS e PA-CS (Tabela 14).

Com relação ao julgamento de atitudes relacionado ao atributo **força**, observa-se correlação positiva forte com a medida CPPS-CS, correlação moderada com as medidas CPP-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, F0SD-CS, F0q3-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterRAP-CS, JitterPPQ5-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS, ShimmerAPQ5-CS, Shimmer dda-CS e AVI-CS, correlação positiva fraca com as medidas F0 min-CS, PER-CS e a medida PA-CS. Por fim, observa-se correlação negativa fraca com as medidas F0 mean-CS, F0 median-CS e a medida JitterLoc-CS (Tabela 14).

Ao analisar o julgamento de atitudes relacionado ao atributo **resistência**, observa-se correlação positiva forte com a medida CPPS-CS, correlação moderada com a medidas CCP-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, F0 SD-CS, F0q3-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterRAP-CS, JitterPPQ5-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, Shimmer dda-CS e AVI-CS, correlação positiva fraca com as medidas F0 min, PER-CS, Ffno-CS e PA-CS. Por fim, observa-se correlação negativa fraca com as medidas F0 mean-CS, F0 median-CS e JitterLoc-CS (Tabela 14).

Com relação ao julgamento de atitudes relacionado ao atributo **extroversão**, observa-se correlação positiva moderada com as medidas CPP-CS, CPPS-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, JitterRAP-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS, ShimmerAPQ5-CS,

Shimmer dda-CS, correlação positiva fraca com as medidas F0 min, Autocorrelation-CS e PA-CS. Por fim, observa-se correlação negativa fraca com as medidas F0SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, ShimmerPPQ5-CS, JitterLoc-CS, JitterPPQ5-CS, AVI-CS (Tabela 14).

Também há correlação positiva moderada entre o julgamento de atitudes relacionados ao atributo **saúde** e as medidas CPP-CS,CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS,PA-CS, HNRmean-CS, PA-CS, correlação negativa forte com as medidas Shimmer APQ3-CS e Shimmer dda-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, F0SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterRAP-CS, Jitterddp-CS, ShimmerPPQ5-CS, ShimmerLoc-CS, ShimmerdB-CS, ShimmerAPQ5-CS, AVI-CS e JitterLoc-CS, correlação positiva fraca com as medidas F0 min-CS e Hfno-CS. Por fim, observa-se correlação negativa fraca com as medidas F0 mean-CS, F0q3-CS, SFR-CS e a medida SNL6-CS (Tabela 14).

Ao analisar o julgamento de atitudes relacionado ao atributo **autoridade**, observa-se correlação positiva moderada com as medidas CPP-CS , CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, correlação negativa moderada com as medidas acústicas Spectral Tilt-CS, F0SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, , JitterRAP-CS, Jitterddp-CS, JitterPPQ5-CS, ShimmerLoc-CS, ShimmerdB-CS, Shimmer APQ3-CS, ShimmerAPQ5-CS, Shimmer dda-CS, AVI-CS, correlação positiva fraca com as medidas acústicas F0 min-CS, PER-CS, PA-CS, correlação negativa fraca com as medidas F0 mean-CS, JitterLoc-CS e a medida SFR-CS (Tabela 14).

Diante do atributo relacionado à **calma**, é observado apenas correlação positiva moderada com as medidas HA1H1-CS, H1H3-CS e H1H2-CS (Tabela 14).

Com relação ao julgamento de atitudes relacionado ao atributo **segurança,** observa-se correlação positiva moderada com as medidas CPP-CS, CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, F0SD-CS, F0medianSD-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterRAP-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, Shimmer dda-CS, AVI-CS, correlação positiva fraca com as medidas F0 min-CS

e PA-CS, por fim, apresentou correlação negativa fraca com as medidas F0q3-CS, JitterLoc-CS e a medida JitterPPQ5-CS (Tabela 14).

Ao analisar o julgamento de atitudes relacionado ao atributo **competência,** observa-se correlação positiva moderada com as medidas CPP-CS,CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, PA-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, F0SD-CS, F0CV-CS, PSD-CS, LNPSD-CS, JitterRAP-CS, Jitterddp-CS, ShimmerLoc-CS, ShimmerdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, Shimmerdda-CS, AVI-CS e PA-CS, correlação positiva fraca com as medidas F0 min, Hfno-CS, HNRmean-CS, por fim, apresentou correlação negativa fraca com as medidas F0 mean-CS, F0 median-CS, F0q3-CS, JitterLoc-CS, JitterPPQ5-CS e a medida SFR-CS (Tabela 14).

Ao analisar o julgamento de atitudes relacionado ao atributo **independência,** observa-se correlação positiva Moderada com as medidas CPP-CS,CPPS-CS, Autocorrelation-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS, HNRmean-CS,PA-CS, correlação negativa moderada com as medidas Spectral Tilt-CS, ShimmerLoc-CS, ShimmerdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, LNPSD-CS, Shimmerdda-CS e a correlação negativa fraca com as medidas F0SD-CS, F0CV-CS, PSD-CS, JitterLoc-CS, JitterPPQ5-CS,AVI-CS, JitterRAP-CS, JitterPPQ5-CS e Jitterddp-CS (Tabela 14).

TABELA 14: Correlação entre média da avaliação/julgamento de atitudes realizado pelos ouvintes e parâmetros acústicos.

VARIÁVEIS		Agradabilidade	Poder	Simpatia	Força	Resistência	Extroversão	Saúde	Autoridade	Calma	Segurança	Competência	Independência
CPP-CS	**Coef.Cor.**	0,615	0,643	0,553	0,611	0,638	0,560	,636	0,586	-0,095	0,594	0,602	0,493
	p-valor	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<,001*	<0,001*	0,543	<0,001*	<0,001*	<0,001*
CPPS-CS	**Coef.Cor.**	0,657	0,713	0,529	0,710	0,701	0,566	,686	0,667	-0,220	0,630	0,631	0,517
	p-valor	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<,001*	<0,001*	0,156	<0,001*	<0,001*	<0,001*
Spectral	**Coef.Cor.**	-0,070	0,077	0,086	0,113	0,081	0,232	-,003	0,139	-0,550	0,075	0,004	0,014
decline-CS	**p-valor**	0,654	0,625	0,582	0,472	0,604	0,135	,982	0,376	<0,001*	0,632	0,979	0,928
Spectral	**Coef.Cor.**	-,0544	-0,556	-0,439	-0,534	-,0544	-0,423	-,570	-,0494	0,061	-0,487	-0,513	-0,418
tilt-CS	**p-valor**	<0,001*	<0,001*	0,003*	<0,001*	<0,001*	0,005*	<,001*	<0,001*	0,700	<0,001*	<0,001*	0,005*
H1A1-CS	**Coef.Cor.**	0,330	0,177	0,141	0,125	0,159	0,038	,278	0,124	0,479	0,189	0,247	0,268
	p-valor	0,033*	0,263	0,372	0,429	0,314	0,811	,074	0,432	0,001*	0,231	0,114	0,086
H1A3-CS	**Coef.Cor.**	0,276	0,084	0,108	0,036	0,063	-0,054	,188	0,001	0,576	0,084	0,159	0,205
	p-valor	0,077	0,598	0,497	0,822	0,692	0,735	,233	0,993	<0,001*	0,598	0,314	0,193
H1H2-CS	**Coef.Cor.**	0-,115	-0,217	-0,201	-0,235	-0,203	-0,302	-,123	-0,222	0,502	-0,165	-0,112	-0,144
	p-valor	0,469	0,167	0,202	0,135	0,198	0,052	,438	0,158	<0,001*	0,295	0,478	0,364
fo **mean-CS**	**Coef.Cor.**	-0,251	-0,329	-0,100	-0,343	-0,372	-0,086	-,319	-0,314	-0,181	-0,289	-0,314	-0,081
	p-valor	0,104	0,031*	0,524	0,024*	0,014*	0,584	,037*	,0041*	0,245	0,060	0,040*	0,606
$F0_{SD\text{-}CS}$	**Coef.Cor.**	-0,562	-0,564	-0,315	-0,542	-0,578	-0,318	-,594	-0,541	-0,133	-0,519	-0,566	-0,340
	p-valor	<0,001*	<0,001*	0,039*	<0,001*	<0,001*	0,038*	<,001*	<0,001*	0,394	<0,001*	<0,001*	0,026*
F0 **median-CS**	**Coef.Cor.**	-0,280	-0,347	-0,122	-0,360	-0,387	-0,101	-,341	-0,325	-0,198	-0,304	-0,332	-0,111
	p-valor	0,069	0,022*	0,437	,018*	,010*	0,519	,025*	0,033*	0,203	0,047*	0,030*	0,479
$F0_{q1\text{-}CS}$	**Coef.Cor.**	-0,0170	-0,245	-0,056	-0,268	-0,292	-0,040	-,239	-0,232	-0,184	-0,217	-0,237	-0,045
	p-valor	0,276	0,114	0,723	0,083	0,058	0,801	,122	0,134	0,237	0,161	0,126	0,775
$F0_{q3\text{-}CS}$	**Coef.Cor.**	-0,334	-0,398	-0,151	-0,403	-0,433	-0,137	-,390	-0,374	-0,173	-0,348	-0,376	-0,142
	p-valor	0,028*	0,008*	0,334	0,007*	0,004*	0,381	,010*	0,013*	0,267	0,022*	0,013*	0,363
$F0_{CV\text{-}CS}$	**Coef.Cor.**	-0,500	-0,475	-0,289	-0,443	-0,474	-0,302	-,513	-0,461	-0,053	-0,440	-0,483	-0,312
	p-valor	<0,001*	0,001*	0,060	0,003*	0,001*	0,049*	<,001*	0,002*	0,737	0,003*	0,001*	0,042*
$F0_{min\text{-}CS}$	**Coef.Cor.**	0,368	0,334	0,340	0,338	0,331	0,322	,361	0,352	-0,169	0,357	0,381	0,105

$F0_{max-CS}$	**p-valor**	0,015*	0,028*	0,025*	0,026*	0,030*	0,035*	,017*	0,021*	0,280	0,019*	0,012*	0,503
	Coef.Cor.	-0,078	-0,108	0,033	-0,114	-0,161	-0,003	-,172	-0,169	-0,010	-0,120	-0,150	0,110
PER-CS	**p-valor**	0,620	0,491	0,834	0,466	0,302	0,987	,269	0,279	0,951	0,442	0,337	0,483
	Coef.Cor.	0,219	0,332	0,090	0,346	0,373	0,091	,298	0,327	0,102	0,287	0,296	0,052
PSD-CS	**p-valor**	0,159	0,030*	0,565	0,023*	0,014*	0,563	,052	0,032*	0,517	0,062	0,054	0,740
	Coef.Cor.	-0,526	-0,466	-0,356	-0,431	0-,441	-0,348	-,491	-0,426	0,008	-0,417	-0,451	-0,394
LNPSD-CS	**p-valor**	<0,001*	0,002*	0,019*	0,004*	0,003*	0,022*	<,001*	0,004*	0,961	0,005*	0,002*	0,009*
	Coef.Cor.	-0,562	-0,467	-0,383	-0,431	-0,438	-0,349	-,520	-0,416	-0,060	-0,425	-0,465	-0,447
JitterLoc-CS	**p-valor**	<0,001*	0,002*	0,011*	0,004*	0,003*	0,022*	<0,001*	0,006*	0,704	0,005*	0,002*	0,003*
	Coef.Cor.	-0,432	-0,404	-0,326	-0,389	-0,376	-0,361	-0,410	-0,383	0,139	-0,361	-0,361	-0,313
JitterABS-CS	**p-valor**	0,004*	0,007*	0,033*	0,010*	0,013*	0,017*	0,006*	0,011*	0,374	0,017*	0,017*	0,041*
	Coef.Cor.	-0,185	-0,113	-0,180	-0,101	-0,079	-0,196	-0,139	-0,105	0,130	-0,109	-0,110	-0,170
JitterRAP-CS	**p-valor**	0,235	0,471	0,248	0,521	0,613	0,208	0,373	0,501	0,407	0,487	0,483	0,277
	Coef.Cor.	-0,510	-0,516	-0,440	-0,508	-0,489	-0,463	-0,498	-0,493	0,226	-0,471	-0,466	-0,353
JitterPPQ5-CS	**p-valor**	<0,001*	<0,001*	0,003*	<0,001*	<0,001*	0,002*	<0,001*	<0,001*	0,145	0,001*	0,002*	0,020*
	Coef.Cor.	-0,470	-0,447	-0,363	-0,433	-0,423	-0,398	-0,451	-0,418	0,160	-0,396	-0,396	-0,347
Jiterddp-CS	**p-valor**	0,001*	0,003*	0,017*	0,004*	0,005*	0,008*	0,002*	0,005*	0,305	0,008*	0,009*	0,023*
	Coef.Cor.	-0,495	-0,499	-0,406	-0,494	-0,476	-0,449	-0,490	-0,480	0,218	-0,455	-0,448	-0,346
ShimmerLoc-CS	**p-valor**	<0,001*	<0,001*	0,007*	<0,001*	0,001*	0,003*	<0,001*	0,001*	0,161	0,002*	0,003*	0,023*
	Coef.Cor.	-0,621	-0,583	-0,442	-0,556	-0,561	-0,431	-0,602	-0,523	-0,025	-0,517	-0,544	-0,453
ShimmerdB-CS	**p-valor**	<0,001*	<0,001*	0,003*	<0,001*	<0,001*	0,004*	<0,001*	<0,001*	0,874	<0,001*	<0,001*	0,002*
	Coef.Cor.	-0,616	-0,586	-0,429	-0,559	-0,565	-0,426	-0,601	-0,533	-0,055	-0,527	-0,558	-0,453
ShimmerAPQ3-CS	**p-valor**	<0,001*	<0,001*	0,004*	<0,001*	<0,001*	0,004*	<0,001*	<0,001*	0,725	<0,001*	<0,001*	0,002*
	Coef.Cor.	-0,720	-0,699	-0,543	-0,675	-0,676	-0,524	-0,713	-0,621	0,031	-0,615	-0,637	-0,528
ShimmerAPQ5-CS	**p-valor**	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	0,842	<0,001*	<0,001*	<0,001*
	Coef.Cor.	-0,657	-0,603	-0,475	-0,571	-0,579	-0,445	-0,630	-0,530	-0,061	-0,533	-0,564	-0,482

	p-valor	<0,001*	<0,001*	0,001*	<0,001*	<0,001*	0,003*	<0,001*	<0,001*	0,700	<0,001*	<0,001*	0,001*
ShimmerAPQ11-CS	**Coef.Cor.**	-0,356	-0,236	-0,206	-0,208	-0,216	-0,159	-0,293	-0,195	-0,249	-0,212	-0,254	-0,271
	p-valor	0,019*	0,127	0,186	0,180	0,165	0,308	0,057	0,211	0,107	0,173	0,101	0,079
Shimmerdda-CS	**Coef.Cor.**	-0,720	-0,699	-0,543	-0,676	-0,676	-0,524	-0,713	-0,621	0,031	-0,615	-0,637	-0,528
	p-valor	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	0,843	<0,001*	<0,001*	<0,001*
AVI-CS	**Coef.Cor.**	-0,475	-0,523	-0,316	-0,508	-0,537	-0,350	-0,502	-0,506	0,065	-0,485	-0,512	-0,223
	p-valor	0,001*	<0,001*	0,039*	<0,001*	<0,001*	0,022*	<0,001*	<0,001*	0,679	<0,001*	<0,001*	0,151
Autocorrelation-CS	**Coef.Cor.**	0,601	0,497	0,379	0,453	0,471	0,322	0,547	0,424	0,230	0,438	0,491	0,429
	p-valor	<0,001*	<0,001*	0,012*	0,002*	0,001*	0,035*	<0,001*	0,005*	0,138	0,003*	<0,001*	0,004*
GNE1000Hz-CS	**Coef.Cor.**	0,605	0,610	0,466	0,589	0,582	0,441	0,613	0,528	0,010	0,525	0,546	0,531
	p-valor	<0,001*	<0,001*	0,002*	<0,001*	<0,001*	0,003*	<0,001*	<0,001*	0,949	<0,001*	<0,001*	<0,001*
GNE2000Hz-CS	**Coef.Cor.**	0,600	0,609	0,470	0,588	0,581	0,455	0,614	0,530	-0,026	0,520	0,543	0,557
	p-valor	<0,001*	<0,001*	0,001*	<0,001*	<0,001*	0,002*	<0,001*	<0,001*	0,869	<0,001*	<0,001*	<0,001*
GNE3000Hz-CS	**Coef.Cor.**	0,596	0,603	0,502	0,581	0,579	0,485	0,614	0,532	-0,073	0,526	0,547	0,565
	p-valor	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	<0,001*	0,641	<0,001*	<0,001*	<0,001*
Hfno-CS	**Coef.Cor.**	0,403	0,321	0,265	0,292	0,307	0,182	0,377	0,244	0,190	0,269	0,331	0,292
	p-valor	0,007*	0,036*	0,086	0,057	0,045*	0,243	0,013*	0,114	0,223	0,082	0,030*	0,058
HNRmean-CS	**Coef.Cor.**	0,487	0,334	0,267	0,286	0,299	0,180	0,412	0,254	0,331	0,290	0,346	0,404
	p-valor	<0,001*	0,028*	0,083	0,063	0,051	0,247	0,006*	0,100	0,030*	0,059	0,023*	0,007*
HNRDP-CS	**Coef.Cor.**	0,225	0,180	0,203	0,157	0,164	0,115	0,198	0,161	0,125	0,175	0,215	0,159
	p-valor	0,146	0,247	0,191	0,315	0,292	0,463	0,204	0,301	0,426	0,261	0,167	0,308
HNRD-CS	**Coef.Cor.**	0,258	0,192	0,244	0,170	0,150	0,268	0,201	0,165	-0,180	0,180	0,162	0,273
	p-valor	0,095	0,218	0,115	0,277	0,338	0,082	0,197	0,291	0,247	0,249	0,300	0,077
PA-CS	**Coef.Cor.**	0,520	0,410	0,327	0,376	0,387	0,322	0,469	0,387	0,119	0,390	0,431	0,430
	p-valor	<0,001*	0,006*	0,033*	0,013*	0,010*	0,035*	0,002*	0,010*	0,449	0,010*	0,004*	0,004*
SFR-CS	**Coef.Cor.**	-0,354	-0,298	-0,096	-0,274	-0,339	-0,054	-0,356	-0,257	-0,364	-0,267	-0,318	-0,112
	p-valor	0,020*	0,052	0,539	0,076	0,026	0,729	0,019*	0,096	0,016*	0,083	0,038*	0,476
SNL1-CS	**Coef.Cor.**	0,010	0,199	0,085	0,228	0,195	0,245	0,071	0,252	-0,613	0,166	0,082	-0,025
	p-valor	0,950	0,200	0,589	0,142	0,210	0,113	0,649	0,103	<0,001*	0,289	0,602	0,874
SNL2-CS	**Coef.Cor.**	-0,165	0,024	-0,060	0,060	0,020	0,101	-0,116	0,087	-0,583	0,003	-0,092	-0,173
	p-valor	0,290	0,877	0,702	0,704	0,898	0,520	0,458	0,577	<0,001*	0,984	0,557	0,268
SNL3-CS	**Coef.Cor.**	0,094	0,275	0,156	0,301	0,272	0,295	0,154	0,314*	-0,597	0,233	0,157	0,025
	p-valor	0,550	0,075	0,319	0,050	0,077	0,055	0,323	0,041	<0,001*	0,133	0,316	0,871
SNL4-CS	**Coef.Cor.**	0,071	0,255	0,140	0,281	0,254	0,286	0,136	0,300	-0,608	0,217	0,140	0,013
	p-valor	0,649	0,099	0,371	0,068	0,101	0,063	0,384	0,051	<0,001*	0,163	0,371	0,936
SNL5-CS	**Coef.Cor.**	-0,062	0,123	0,020	0,152	0,118	0,188	-0,003	0,185	-0,588	0,098	0,013	-0,066

	p-valor	0,693	0,431	0,901	0,331	0,452	0,229	0,984	0,236	<0,001*	0,531	0,934	0,674
SNL6-CS	**Coef.Cor.**	-0,389	-0,244	-0,265	-0,206	-0,247	-0,141	-0,371	-0,181	-0,399	-0,242	-0,332	-0,355
	p-valor	0,010*	0,114	0,086	0,184	0,110	0,366	0,014*	0,246	0,008*	0,119	0,030*	0,020*

*Valores significativos (p<0,005) – Teste T Student para amostras independentes.

Legenda: CCP: Proeminência do pico cepstral; **CPPS:** Proeminência do pico cepstral suavizada; **DECLINIO:** Inclinação geral do espectrograma; **TILT:** Inclinação da linha de regressão do espectrograma. **H1A1:** Amplitude do primeiro harmônico; **H1A3:** Amplitude do primeiro harmônico em relação ao terceiro harmônico. **H1H2:** Razão da amplitude do primeiro e segundo harmônico. **Fo média**: Média da Frequência Fundamental. **Fo DP**: Desvio Padrão da Frequência Fundamental. **Fo mediana**: Mediana da Frequência Fundamental. **Fo q1:** 1º Quartil da Frequência Fundamental. **Fo q3:** 3º Quartil da Frequência Fundamental. **Fo CV:** Coeficiente de variação da Frequência Fundamental. **Fo min:** Mínima da Frequência Fundamental. **Fo max**: Máxima da Frequência Fundamental. **PSD:** Densidade Espectral de potência de ruído. **LNPSD:** Logaritmo do desvio padrão do período. **Jitter LOC:** Média de perturbação relativa. **Jitter ABS:** perturbação média relativa. **Jitter RAP:** Perturbação média relativa em porcentagem. **Jitter PPQ5:** Coeficiente médio de perturbação. **Jitter DDP**: Diferença média absoluta entre períodos consecutivos. **Shimmer Loc:** Média de perturbação relativa. **Shimmer dB**: Mudança da amplitude pico a pico (Em dB); **Shimmer APQ3:** Quociente de perturbação de amplitude-3. **Shimmer APQ5:** Quociente de perturbação de amplitude-5; **ShimmerAPQ11**: Quociente de perturbação de amplitude-11; **Shimmer DDA:** Diferença média absoluta; **AVI:** Índice de variabilidade da amplitude; **GNE1:** Taxa de excitação lottal-para-ruído com largura de banda de 1000 Hz. **GNE2:** Taxa de excitação lottal-para-ruído com largura de banda de 2000 Hz. **GNE3:** Taxa de excitação lottal-para-ruído com largura de banda de 3000 Hz. **Hfno:** Nível de energia relativa do ruído de alta frequência. **HNR média:** Média da relação harmônico-ruído. **HNRD:** Desvio Padrão da relação harmônico ruído. **PA:** Amplitude da Frequência. **SFR:** Planicidade espectral do sinal residual. **SNL1:** Nível de ruído espectral = 100–2600 Hz; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SNL3:** Nível de ruído espectral = 100–5100 Hz; **SNL4:** Nível de ruído espectral = 100–8000 Hz**). SNL5:** nível de ruído espectral = 2600–5100 Hz**; SNL6:** Nível de ruído espectral = 5100–8000 Hz.

5.4 MODELOS DE PREDIÇÃO DA VARIABILIDADE NO JULGAMENTO DE ATITUDES DOS OUVINTES A PARTIR DAS MEDIDAS ACÚSTICAS E DO JPA

O modelo produzido cumpriu os pré-requisitos definidos. A análise resultou em um modelo estatisticamente significativo (Tabela 15). O modelo 3 demonstrou que o EAV-GG (β = -0,032; t = -11,003; p < 0,001), a f_0 min-CS (β = 0,011; t = 2,410; p = 0,021) e o SNL3-CS (β = 0,032; t = 2,291; p = 0,028) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas. O modelo com essas três medidas é capaz de explicar em 80,0% (R^2 = 0,800) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas.

TABELA 15: Resultado do modelo de regressão linear para predição da variabilidade do julgamento geral das atitudes dos ouvintes para vozes disfônicas e não disfônicas, a partir de medidas perceptivo-auditivas e acústicas.

VALÊNCIA		β	Erro	T	Sig.	R	R^2
Modelo 1	(Constante)	5,206	0,190	27,339	<0,001*		
	EAV - GG	-0,033	0,003	-10,621	<0,001*	0,859	0,732
Modelo 2	(Constante)	4,254	0,474	8,970	<0,001*		
	EAV - GG	-0,032	0,003	-10,461	<0,001*	0,876	0,755
	$f_0$min-CS	0,011	0,005	2,174	0,036*		
Modelo 3	(Constante)	3,379	0,590	5,723	<0,001*		
	EAV - GG	-0,032	0,003	-11,003	<0,001*	0,892	0,800
	$f_0$min-CS	0,011	0,005	2,410	0,021*		
	SNL3-CS	0,032	0,014	2,291	0,028*		

Significância p<0,05* - Modelo de Regressão Linear (Técnica Stepwise).
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; fomin-CS:** Mínima da Frequência Fundamental; **SNL3-CS:** nível de ruído espectral = 100–5100 Hz.

A análise resultou em um modelo estatisticamente significativo (Tabela 16). O modelo 3 demonstrou que a EAV-GG (β = -0,041; t = -11,349; p < 0,001), a f_0 min-CS (β = 0,014; t = 2,460; p = 0,019) e o JitterPPQ5-CS (β = - 0,454; t =

-2,142; p = 0,039) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão avaliação para o atributo de **agradabilidade**. O modelo com essas três medidas é capaz de explicar em 83,1% (R^2 =0,831) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 19).

A EAV-GG (β = -0,024; t = -6,887; p < 0,001) foi preditora do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, mediante o atributo de **simpatia**. O modelo com EAV-GG é capaz de explicar em 54,2% (R^2 = 0,542) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 19).

Com relação ao atributo **saúde**, o modelo 2 demonstrou que a EAV-GG (β = -0,053; t = -11,153; p < 0,001) e a f_0 min-CS (β = 0,020; t = 2,460; p = 0,014) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas. O modelo com essas duas medidas é capaz de explicar em 79,2% (R^2 =0,792) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 19).

Por fim, o modelo 3 demonstrou que a EAV-GG (β = -0,026; t = -6,276; p < 0,001), a Spectral Decline - CS (β = 0,078; t = 3,102; p = 0,004) e a F0min -CS (β = 0,015; t = 2,196; p = 0,034) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão avaliação para o atributo de extroversão. O modelo com essas três medidas é capaz de explicar em 59,9% (R^2 =0,599) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 19).

TABELA 16: Resultado do modelo de regressão linear para predição da variabilidade do julgamento das atitudes dos ouvintes na dimensão avaliação (agradabilidade, simpatia, saúde e extroversão), a partir de medidas perceptivo-auditivas e acústicas.

AGRADABILIDADE							
Modelo		**β**	**Erro**	**t**	**Sig.**	**R**	**R^2**
1	(Constante)	5,579	0,237	23,551	<0,001*	0,879	0,773
	EAV - GG	-0,045	0,004	-11,672	<0,001*		
2	(Constante)	4,114	0,571	7,209	<0,001*	0,900	0,811
	EAV - GG	-0,043	0,004	-11,825	<0,001*		
	$f_0$min-CS	0,017	0,006	2,780	0,008*		
3	(Constante)	4,665	0,604	7,729	<0,001*	0,912	0,831
	EAV - GG	-0,041	0,004	-11,349	<0,001*		
	F_0min-CS	0,014	0,006	2,460	0,019*		
	JitterPPQ5-CS	-0,454	0,212	-2,142	0,039*		
SIMPATIA							
1	(Constante)	4,711	0,212	22,271	<,001*	0,737	0,542
	EAV - GG	-0,024	0,003	-6,887	<,001*		
SAÚDE							
1	(Constante)	6,419	0,303	21,165	<0,001*	0,870	0,756
	EAV - GG	-0,055	0,005	-11,134	<0,001*		
2	(Constante)	4,660	0,739	6,303	<0,001*	0,890	0,792
	EAV - GG	-0,053	0,005	-11,156	<0,001*		
	F_0min-CS	0,020	0,008	2,577	0,014*		
EXTROVERSÃO							
1	(Constante)	4,692	0,277	16,921	<0,001*	0,680	0,463
	EAV - GG	-0,027	0,005	-5,873	<0,001*		
2	(Constante)	6,310	0,650	9,702	<0,001*	0,740	0,548
	EAV - GG	-0,027	0,004	-6,500	<0,001*		
	Spectral decline-CS	0,070	0,026	2,709	0,010*		
3	(Constante)	5,166	0,810	6,375	<0,001*	0,774	0,599
	EAV - GG	-0,026	0,004	-6,276	<0,001*		
	Spectral decline-CS	0,078	0,025	3,102	0,004*		
	F_0min-CS	0,015	0,007	2,196	0,034*		

Significância p<0,05* - Modelo de Regressão Linear (Técnica Stepwise)
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; fomin-CS:** Mínima da Frequência; Fundamental; **JitterPPQ5-CS:** Diferença média absoluta entre um período e a média; **Spectral Decline:** Inclinação geral do espectograma

O modelo produzido cumpriu os pré-requisitos definidos. A análise resultou em um modelo estatisticamente significativo (Tabela 17). O modelo 4 demonstrou que a EAV-GG (β = -0,037; t = -9,823; p < 0,001), a SNL3-CS (β = 0,067; t = 3,920; p = 0,001), a f_0 min-CS (β = 0,017; t = 2,915; p = 0,006) e a f_0q1-CS (β = -0,005; t = -2,152; p = 0,038) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão **potência** para o atributo de poder. O modelo com essas quatro medidas é capaz de explicar

em 81,7% (R^2 =0,817) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 20).

A EAV-GG (β = -0,036; t = -8,919; p < 0,001), a medida SNL3-CS (β =-0,077; t = 4,247; p < 0,001), a f0min-CS (β = 0,019; t = 3,038; p=0,004) e a f0q1 (β = - 0,006; t = -2,527; p=0,038) -foi preditora do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, mediante o atributo de **força**. O modelo com as quatro medidas é capaz de explicar em 80,2% (R^2 = 0,802%) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 20).

Com relação ao atributo **autoridade**, o modelo 4 demonstrou que a EAV-GG (β = -0,023; t = -6,561; p < 0,001) a medida SNL4-CS (β = 0,058; t = 3,663; p < 0,001), f0min-CS (β = 0,016; t = 2,828; p=0,008), f_0q1-CS (β = -0,004; t = -2,121; p = 0,041) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas. O modelo com essas quatro medidas é capaz de explicar em 70,8% (R^2 =0,708%) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 20).

Por fim, o modelo 3 demonstrou que a EAV-GG (β = -0,026; t = -8,894; p < 0,001) e a F0min -CS (β = 0,012; t = 2,588; p = 0,013) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão avaliação para o atributo de **competência**. O modelo com essas duas medidas é capaz de explicar em 70,2% (R^2 =0,702) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas (Tabela 20).

TABELA 17: Resultado do modelo de regressão linear para predição da variabilidade do julgamento das atitudes dos ouvintes na dimensão potência (poder, força, autoridade e competência), a partir de medidas perceptivo-auditivas e acústicas.

PODER							
Modelo		**B**	**Erro**	**t**	**Sig.**	**R**	**R^2**
1	(Constante)	5,375	0,261	20,632	<0,001*	0,837	0,700
	EAV - GG	-0,041	0,004	-9,666	<0,001*		
2	(Constante)	3,833	0,543	7,065	<0,001*	0,872	0,761
	EAV - GG	-0,041	0,004	-10,707	<0,001*		
	SNL3-CS	0,059	0,019	3,153	0,003*		
3	(Constante)	2,463	0,757	3,254	0,002*	0,891	0,794
	EAV - GG	-0,039	0,004	-10,672	<0,001*		
	SNL3-CS	0,061	0,018	3,477	0,001*		
	F0min-CS	0,015	0,006	2,455	0,019*		
4	(Constante)	2,770	0,737	3,759	<0,001*	0,904	0,817
	EAV - GG	-0,037	0,004	-9,823	<0,001*		
	SNL3-CS	0,067	0,017	3,920	<0,001*		
	F0min-CS	0,017	0,006	2,915	0,006*		
	f0q1-CS	-0,005	0,002	-2,152	0,038*		
FORÇA							
1	(Constante)	5,481	0,287	19,076	<0,001	0,809	0,654
	EAV - GG	-,041	0,005	-8,704	<0,001		
2	(Constante)	3,708	0,592	6,267	<0,001	0,855	0,731
	EAV - GG	-,041	0,004	-9,753	<0,001		
	SNL3-CS	,068	0,020	3,327	0,002		
3	(Constante)	2,211	0,825	2,680	0,011	0,876	0,768
	EAV - GG	-,039	0,004	-9,680	<0,001		
	SNL3-CS	,070	0,019	3,662	<0,001		
	fomin-CS	,016	0,007	2,458	0,019		
4	(Constante)	2,597	0,787	3,300	0,002		
	EAV - GG	-0,036	0,004	-8,919	<0,001	0,896	0,802
	SNL3-CS	0,077	0,018	4,247	<0,001		
	F0min-CS	0,019	0,006	3,038	0,004		
	foq1-CS	-,006	0,002	-2,527	0,016		
AUTORIDADE							
1	(Constante)	4,826	0,242	19,970	<0,001*	0,735	0,541
	EAV - GG	-0,027	0,004	-6,865	<0,001*		
2	(Constante)	3,588	0,477	7,518	<0,001*	0,790	0,624
	EAV - GG	-0,027	0,004	-7,544	<0,001*		
	SNL4-CS	0,050	0,017	2,929	0,006*		
3	(Constante)	2,341	0,691	3,386	0,002*	0,820	0,672
	EAV - GG	-0,026	0,003	-7,362	<0,001*		
	SNL4-CS	0,052	0,016	3,214	0,003*		
	fomin-CS	0,014	0,006	2,378	0,023*		
4	(Constante)	2,627	0,675	3,890	<0,001*	0,841	0,708
	EAV - GG	-0,023	0,004	-6,561	<0,001*		
	SNL4-CS	0,058	0,016	3,663	<0,001*		
	F0min-CS	0,016	0,006	2,828	0,008*		
	F0q1-CS	-0,004	0,002	-2,121	0,041*		
COMPETÊNCIA							
1	(Constante)	5,330	0,186	28,585	<0,001*	0,820	0,672
	EAV - GG	-0,028	0,003	-9,057	<0,001*		
2	(Constante)	4,244	0,454	9,344	<0,001*	0,849	0,720
	EAV - GG	-0,026	0,003	-8,984	<0,001*		
	F0min-CS	0,012	0,005	2,588	0,013*		

Significância p<0,05* - Modelo de Regressão Linear (Técnica Stepwise).
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; SNL3:** nível de ruído espectral = 100–5100 Hz **fomin-CS:** Mínima da Frequência Fundamental; **F0q1:** 1º Quartil da Frequência Fundamental; **SNL4:** nível de ruído espectral = 100–8000 Hz

A análise resultou em um modelo estatisticamente significativo (Tabela 18). O modelo 4 demonstrou que a EAV-GG (β = -0,035; t = -0,706; p < 0,001), a SNL3-CS (β = 0,070; t = 0,287; p < 0,001), a f_0 min-CS (β = 0,018; t = 0,224; p = 0,006) e a f_0q1-CS (β = -0,007; t = -0,222; p = 0,008) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão atividade para o atributo de **resistência**. O modelo com essas três medidas é capaz de explicar em 80,1% (R^2 =0,801%) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas.

A medida SNL3-CS (β = 0,044; t = 0,217; p=0,022), e a f0min-CS (β = 0,015; t = 0,220; p=0,023) foi preditora do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, mediante o atributo de **segurança**. O modelo com as 02 medidas é capaz de explicar em 68,5% (R^2 = 0,685) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas.

Com relação ao atributo **calma**, as medidas SNL1-CS (β = 0,044; t = -0,497; p < 0,001) e a medida H1A1-CS (β = 0,065; t = 0,279; p=0,040) foram preditoras do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, mediante o atributo calma. O modelo com as 02 medidas é capaz de explicar em 43,6% (R^2 =0,436) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas.

Por fim, o modelo 2 demonstrou que a EAV-GG (β = -0,037; t = -0,824; p < 0,001) e a F0max -CS (β = 0,002; t = 0,208; p = 0,027) foram preditores do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas, na dimensão avaliação para o atributo de independência. O modelo com essas duas medidas é capaz de explicar em 68,2% (R^2 =0,682) a variabilidade no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas.

TABELA 18: Resultado do modelo de regressão linear para predição da variabilidade do julgamento das atitudes dos ouvintes na dimensão atividade (resistência, segurança, calma e independência), a partir de medidas perceptivo-auditivas e acústicas.

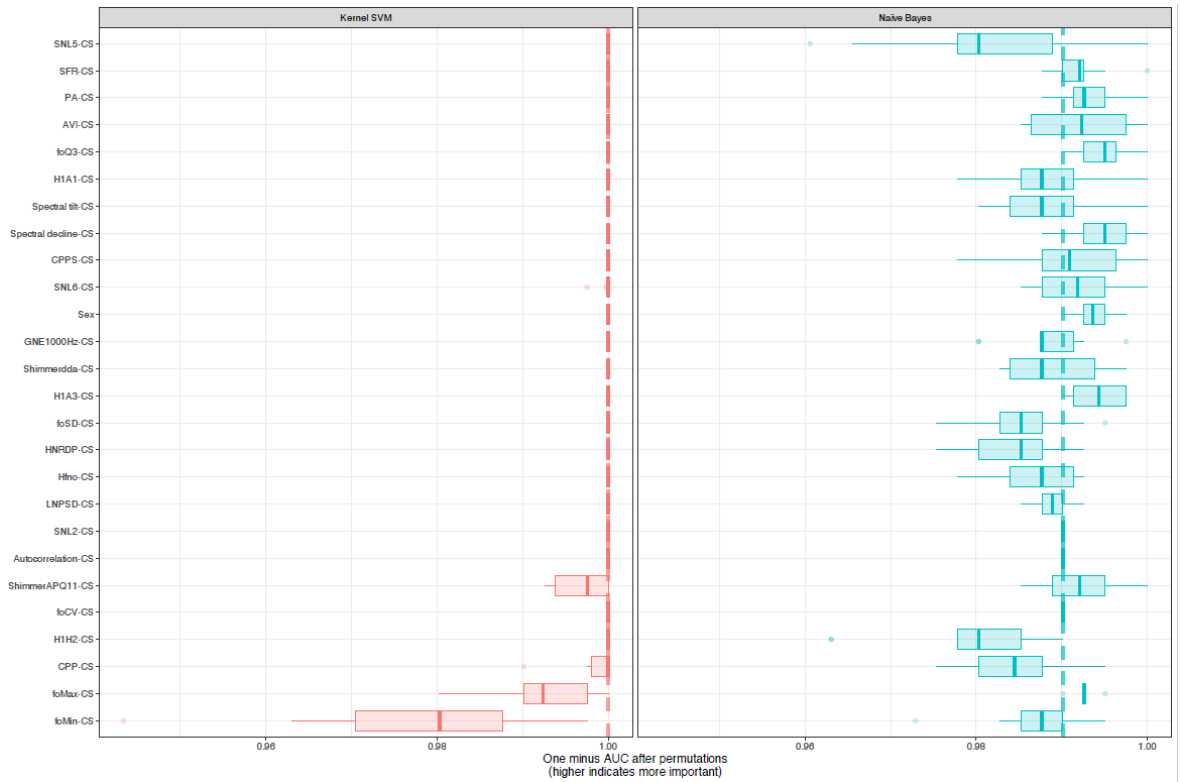
RESISTÊNCIA							
Modelo		**B**	**Erro**	**t**	**Sig.**	**R**	**R^2**
1	(Constante)	5,502	0,281	19,608	<0,001*	0,815	0,665
	EAV - GG	-0,041	0,005	-8,902	<0,001*		
2	(Constante)	3,938	0,593	6,644	<0,001*	0,851	0,725
	EAV - GG	-0,041	0,004	-9,722	<0,001*		
	SNL3-CS	0,060	0,020	2,929	0,006		
3	(Constante)	2,537	0,835	3,039	0,004	0,871	0,758
	EAV - GG	-0,039	0,004	-9,583	<0,001*		
	SNL3-CS	0,062	0,019	3,199	0,003		
	fomin-CS	0,015	0,007	2,273	0,029		
4	(Constante)	2,965	0,783	3,788	<0,001	0,895	0,801
	EAV - GG	-0,035	0,004	-8,895	<0,001*		
	SNL3-CS	0,070	0,018	3,860	<0,001*		
	F0min-CS	0,018	0,006	2,938	0,006*		
	F0q1-CS	-0,007	0,002	-2,816	0,008*		
SEGURANÇA							
1	(Constante)	5,245	0,256	20,455	<0,001*	0,772	0,596
	EAV - GG	-0,032	0,004	-7,685	<0,001*		
2	(Constante)	4,154	0,566	7,340	<0,001*	0,799	0,639
	EAV - GG	-0,032	0,004	-8,031	<0,001*		
	SNL3-CS	0,042	0,020	2,139	0,039		
3	(Constante)	2,771	0,793	3,493	0,001	0,828	0,685
	EAV - GG	-0,030	0,004	-7,860	<0,001*		
	SNL3-CS	0,044	0,018	2,385	0,022*		
	F0min-CS	0,015	0,006	2,363	0,023*		
CALMA							
1	(Constante)	5,432	0,294	18,452	<,0001*	0,609	0,371
	SNL1-CS	-0,070	0,014	-4,857	<0,001*		
2	(Constante)	5,415	0,282	19,177	<0,001*	0,661	0,436
	SNL1-CS	-0,057	0,015	-3,783	<0,001*		
	H1A1-CS	0,065	0,031	2,126	0,040*		
INDEPENDÊNCIA							
1	(Constante)	5,261	0,262	20,078	<0,001*	0,799	0,639
	EAV - GG	-0,036	0,004	-8,410	<0,001*		
2	(Constante)	4,732	0,340	13,932	<0,001*	0,826	0,682
	EAV - GG	-0,037	0,004	-9,057	<0,001*		
	F0max-CS	0,002	0,001	2,290	0,027*		

Significância p<0,05* - Modelo de Regressão Linear (Técnica Stepwise).
Legenda: **EAV:**Escala Analógico-visual**; GG**: grau geral**; SNL3:** nível de ruído espectral = 100–5100 Hz **fomin-CS:** Mínima da Frequência Fundamental; **F0q1:** 1º Quartil da Frequência Fundamental; **SNL1:** nível de ruído espectral = 100–2600 Hz; **SNL3:** nível de ruído espectral = 100–5100 Hz; **F0max-CS:** Frequência Fundamental Máxima; **H1A1:** Amplitude do primeiro harmônico.

5.5 MODELO BASEADO EM ML PARA PREDIÇÃO DA VALÊNCIA DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS E NÃO DISFÔNICAS DE FALANTES NATIVOS DO PORTUGUÊS BRASILEIRO, A PARTIR DAS MEDIDAS ACÚSTICAS

Considerando os critérios de desempenho estabelecidos no presente estudo, apenas os modelos utilizando os métodos de ML Kernel SVM e Naive Bayes foram selecionados. Ambos os modelos apresentaram acurácia e AUC ≥ 0.90, indicando um excelente desempenho na predição da valência das atitudes dos ouvintes em relação a vozes disfônicas e não disfônicas (Tabela 1). Na Figura 4 podem ser observados os Box Plots com relação à contribuição das medidas acústicas e da variável “sexo” nesses dois modelos.

FIGURA 06: Blox Pot representativo da importância das variáveis acústicas para o julgamento positivo ou negativo de vozes disfônicas e não disfônicas.



Legenda: SNL5: Nível de ruído espectral = 2600–5100 Hz; **SFR:** Planicidade espectral do sinal residual; **PA:** Amplitude da frequência; **AVI:** Índice de variabilidade da amplitude; **F0Q3:** 3º Quartil da Frequência Fundamental; **H1A1:** Razão da amplitude do primeiro e segundo harmônico; **SPECTRAL TILT:** Inclinação da linha de regressão do espectrograma ; **SPECTRAL DECLINE:** Inclinação geral do espectograma; **CPPS:** Proeminência do pico cepstral suavizada; **SNL6:** nível de ruído espectral = 5100–8000 Hz; **GNE1:** Taxa de excitação glotal-para-ruído com largura de banda de 1000 Hz.; **SHIMMER DDA:** Mudança da amplitude pico a pico; **HIA3:** Razão

da amplitude do primeiro e terceiro harmônico; **F0SD:** Desvio Padrão da frequência fundamental; **HNRDP:** Desvio Padrão da Relação Harmônico Ruído; **Hfno:** Nível de energia relativa do ruído de alta frequência; **LNPSD:** Logaritmo do desvio padrão do período; **SNL2:** Nível de ruído espectral = 100–3000 Hz; **SHIMMER apq11:** Quociente de perturbação de amplitude-11; **F0CV:** Coeficiente de variação da Frequência Fundamental. **H1H2:** Razão da amplitude do primeiro e segundo harmônico. **CCP:** Proeminência do pico cepstral. **f0 Max:** Máxima da Frequência Fundamental; **f0 Min:** Mínima da Frequência Fundamental.

O modelo de predição da valência das atitudes dos ouvintes baseado no método Naïve Bayes selecionou 22 medidas acústicas, entre as 47 medidas acústicas extraídas nesse estudo, além da variável “sexo”. Foram selecionadas as seguintes medidas acústicas: SNL5, SFR, PA, AVI, f_oQ3, H1A1, spectral tilt, declínio espectral, CPPS, SNL6, GNE 1000Hz, shimmer DDA, H1A3, f_o SD, HNRDP, Hfno, LNPSD, autocorrelação, shimmer apq11, H1H2, CPP e f_o min.

A partir da análise do boxplot do modelo baseado no método Naïve Bayes pode-se observar que todas as medidas acústicas e a variável “sexo” apresentaram medianas de acurácia superiores a 0.98, o que indica uma elevada capacidade preditiva geral do modelo. A consistência das medianas altas sinaliza uma seleção apropriada de características para a predição das atitudes dos ouvintes em relação a vozes disfônicas. A análise detalhada do boxplot será realizada na seção de discussão.

Por outro lado, o modelo de predição da valência das atitudes dos ouvintes baseado no método Kernel SVM foi mais parcimonioso. Tal modelo selecionou apenas quatro medidas acústicas, a saber: shimmer apq11, CPP, f_o max e f_o min (Figura 04). A escolha de um número menor de variáveis pode ser benéfica em termos de simplicidade do modelo e facilidade de interpretação, além de potencialmente reduzir o risco de sobreajuste.

A medida Shimmer apq11 demonstrou uma contribuição estável e consistente à acurácia do modelo, sugerida pelo seu intervalo interquartílico intermediário e pela simetria. O CPP destaca-se pela sua concentração de valores em torno de uma acurácia quase perfeita, indicada por um intervalo interquartílico muito reduzido e uma mediana altamente deslocada para o máximo da acurácia, reforçando sua confiabilidade como preditor. A medida fo max, com um intervalo interquartílico intermediário e assimetria do boxplot, reflete uma sensibilidade maior do modelo a valores mais baixos dessa medida.

Por fim, a medida fo min apresenta a maior variabilidade nas contribuições para a acurácia do modelo, como mostrado pelo seu amplo intervalo interquartílico e grandes bigodes, indicando que esta medida pode ser uma variável chave, sensível a nuances na percepção das vozes disfônicas pelos ouvintes e o seu julgamento de atitudes.

6 DISCUSSÃO

Na sessão de discussão desta tese, propõe-se uma análise e interpretação dos dados quanto à relação entre as medidas acústicas, JPA e as atitudes dos ouvintes para vozes disfônicas, e sobre o modelo encontrado para predição da valência das atitudes de ouvintes para vozes disfônicas. Esta investigação é importante para aprofundar a compreensão de como as variações acústicas incidem sobre a percepção e a valência atribuída à voz humana, e, consequentemente, como podem influenciar o julgamento social e o impacto da disfonia nesse contexto.

A proposta de entendimento destas dinâmicas é importante para o campo da fonoaudiologia e da comunicação humana, visto que a disfonia pode acarretar implicações negativas não apenas para a comunicação eficaz, mas também para a autoestima e o bem-estar psicossocial do indivíduo (Nemr K, *et al*., 2018; Krischke, *et al*., 2005). Além disso, este estudo contribui significativamente para a literatura existente, ao explorar a interrelação entre a avaliação objetiva da voz e a percepção dos ouvintes dentro de um contexto cultural específico, incluindo falantes nativos do português brasileiro.

A compreensão das nuances envolvidas nesse processo é fundamental para aprimorar a capacidade de diagnóstico e tratamento de indivíduos disfônicos. Além disso, os achados do presente estudo podem ser úteis para desenvolver intervenções mais fundamentadas e direcionadas para minimizar as repercussões no impacto social de vozes disfônicas. Dessa forma, conforme foi apresentado na sessão de resultados, os temas serão discutidos em subseções, a fim de facilitar o fluxo de leitura e compreensão dos temas abordados.

6.1 DIFERENÇAS NAS MEDIDAS ACÚSTICAS E DO JPA ENTRE VOZES COM VALÊNCIA POSITIVA E NEGATIVA

A qualidade vocal é considerada a característica da voz mais estável em um falante (MAIDMENT, 2008). Dessa forma, torna-se um elemento essencial de informação indexical, biológica ou emocional sobre o sujeito (COHEN S. M., 2010). Por sua vez, o desvio da qualidade vocal representa uma das principais manifestações da disfonia, podendo variar em relação ao grau da alteração

(classificada como leve, moderada ou intensa, por exemplo), bem como em relação tipo de desvio vocal predominante (tais como a rugosidade, a soprosidade, a tensão e a instabilidade, por exemplo). Nesse sentido, o desvio da qualidade vocal pode influenciar a percepção dos ouvintes e produzir uma impressão negativa do falante (AMIR e LEVINE, 2013).

No presente estudo, o GG se mostrou como um parâmetro vocal preditivo e que influencia no julgamento de atitude de ouvintes para vozes disfônicas, como também observado na literatura (SMITH E., *et al*., 2022; MORSOMME *et al*., 2011)

Os resultados encontrados demonstraram que os parâmetros de rugosidade, soprosidade e tensão do JPA são diferentes entre vozes percebidas pelos ouvintes como positivas ou negativas. As vozes sem alteração nesses parâmetros receberam julgamentos significativamente mais positivos do que as com alteração, bem como as valências das atitudes para estas vozes foram mais negativas.

Os achados referentes ao JPA demonstraram que um maior componente de ruído e irregularidade tem relação com a percepção e julgamento do ouvinte. A literatura (JIEUM *et al*., 2023; LEVITT e LUCAS, 2016; FRACARRO, 2011 2010 2013), apresenta discussões neste sentido, no entanto, eles não se referem ao predomínio da alteração, pois consideraram apenas os componentes rugosidade, soprosidade ou tensão de forma isolada. Apesar disso, de modo geral, é apresentado que esses parâmetros estão associados ao julgamento negativo dos ouvintes.

Um estudo prévio, demonstrou que indivíduos disfônicos nativos do português brasileiro foram avaliados de forma mais negativa em relação aos falantes vocalmente saudáveis nativos do português brasileiro (EVANGELISTA *et al.,* 2022). Observou-se que quanto maior a intensidade do desvio vocal, maior o julgamento negativo dos ouvintes em relação aos falantes. Além disso, as vozes predominantemente soprosas e tensas foram julgadas mais negativamente em relação às vozes saudáveis ou predominantemente rugosas.

Deve-se considerar que a percepção auditiva dos desvios da qualidade vocal está associada à informação acústica do sinal vocal que chega ao ouvinte. Desse modo, embora o JPA incorpore o quanto uma voz é considerada socialmente aceitável em termos da variabilidade da qualidade vocal, a

informação acústica, o JPA realizado pelo clínico e o impacto ocasionado no ouvinte são informações indissociáveis e influenciadas pela cultura e língua nativa do falante (ALTENBERG E.P. e FERRAND C.T., 2022; WAARAMAA T. *et al*., 2021). Nesse sentido, além de compreender a associação entre o resultado do JPA da qualidade vocal e as atitudes positivas ou negativas associadas aos falantes, o estudo dos parâmetros acústicos relacionados, ou que predizem, às atitudes dos ouvintes pode elucidar o entendimento sobre qual a informação acústica relevante para determinar a percepção, direcionando a intervenção clínica (LOPES *et al.*, 2019; LOPES *et al.,* 2018; HUGHES, PASTIZZO e GALLUP, 2008).

Os resultados encontrados quanto às medidas acústicas demonstraram que vozes avaliadas com valência positiva pelos ouvintes, apresentaram maiores valores nas medidas de CPP-CS, CPPS-CS, Spectral decline-CS, Spectral tilt-CS, em relação às vozes avaliadas negativamente pelos ouvintes; por outro lado, vozes avaliadas positivamente pelos ouvintes apresentaram menores valores em nas medidas de Jitterddp-CS, ShimmerLoc-CS, ShimmerrdB-CS, ShimmerAPQ3-CS, ShimmerAPQ5-CS, Shimmerdda-CS, GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS.

Tais achados podem indicar que vozes avaliadas positivamente tendem a apresentar maior componente harmônico, além de menor irregularidade e ruído em relação às vozes avaliadas negativamente. Provavelmente, essas características acústicas podem favorecer a percepção de atributos positivos relacionados a um falante, a partir da sua voz. De modo análogo, esses mesmos parâmetros acústicos estão associados à percepção de uma voz como disfônica ou não disfônicas (CONSERVA *et al*., 2022).

A literatura disponível tem demonstrado que as medidas acústicas estão associadas ao julgamento de atitudes dos ouvintes. As medidas da f0, *jitter* e *shimmer* estão entre as mais estudadas no campo das investigações com avaliação de atitudes (ANDERSON *et al.,* 2014; ROBINSON J. L., 2010). A f0 influenciou no julgamento de vozes masculinas, sendo que as vozes mais graves foram julgadas como mais atrativas, e de vozes femininas, em que as mais agudas foram avaliadas mais positivamente (ANDERSON *et al.,* 2014, SKRINDA *et al.,* 2014; MCALEER P., TODOROV A e BELIN P., 2014; BABEL, MCGUIRE

e KING, 2014; PISANSKI e RENDALL 2011; BLIGH M. C.; ROBINSON J. L., 2010).

A relação entre a presença de estrutura harmônica e de perturbação/ruído no sinal vocal também está associado ao julgamento positivo ou negativo dos falantes (BABEL *et al.,* 2014; LATINUS *et al*., 2011; BAUMANN *et al.*, 2008). Destaca-se que a presença de ruído em uma estrutura harmônica do sinal vocal, afeta a resposta cortical do córtex auditivo primário em regiões específicas que processam as vocalizações humanas (LEWIS *et al*., 2009 2022). Quanto maior regularidade desta estrutura, mais relevante é o som para o processamento cortical e maior a possibilidade de um indivíduo ser julgado positivamente (BRUCKKERT *et al*., 2010).

Do ponto de vista clínico, a identificação de parâmetros acústicos específicos associados à percepção negativa de vozes disfônicas fornece diretrizes objetivas para intervenções terapêuticas em fonoaudiologia (LOPES L. W. *et al*., 2019; LOPES L. W. *et al*., 2014). A reabilitação vocal pode se beneficiar do foco em reduzir medidas de *jitter*, *shimmer* e GNE e aumentar a proeminência do pico cepstral, por exemplo, visando a melhoria da qualidade vocal e o julgamento positivo da voz pelo ouvinte. Além disso, a compreensão das características acústicas que influenciam a valência das atitudes dos ouvintes pode auxiliar na criação de protocolos de avaliação vocal mais refinados, que considerem não apenas os aspectos fisiológicos da produção vocal, mas também as consequências psicossociais da disfonia.

Assim, a abordagem terapêutica pode ser mais holística, com alvos referentes à função vocal e metaterapia, englobando a psicodinâmica vocal, considerando o impacto da voz na qualidade de vida e na interação social do indivíduo, buscando estratégias que visem sanar as alterações vocais, mas também a promoção de uma voz que seja percebida positivamente no contexto social e comunicativo, em relação aos atributos das dimensões avaliação, potência e atividade. Dessa forma, as medidas acústicas e perceptivoauditivas se configuram como uma forma de monitoramento dos alvos propostos.

6.2 DIFERENÇAS NAS MEDIDAS ACÚSTICAS E DO JPA ENTRE VOZES COM VALÊNCIA POSITIVA E NEGATIVA NAS DIMENSÕES AVALIAÇÃO, POTÊNCIA E ATIVIDADE.

Esta subseção aborda as diferenças nas medidas acústicas e do JPA entre vozes com valência positiva e negativa nas três dimensões que são abordas no questionário para o julgamento de atitudes dos ouvintes utilizado nesta pesquisa: avaliação, potência e atividade. Essa análise buscou compreender as medidas acústicas influenciam, especificamente na percepção de cada atitude do ouvinte em relação ao falante. A compreensão dessas relações pode elucidar sobre o impacto de vozes disfônicas na comunicação interpessoal.

A presença do desvio da qualidade vocal em indivíduos disfônicos pode tornar a transmissão da mensagem menos efetiva, culminando em diferentes atitudes por parte do ouvinte (ALLARD E. R. & WILLIAMS D. F., 2008). Por isso, é importante conhecer os aspectos vocais que mais se relacionam com a percepção do ouvinte, para além da valência positiva ou negativa, mas observar a percepção de diferentes atributos relacionados às dimensões citadas anteriormente (AMIR e LEVINE, 2013; BORKOWSKA e PAWLOWSKI, 2011).

O presente estudo identificou que vozes avaliadas positivamente quanto aos atributos da dimensão avaliação, isto é, a agradabilidade, simpatia, saúde e extroversão, apresentaram menores GG, GR, GS e GT. Dessa forma, as vozes menos desviadas foram avaliadas mais positivamente pelos ouvintes em relação às vozes mais desviadas. Além disso, as medidas acústicas denotaram que vozes com valência positiva apresentaram maior componente de regularidade, estabilidade e menor componente de ruído, em comparação às vozes julgadas com valência negativa.

A partir desses resultados, podemos inferir que a qualidade vocal é um fator determinante na percepção de atitudes dos ouvintes. O desvio da qualidade vocal pode gerar conclusões negativas acerca competência comunicativa do indivíduo (AMIR e LEVINE-YUNDOF, 2013; ALLARD E. R. e WILLIAMS, 2008), tanto que vozes percebidas como mais periódicas e com mais componentes harmônicos, tendem a ser avaliadas mais positivamente, o que pode facilitar a comunicação interpessoal e a construção de relações sociais.

Por sua vez, as vozes com maior grau de rugosidade, de soprosidade e de tensão, podem ser associadas a uma atitude negativa dos ouvintes, o que pode afetar negativamente a interação social (PORCARO C., *et al*., 2020;

AHLANDER *et al*., 2015; XU, *et al*, 2011; VAN BORSEL, *et al*., 2009). Sendo assim, indivíduos com maior desvio da qualidade vocal identificado no JPA podem ser percebidos pelos ouvintes como menos competentes, fracos, frágeis e, consequentemente, apresentar menor atratividade vocal (BORKOWSKA e PAWLOWSKI 2011; SAXTON, *et al.*, 2006; MENDOZA, 2007; KIMBLE e SEIDEL, 1991).

Na comunicação interpessoal, a voz é um dos principais veículos de expressão das intenções, emoções e da identidade do falante (COHEN, 2010). As características vocais podem influenciar a primeira impressão, a credibilidade e a empatia entre os interlocutores. Portanto, entender como as características vocais afetam a percepção pode ajudar indivíduos a aprimorar suas habilidades comunicativas, especialmente em contextos profissionais ou sociais nos quais a voz desempenha um papel central.

Para fonoaudiólogos e otorrinolaringologistas, profissionais que atuam diretamente com a saúde vocal de profissionais da voz e da população em geral, esses achados reforçam a importância de avaliar e tratar as disfonias não apenas com foco nas alterações estruturais ou funcionais da laringe. O tratamento, seja ele médico e/ou fonoaudiológico, deve considerar o impacto psicossocial e causado pela disfonia, uma vez que as intervenções que visam melhorar a qualidade vocal também podem contribuir para uma melhor qualidade de vida e bem-estar dos pacientes.

Os resultados encontrados também contribuem para a Linguística, ao fornecer evidências empíricas de como as propriedades acústicas da voz estão relacionadas com a percepção social e com a comunicação eficaz. Dessa forma, pode-se expandir o entendimento da “pragmática vocal”, que diz respeito à forma como a voz é usada e percebida em contextos comunicativos e podem contribuir nas teorias sobre a relação entre linguagem, cognição e interação social (ALKMIM T. *et al*., 2004).

A dimensão “avaliação” relacionada às atitudes diz respeito ao julgamento de valor mais global que os ouvintes fazem sobre um falante, classificando-o em um continuum entre positivo e negativo. Na presente pesquisa, a dimensão “avaliação” foi representada pelas atitudes de agradabilidade, simpatia, saúde e extroversão.

Quanto à dimensão “avaliação”, nós observamos que indivíduos avaliados negativamente quanto às atitudes de agradabilidade, simpatia, saúde e extroversão apresentaram maior grau geral e graus de rugosidade, de soprosidade e de tensão em relação aos indivíduos avaliados positivamente nessas mesmas atitudes.

De modo geral, em relação às medidas acústicas, as vozes avaliadas negativamente nas atitudes da dimensão “avaliação” demonstraram valores das medidas acústicas mais compatíveis com aumento dos componentes de perturbação e ruído, assim como redução dos componentes harmônicos, em relação às vozes avaliadas positivamente nessas mesmas atitudes.

A agradabilidade refere-se à percepção dos ouvintes sobre a amabilidade e o prazer associado à voz do falante. Esta atitude engloba a sensação de conforto e satisfação ao escutar a voz, podendo influenciar a percepção de amizade e cordialidade que os ouvintes têm em relação ao falante. A simpatia está relacionada com a empatia e a capacidade do falante em estabelecer uma conexão emocional com o ouvinte. A simpatia envolve a percepção de afeto e de acessibilidade do falante, fazendo com que os ouvintes se sintam mais próximos e compreendidos. A saúde refere-se à percepção dos ouvintes sobre o estado de saúde do falante, baseada em características vocais. Por fim, a extroversão corresponde à percepção de sociabilidade e abertura do falante para com os outros (KUHNE, K. *et al*., 2022 2020; WAARAMAA, T., 2021; MCALEER P. *et al*., 2014; WILSON J. A., *et al*., 2002 2021)

Dessa forma, a partir dos achados do presente estudo, nós podemos indicar que indivíduos com vozes menos desviadas podem ser avaliados como mais agradáveis, simpáticos, saudáveis e extrovertidos em relação a indivíduos com vozes mais desviadas. Nesse contexto, a presença de menor componente de ruído e perturbação em uma voz parece influenciar a construção de relações sociais positivas e até a empatia entre os interlocutores (BORKOWSKA e PAWLOWSKI, 2011; MENDOZA, 2007; SAXTON *et al.*, 2006;).

Entre as medidas com diferença estatística significante na dimensão avaliação, destacou-se as relacionadas à f0 e de perturbação, em que as mais agudas e com maior variação de frequência foram consideradas mais negativas. Uma voz com uma frequência fundamental mais estável e com um maior componente de periodicidade pode ser percebida como mais agradável, o que

também pode influenciar positivamente a percepção de simpatia e extroversão (KYRIAKOU, PETINOU e PHINIKETTOS, 2018; VILLAR, KORN e AZEVEDO, 2016; RUMBACH *et al*., 2015).

A dimensão “potência” relaciona-se com a percepção de autoridade ou influência que um falante exerce sobre os ouvintes, sendo percebido como dominante ou submisso. Na presente pesquisa foram a dimensão “potência” incluiu as atitudes de poder, força, autoridade e competência.

Neste estudo, observou-se que as vozes avaliadas positivamente quanto às atitudes de força e autoridade exibiam menores valores no grau desvio vocal geral, além de menores graus de rugosidade, de soprosidade e de tensão quando comparadas às vozes com valência negativa nessas mesmas atitudes. Em relação às atitudes de poder e competência, as diferenças foram restritas ao grau geral de desvio vocal, e aos graus de rugosidade e de soprosidade. Especificamente, as vozes percebidas positivamente quanto às atitudes de poder e competência demonstraram valores inferiores nesses parâmetros em relação àquelas percebidas negativamente nessas atitudes.

As atitudes de poder, força, autoridade e competência são conceitos-chave na análise da comunicação e da percepção social, representando diferentes dimensões de avaliação que influenciam a forma como os indivíduos são percebidos pelos outros (KYRIAKOU, PETINOU e PHINIKETTOS, 2018). A atitude de poder refere-se à capacidade de influenciar ou controlar o comportamento de outras pessoas ou o curso dos eventos (KISSEL *et al*., 2022; EADIE *et al*., 2017).

Na comunicação, uma voz que transmite poder pode ser percebida como dominante ou assertiva. Indivíduos percebidos como poderosos são frequentemente vistos como capazes de impor sua vontade aos outros, seja através de status social, recursos ou força de personalidade. A atitude força está frequentemente associada à robustez física ou mental e à capacidade de resistir a desafios ou exercer influência em situações adversas (KORN e AZEVEDO, 2016).

Na comunicação interpessoal, a força pode ser percebida através de qualidades como a intensidade vocal, as ênfases e a precisão articulatória, transmitindo resiliência e determinação (VILLAR, KORN e AZEVEDO, 2016). A atitude autoridade relaciona-se com o direito de governar, comandar ou decidir,

muitas vezes concedido por uma posição social ou por conhecimento especializado (RUMBACH *et al*., 2015). Uma voz que reflete autoridade pode inspirar confiança e respeito, indicando que o falante possui conhecimento, experiência ou posição que justifica a liderança ou o respeito dos outros. Por fim, a atitude competência refere-se à habilidade, conhecimento ou capacidade para realizar determinadas tarefas eficazmente A competência pode ser comunicada através da forma como alguém fala, incluindo o uso de terminologia específica, a clareza de expressão e a capacidade de articular ideias complexas de maneira compreensível. Indivíduos percebidos como competentes são vistos como capazes e eficientes em suas áreas de expertise.

Essas atitudes da dimensão “potência” não só moldam a percepção de um indivíduo pelos outros, mas também podem afetar o comportamento e as interações sociais. Entender como essas dimensões de avaliação são comunicadas e percebidas é crucial para a análise da comunicação humana e para o desenvolvimento de estratégias eficazes de comunicação e persuasão.

A partir dos achados do presente estudo, é evidente que indivíduos avaliados como fortes e autoritários apresentam vozes menos desviadas em todos os parâmetros do JPA. Isso sugere uma voz menos desviada pode estar associada a indivíduos que transmitem a impressão de força e autoridade. Tal percepção pode ser atribuída à expectativa social de que indivíduos em posições de liderança ou com atributos de força devem ser capazes de comunicar suas mensagens de forma clara e inequívoca, sem indícios de fraqueza ou hesitação, que poderiam ser inferidos a partir de vozes disfônicas (RUMBACH *et al*., 2015).

Por outro lado, nas atitudes de poder e competência, não houve diferença quanto ao grau de tensão. Isso indica que, embora o menor componente de rugosidade e soprosidade também seja valorizado nessas atitudes, a tensão fonatória pode não ser um fator decisivo na avaliação da competência ou do poder. A ausência de diferença significativa no grau de tensão fonatória entre vozes avaliadas com valência positiva e negativa nas dimensões de poder e competência pode ser atribuída a diversos fatores relacionados à complexidade da percepção vocal e à especificidade dos contextos de comunicação. Primeiramente, é importante considerar que a percepção de poder e competência pode estar mais associada à redução de ruído e irregularidade em uma voz (refletidas em medidas dos graus de

rugosidade e de soprosidade) e menos dependente da tensão vocal, que muitas vezes é associada a estados emocionais ou a esforço vocal (KLOFSTAD *et al*., 2015).

Além disso, a tensão fonatória pode ser percebida de maneira ambígua, dependendo do contexto comunicativo e das expectativas do ouvinte (EVANGELISTA *et al*., 2022). Em situações em que a assertividade e a confiança são valorizadas, um certo nível de tensão vocal pode até ser interpretado positivamente, como um sinal de engajamento ou paixão. Por outro lado, em contextos em que a calma e a estabilidade emocional são mais valorizadas, a mesma tensão pode ser interpretada negativamente, por exemplo. Essa ambiguidade na interpretação da tensão fonatória pode contribuir para a ausência de diferenças claras relacionadas à valência nas atitudes de poder e competência.

Adicionalmente, a complexidade intrínseca da comunicação humana sugere que a avaliação das atitudes de poder e competência não se baseia exclusivamente em um único parâmetro vocal, mas em um conjunto de sinais acústicos e no contexto mais amplo da interação. Assim, enquanto a rugosidade e a soprosidade podem servir como indicadores mais diretos da valência atribuída a essas atitudes, a tensão fonatória pode não desempenhar um papel tão central nessa avaliação específica.

Portanto, a falta de diferença significativa no grau de tensão fonatória entre as avaliações de valência positiva e negativa nas atitudes de poder e competência pode refletir a multifacetada natureza da percepção vocal, onde a tensão fonatória não é o principal determinante da valência atribuída pelo ouvinte nesses contextos. Isso destaca a importância de considerar um espectro mais amplo de características vocais e contextuais ao investigar a forma como as vozes são percebidas e avaliadas socialmente.

Essas observações reforçam a ideia de que as vozes não são apenas veículos para a comunicação verbal, mas também poderosos indicadores de características sociais e psicológicas. As nuances na percepção vocal refletem as complexas dinâmicas de poder, autoridade, força e competência dentro das interações sociais. A capacidade de projetar uma voz que alinha com essas expectativas pode, portanto, influenciar significativamente a eficácia da

comunicação e as dinâmicas de poder em diversos contextos, desde o ambiente profissional até interações mais casuais.

A dimensão “atividade” diz respeito ao grau de dinamismo ou energia que um falante demonstra, influenciando o nível de engajamento e interesse dos ouvintes. Nessa dimensão foram incluídas as atitudes resistência, segurança, calma e independência.

Nós observamos diferenças em todas as medidas do JPA entre vozes avaliadas com valência positiva e negativa pelos ouvintes nas atitudes de resistência, segurança e independência. Indivíduos avaliados positivamente quanto às atitudes de resistência, segurança e independência apresentaram menores valores no grau geral de desvio vocal, assim como nos graus de rugosidade, de soprosidade e de tensão em relação às vozes avaliadas com valência negativa nessas mesmas atitudes. De modo geral, indivíduos julgados positivamente quanto às atitudes de resistência, segurança e independência apresentaram maiores valores nas medidas cepstrais, assim como menores valores nas medidas de perturbação e ruído em relação aos indivíduos avaliados negativamente nessas três atitudes. Especificamente, falantes avaliados positivamente quanto às atitudes de resistência e segurança apresentaram menor valor máximo da f_0 em relação aos falantes avaliados negativamente nessas duas atitudes.

Por sua vez, houve diferença apenas no grau de soprosidade entre os indivíduos avaliados com valência positiva e negativa na atitude calma. Pessoas avaliadas positivamente quanto à atitude calma apresentaram maior grau de soprosidade em relação aos falantes avaliados negativamente nessa atitude. Em termos acústicos, quanto à atitude calma, houve diferença entre falantes avaliados positiva e negativamente nessa atitude apenas em relação às medidas do CPPS e H1A3. Falantes avaliados positivamente quanto à atitude calma apresentaram menores valores do CPPS e maiores valores de H1A3 em relação aos falantes avaliados positivamente na atitude calma.

Na dimensão "atividade", as atitudes de resistência, segurança, calma e independência refletem diferentes aspectos do comportamento e da disposição de um indivíduo frente a situações e interações. A resistência pode ser interpretada como a capacidade de um indivíduo de se opor ou suportar circunstâncias adversas, pressões externas ou desafios (DUTTA M., 2013). Em

um contexto comportamental e comunicativo, a resistência pode se manifestar através de uma postura resiliente, uma fala com maior uso de ênfases, e a capacidade de manter a posição ou opinião diante da oposição. Características vocais associadas à resistência podem incluir a manutenção da intensidade vocal e uma articulação precisa, transmitindo determinação e força de vontade (SEBESTA P., 2017; BABEL, M., 2014) A segurança está relacionada à confiança e à ausência de dúvida sobre as próprias capacidades, decisões ou opiniões (MCCLEAN *et al*., 2021). A segurança pode ser expressa por meio de uma voz estável, ritmo constante e ausência de hesitações ou inflexões interrogativas, sugerindo autoconfiança e autoridade (EGELMAN & PEER, 2015). Em termos de atitude, a segurança implica uma postura assertiva e uma abordagem direta nas interações, refletindo a crença na validade das próprias ideias ou ações. A calma pode refletir serenidade, estabilidade emocional e a capacidade de manter a compostura em situações de tensão ou incerteza (LEVITT A. G. & LUCAS M., 2016) A calma pode ser evidenciada por características vocais como uma intensidade vocal reduzida, ritmo moderado e pausas controladas, que, juntos, transmite tranquilidade e controle. Em termos comportamentais, a calma se manifesta através de uma abordagem ponderada e reflexiva, evitando reações precipitadas ou emocionais. A independência denota autonomia, confiança na própria capacidade de tomar decisões e agir de acordo com as próprias convicções, independentemente das opiniões ou influências externas (STOKLASA J. *et al*., 2019; ZHENG *et al*. ,2020) Vocalmente, a independência pode ser indicada por um estilo de fala distintivo ou não convencional, que destaca a singularidade do indivíduo e sua disposição para se destacar. A atitude de independência é caracterizada por uma tendência à auto-suficiência, inovação e uma preferência por abordagens individuais em detrimento da conformidade.

Essas atitudes, quando consideradas dentro da dimensão "atividade", fornecem um entendimento sobre como os indivíduos se posicionam e agem em relação ao seu ambiente e às interações sociais. Elas refletem não só a maneira como as pessoas respondem fisicamente e emocionalmente a diferentes contextos, mas também como essas reações são percebidas e interpretadas por outros, influenciando a dinâmica social e a eficácia comunicativa.

Os resultados obtidos para a dimensão “atividade” evidenciam uma relação significativa entre as características acústicas da voz e a percepção das atitudes de resistência, segurança, independência e calma por parte dos ouvintes. A análise demonstrou que vozes avaliadas positivamente em relação às atitudes de resistência, segurança e independência são caracterizadas por qualidades vocais específicas, que incluem menores grau de desvio vocal, e dos graus de rugosidade, de soprosidade e de tensão. Além disso, houve maior componente de estrutura harmônica e menor componente de perturbação e ruído nas vozes avaliadas positivamente quanto à resistência, segurança e independência em relação àquelas avaliadas negativamente nessas atitudes. Essas características podem ser interpretadas à luz das funções comunicativas e emocionais que a voz desempenha no contexto social.

Vozes percebidas como menos desviadas, menos ruidosas e mais estáveis, as quais são associadas a avaliações positivas nas atitudes de resistência, segurança e independência, podem refletir uma maior competência vocal e emocional. Isso porque essas qualidades vocais tendem a ser interpretadas como sinais de confiança, assertividade e autocontrole, atributos intimamente vinculados às atitudes em questão (SMITH E. *et al*.,2022; MCALEER P. *et al*., 2014) O menor grau de rugosidade e de soprosidade, juntamente com os menores valores nas medidas de perturbação e ruído, sugerem uma maior regularidade na produção vocal, o que pode ser percebido como uma expressão de estabilidade emocional e firmeza (EVANGELISTA *et al*., 2022; AMIR *et al*., 2013).

A distinção entre as vozes avaliadas positivamente nas atitudes de resistência e segurança, que apresentaram menor valor máximo de f_o, pode ser interpretada como uma manifestação de autoridade e assertividade. Um valor máximo de f_o mais baixo sugere uma voz mais grave ou com menor gama tonal na fala conectada (SHIRAZI T. N. *et al*., 2018; SEBESTA P., *et al*., 2017; TSANTANI M. S. *et al*., 2017. Em termos de comunicação interpessoal e percepção social, um tom de voz mais grave tem sido associado a uma série de qualidades positivas, como autoridade, confiança e calma.

No contexto de atitudes de resistência e segurança, a apresentação vocal com um valor máximo de f_o mais baixo pode ser interpretada como uma expressão de confiança e firmeza (RAYMOND, M. *et al*., 2019; SKRINDA I. *et*

al., 2014) Isso se alinha com a percepção social que associa uma voz mais grave a maior autoridade e seriedade. Em situações que requerem a expressão de resistência ou segurança, uma voz mais grave pode ser mais eficaz em transmitir determinação e autoconfiança, características essenciais para persuadir ou influenciar a percepção do ouvinte (ZHENG Y., 2020).

A resistência, entendida como a capacidade de opor-se ou manter-se firme diante de adversidades, e a segurança, no sentido de confiança ou certeza em suas ações ou palavras, são ambas atitudes que se beneficiam da projeção de força e estabilidade. Uma voz que naturalmente carrega um tom mais grave pode, portanto, ser mais convincente e eficaz ao comunicar essas atitudes, possivelmente levando a uma avaliação mais positiva por parte dos ouvintes (ZHENG Y., 2020)

Além disso, um tom de voz mais grave pode também ser menos intrusivo e mais agradável aos ouvidos dos ouvintes, contribuindo para uma maior aceitação e apreciação da mensagem transmitida. Isso sugere que, além de transmitir confiança e autoridade, falantes com um valor máximo de f_o mais baixo podem ser percebidos como mais agradáveis e acessíveis, reforçando ainda mais a valência positiva associada às atitudes de resistência e segurança.

Portanto, o achado de que falantes avaliados positivamente em resistência e segurança apresentam um valor máximo de f_o mais baixo pode refletir uma adaptação natural da expressão vocal para maximizar a eficácia comunicativa em contextos que demandam a projeção de autoridade, confiança e estabilidade. Este resultado ressalta a complexidade da comunicação vocal e como as propriedades acústicas da voz podem ser finamente ajustadas para atender às exigências situacionais e às expectativas dos ouvintes.

A atitude “calma” comportou-se de forma diferente das demais atitudes da dimensão atividade. As pessoas avaliadas positivamente quanto à atitude calma apresentaram um maior grau de soprosidade em sua voz, em comparação aos indivíduos avaliados negativamente nessa atitude. Um dos argumentos centrais para justificar este achado é que a soprosidade pode ser percebida como uma qualidade vocal mais suave e menos agressiva (LIU X, 2011; KREIMAN J., 2011; VAN BORSEL *et al*., 2009). Em contextos sociais e comunicativos, uma voz soprosa pode transmitir calma e serenidade, sugerindo que o falante está relaxado e, por extensão, em um estado emocional estável

(LIU X., 2011). Esta qualidade da voz pode ser particularmente reconfortante para os ouvintes, criando um ambiente comunicativo que é percebido como mais acolhedor e menos ameaçador.

Além disso, a soprosidade pode diminuir a percepção de agressividade ou dominância em uma voz (BORSEL *et al*., 2009). Em contraste com vozes projetadas e em maior intensidade, que podem ser associadas à assertividade ou à agressividade, uma voz soprosa pode ser percebida como mais gentil e acessível em alguns contextos (WAARAMAA T. *et al*., 2022). Isso pode ser particularmente valorizado em contextos em que a calma é uma atitude esperada e avaliada positivamente, como em situações de resolução de conflitos ou na mediação de discussões acaloradas.

Obviamente, é possível que existam influências culturais e contextuais na forma como a soprosidade é percebida (PEARSEL & PAPE, 2023; WAARAMAA T. *et al*., 2022; BRUCKERT L. *et al*., 2010). Em algumas culturas ou subculturas, vozes mais suaves e sussurradas podem ser especialmente valorizadas em certos contextos, como na poesia, na narrativa ou em práticas meditativas (VERDUYCKT I & MORSOMME, 2022; PISANSKI, 2019). Portanto, a valoração positiva da soprosidade em vozes consideradas calmas pode refletir normas culturais específicas sobre o que é considerado uma expressão vocal apropriada e agradável para transmitir calma.

Em suma, a maior soprosidade em vozes avaliadas positivamente quanto à atitude calma pode ser justificada por sua capacidade de transmitir tranquilidade, diminuir a percepção de agressividade, evocar respostas psicoacústicas positivas e alinhar-se com preferências culturais ou contextuais. Essa qualidade vocal, portanto, pode ser um componente chave na comunicação eficaz de calma e serenidade, contribuindo para a formação de julgamentos positivos sobre a atitude do falante, especificamente para falantes do português brasileiro.

Nós não observamos diferenças significativas nas medidas acústicas para vozes avaliadas com valência positiva e negativa na atitude de calma, com exceção do CPPS e H1A3. A justificativa para o achado específico relacionado à atitude calma pode ser aprofundada considerando os aspectos acústicos e perceptivos associados a essas medidas. O CPPS é um parâmetro que reflete a projeção vocal e a qualidade da voz (SAMLAN *et al*., 2013). Menores valores de

CPPS podem indicar uma voz com menor componente harmônico ou com menor projeção vocal (PATEL *et al*, 2018). Considerando que as vozes avaliadas positivamente quanto à calma, apresentaram uma média do CPPS de 13,52±5,07 dB, acima dos valores de corte de 10,90 dB do CPPS para as sentenças do CAPE-V em falantes do português brasileiro (LOPES e ABREU, 2024), os achados do presente estudo podem indicar a presença de vozes menos projetadas associadas à atitude calma. Sendo assim, os falantes avaliados positivamente quanto à atitude calma podem apresentar uma produção vocal mais suave, o que pode ser percebido como mais agradável e relaxante pelos ouvintes.

Por outro lado, o H1A3 é uma medida relacionada ao timbre da voz, especificamente à relação entre as amplitudes dos harmônicos, que pode influenciar a qualidade percebida da voz. Maiores valores de H1A3 sugerem uma maior riqueza harmônica (EKBERG M., 2023), o que pode ser mais eficaz em transmitir a sensação de tranquilidade e conforto, possivelmente devido à associação dessas qualidades acústicas com a presença de o apoio emocional.

Assim, a combinação de menores valores de CPPS com maiores valores de H1A3 em falantes avaliados positivamente quanto à atitude calma pode refletir uma otimização acústica para a transmissão dessa atitude específica. Essas características vocais podem facilitar a comunicação de um estado de serenidade e controle emocional, que são centrais para a expressão de calma. Portanto, a diferença nas medidas acústicas entre avaliações positivas e negativas para a atitude calma pode ser interpretada como um alinhamento mais preciso das propriedades vocais com as qualidades percebidas como desejáveis para a expressão efetiva de calma, evidenciando o papel da voz como um veículo expressivo de nuances emocionais e atitudinais.

Em síntese, os achados deste estudo reforçam a importância das características acústicas da voz na comunicação das atitudes sociais e emocionais. A análise sugere que a voz não apenas transmite informações linguísticas, mas também funciona como um veículo poderoso para a expressão de estados psicológicos e traços de personalidade. Estes resultados contribuem para a compreensão das nuances na percepção vocal e oferecem subsídios para futuras investigações sobre a interação entre as propriedades acústicas da voz

e a percepção social, bem como para o desenvolvimento de intervenções fonoaudiológicas e tecnológicas visando aprimorar a comunicação humana.

A análise das dimensões de avaliação, potência e atividade no contexto das vozes disfônicas fornece insights valiosos sobre como as percepções auditivas e acústicas influenciam o julgamento de atitudes dos ouvintes. Esta análise, imersa no estudo das reações aos diversos aspectos da voz humana, revela a complexidade da comunicação vocal e sua interpretação. A interpretação das vozes, permeada por estas três dimensões, revela a multifacetada natureza da percepção vocal e destaca a importância de considerar a voz como um todo complexo, em que cada característica contribui de forma única para a impressão geral transmitida. Assim, a análise detalhada dessas dimensões não só enriquece nossa compreensão das dinâmicas da comunicação vocal, mas também oferece pistas importantes para o desenvolvimento de intervenções terapêuticas e tecnológicas destinadas a melhorar a comunicação em indivíduos com disfonia.

Destaca-se ainda que vozes avaliadas positivamente em termos de resistência, segurança e independência, apresentaram menor grau de *shimmer* e maior GNE e HNR, indicando menor irregularidade na produção vocal, menos ruído, e um maior componente de periodicidade, respectivamente, o que pode refletir uma percepção de maior controle e capacidade vocal, corroborando com a valência positiva atribuída. Vozes com essas características podem ser interpretadas como mais saudáveis e eficientes, o que pode influenciar positivamente a comunicação e relações sociais do indivíduo.

Os atributos resistência, segurança e independência são percebidos de forma mais positiva em vozes menos alteradas para tensão, rugosidade e soprosidade, reforçando a percepção de um falante mais firme e autoconfiante, características valorizadas em contextos sociais e profissionais (HENTON, 2017)

Na comparação das médias dos valores dos parâmetros acústicos das vozes em relação às valências das atitudes da dimensão avaliação, potência e atividade, os resultados indicaram que vozes com valência positiva, ou seja, avaliadas como mais agradáveis, simpáticas, saudáveis, extrovertidas, poderosas, autoritárias, competentes, resistentes, seguras e independentes, mais altas para CPP e CPPS, ou seja, apresentaram sinal vocal com maior qualidade.

Entre as medidas com diferença estatística significante na dimensão avaliação, destacou-se as relacionadas à frequência fundamental e de perturbação, em que as mais agudas e com maior variação de frequência e intensidade foram consideradas mais negativas. Uma voz com uma frequência fundamental mais estável e com um maior componente de periodicidade pode ser percebida como mais agradável, o que também pode influenciar positivamente a percepção de simpatia e extroversão (KYRIAKOU, PETINOU e PHINIKETTOS 2018; VILLAR, KORN e AZEVEDO, 2016; RUMBACH *et al*., 2015).

Dessa forma, é possível que a investigação e interpretação efetiva dessas medidas acústicas desde a avaliação vocal, relacionando-as com a percepção social pode auxiliar na definição de metas terapêuticas mais específicas e monitoramento dos resultados do tratamento de disfonias .Por isso, sugere-se que as Intervenções terapêuticas incluam estratégias para melhorar o alvo relacionado ao componente de periodicidade da voz, a fim de melhorar a percepção da voz e, consequentemente, sua impressão comunicativa, impactando na qualidade de vida dos pacientes.

6.3 CORRELAÇÃO ENTRE ATITUDES DOS OUVINTES E AS MEDIDAS ACÚSTICAS E DO JPA

A análise da correlação entre as medidas acústicas das vozes dos falantes e a percepção de atitude dos ouvintes é fundamental para compreender como características acústicas específicas, extraídas a partir da voz, influenciam a maneira como os falantes são percebidos socialmente.

Observou-se correlação negativa forte entre o julgamento das atitudes dos ouvintes e as medidas EAV-GG e EAV-R, assim como uma correlação negativa moderada com as medidas EAV-S e EAV-T. Estes resultados reforçam que quanto maior a percepção de disfonia através do grau geral, rugosidade, soprosidade e tensão na voz, mais negativo é o julgamento dos ouvintes em relação em relação às atitudes de agradabilidade, poder, simpatia, força, resistência, extroversão, saúde, autoridade, segurança, competência e independência.

A correlação negativa forte entre EAV-GG e EAV-R com o julgamento de atitudes indica que a qualidade vocal tem um impacto significativo na percepção social. Vozes com maior grau geral de desvio vocal são frequentemente associadas a características negativas, o que pode levar a estigmatização e preconceito do falante (AMIR, 2013). Isso é particularmente relevante para pacientes disfônicos, cujas vozes podem ser julgadas como menos agradáveis ou competentes, afetando suas interações sociais e oportunidades profissionais.

O atributo de "agradabilidade" apresenta uma das correlações mais fortes com a EAV-GG e EAV-R, indicando que uma voz percebida como mais disfônica é menos provável de ser considerada agradável pelos ouvintes. Isso tem implicações importantes na comunicação interpessoal, pois a agradabilidade vocal pode influenciar a primeira impressão e a disposição dos ouvintes para engajar em interações sociais. Além disso, a agradabilidade está associada à sensação de conforto e satisfação ao escutar a voz, podendo influenciar a percepção de amizade e cordialidade que os ouvintes têm em relação ao falante.

Em relação ao "Poder", a correlação também é fortemente negativa, especialmente com a EAV-GG e EAV-R. Isso sugere que a qualidade vocal pode afetar a percepção de autoridade e confiança de um indivíduo, o que é particularmente relevante em contextos profissionais, em que a projeção de poder e confiança pode ser crucial.

A atitude "Simpatia" segue um padrão semelhante, com uma correlação negativa forte com a EAV-GG e EAV-R. Isso implica que a percepção de simpatia pode ser comprometida por uma qualidade vocal disfônica, potencialmente afetando a capacidade de um indivíduo de formar relações sociais positivas.

Os atributos como "Força" e "Resistência" também mostram correlações negativas fortes com a EAV-GG e EAV-R, sugerindo que a robustez percebida da voz pode ser um indicador importante de vitalidade e estabilidade, que são afetadas negativamente pela presença de disfonia.

A "Extroversão" apresenta uma correlação negativa moderada a forte com EAV-GG E EAV-R, o que pode indicar que uma voz disfônica pode ser interpretada como um sinal de introversão ou timidez, mesmo que isso não corresponda à personalidade do falante. A "Saúde" é outra atitude com correlação negativa forte com a EAV-GG e EAV-R, o que é compreensível, pois

a qualidade da voz é frequentemente associada ao bem-estar físico e pode influenciar a percepção de saúde do falante. Por fim, a "Autoridade" também apresenta uma correlação negativa forte, ressaltando como a voz pode impactar a percepção de liderança e controle.

Do ponto de vista clínico, esses achados ressaltam a importância de uma abordagem holística no tratamento da disfonia. Além de focar na melhoria das características acústicas da voz, os profissionais que atuam no diagnóstico e reabilitação das alterações vocais, devem estar cientes das implicações sociais e emocionais da disfonia (BARBOSA *et al*., 2022). O processo de reabilitação vocal que visa reduzir a rugosidade, soprosidade e tensão podem não apenas melhorar a funcionalidade e qualidade vocal, mas também a percepção social e a autoestima do paciente (EVANGELISTA, 2022).

Além disso, esses achados contribuem para a linguística ao fornecer insights sobre como as características vocais afetam a construção social de atitudes e identidades, isto é, fornecem insights sobre como a voz é um veículo de informação social e como as propriedades acústicas da fala podem influenciar o julgamento e as atitudes dos ouvintes. Isso pode contribuir para o desenvolvimento de modelos de comunicação mais inclusivos e compreensivos, que considerem as variações vocais e suas percepções associadas.

Os dados de correlação apresentados mostram uma variedade de correlações entre os parâmetros acústicos e o julgamento de atitudes. Por exemplo, o CPP-CS (Coeficiente de Perturbação de Pitch) e o CPPS-CS (Coeficiente de Perturbação de Pitch Suavizado) apresentam correlações positivas moderadas, indicando que quanto maior a estabilidade da frequência fundamental (F0), mais positivo é o julgamento de atitudes. Isso sugere que uma voz estável é percebida como mais agradável e confiável.

Por outro lado, medidas como Spectral Tilt-CS, F0SD-CS (Desvio Padrão da Frequência Fundamental), F0CV-CS (Coeficiente de Variação da Frequência Fundamental), PSD-CS (Densidade Espectral de Potência), LNPSD-CS (Logaritmo Natural da Densidade Espectral de Potência), e diversas medidas de *jitter* e shimmer apresentaram correlações negativas moderadas com a média do julgamento feito pelos ouvintes. Estes resultados indicam que a presença de irregularidades na frequência e na amplitude da voz, que são características

comuns em vozes disfônicas, pode levar a julgamentos negativos de atitudes como agradabilidade, competência e autoridade.

Ao analisar detalhadamente as correlações entre as medidas acústicas e o julgamento das atitudes dos ouvintes, as medidas CPP-CS e CPPS-CS estão relacionadas à estabilidade da frequência fundamental da voz. A correlação positiva moderada sugere que vozes com menor perturbação de pitch são julgadas mais favoravelmente e isso deve-se à percepção de que uma voz estável é mais agradável e confiável. Já a medida Autocorrelation-CS, relaciona-se com a periodicidade do sinal de voz. Uma correlação positiva indica que vozes com maior periodicidade são percebidas como mais positivas, o que pode estar associado à regularidade do sinal e à qualidade vocal.

As medidas de GNE1000Hz-CS, GNE2000Hz-CS, GNE3000Hz-CS referem-se ao *Glottal to Noise Excitation ratio* em diferentes frequências e estão relacionadas à qualidade do sinal glótico (LOPES L. W., *et al*., 2018). As correlações positivas sugerem que uma voz com menos ruído é percebida como mais saudável e atraente. A medida PA-CS, isto é, prominência de pico (Peak Amplitude) correlaciona-se positivamente, indicando que uma maior regularidade dos picos de amplitude na voz pode contribuir para uma percepção mais positiva.

Ao analisar as medidas espectrais Spectral Tilt-CS, F0SD-CS, F0CV-CS, PSD-CS e LNPSD-CS, que estão relacionadas à inclinação espectral e à variabilidade da frequência fundamental, constatam-se correlações negativas moderadas, isto é, indicam que uma maior irregularidade e dispersão na frequência fundamental são percebidas negativamente, o que pode afetar a percepção de confiança e competência. As diversas medidas de *Jitter* e *Shimmer* que se relacionam com a irregularidade na frequência e amplitude do sinal vocal, respectivamente, apresentam correlações negativas e sugerem que as vozes com maiores valores de tais medidas são julgadas como menos agradáveis e menos saudáveis, o que pode ser devido à percepção de instabilidade vocal.

Por fim, as medidas AVI-CS (Índice de Variabilidade Acústica) apresenta tem uma correlação negativa, indicando que uma maior variabilidade acústica geral é associada a julgamentos menos favoráveis. A relação harmônico-ruído média (HNRmean-CS) quando correlacionada negativamente, sugere que uma menor clareza harmônica pode levar a uma percepção negativa da voz. Já a

frequência fundamental mínima (f0min-CS) apresentou uma correlação fraca, o que pode indicar que vozes com tons mais baixos, podem ser julgadas de maneira ligeiramente mais positiva, embora essa correlação seja menos significativa em comparação com outras medidas.

As correlações observadas revelam que a percepção da voz e o julgamento das atitudes dos ouvintes são multifacetados e influenciados por uma combinação de fatores acústicos. Do ponto de vista da linguística, esses dados fornecem insights sobre como a voz é um veículo poderoso para a comunicação não apenas de informações, mas também de traços sociais e emocionais. Isso pode levar a pesquisas adicionais sobre a relação entre a fala, a identidade social e as atitudes, bem como sobre como esses elementos interagem em diferentes contextos comunicativos.

De modo geral observa-se que as medidas de CCP-CS e CPPS-CS apresentam correlações positivas com a média do julgamento de atitudes realizado pelos ouvintes, sugerindo que os falantes com vozes mais estáveis e menos variáveis em termos da frequência são percebidos como mais agradáveis, poderosos, fortes, resistentes, extrovertidos, saudáveis, autoritários, simpáticos, seguros, competentes e independentes. A medida Spectral Tilt-CS apresenta uma correlação negativa significativa apenas com o atributo calma, o que sugere que um declínio espectral mais acentuado pode ser associado a uma percepção de menos calma.

6.4 MODELOS DE PREDIÇÃO DA VARIABILIDADE NO JULGAMENTO DE ATITUDES DOS OUVINTES A PARTIR DAS MEDIDAS ACÚSTICAS E DO JPA

Esta subseção aborda os modelos de predição da variabilidade no julgamento de atitudes dos ouvintes a partir das medidas acústicas e do JPA, por meio de um modelo de regressão linear. Esse modelo assumiu uma relação linear entre os parâmetros acústicos, que são as variáveis independentes, e as atitudes, variável dependente.

Inicialmente, nós obtivemos um modelo capaz de explicar em 80,0% (R^2 = 0,800) a **variabilidade global** no julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas. Esse modelo incluiu uma medida do JPA (o grau geral do desvio vocal) e duas medidas acústicas (a f_o mínima e o SNL na

faixa de 100-5100Hz) como preditores significativos do julgamento de atitudes dos ouvintes para vozes disfônicas e não disfônicas. Tais achados podem oferecer insights acerca dos elementos cruciais que influenciam a percepção vocal. A capacidade desse modelo de explicar 80% da variabilidade nos julgamentos de atitudes destaca a importância de uma abordagem multidimensional na análise da voz, combinando medidas perceptivo-auditivas e acústicas.

O grau geral de desvio vocal, baseado no JPA realizado pelo clínico, é um indicador compreensivo da qualidade vocal percebida e da aceitabilidade social de uma voz (ALTENBERG EP & FERRAND C. T., 2022; IRANI F. *et al*., 2014; ISETTI D. *et al*., 2014). A percepção de desvio da qualidade vocal é, frequentemente, o primeiro sinal que os ouvintes identificam como indicativo de uma valência negativa. Isso ocorre porque a qualidade vocal pode estar associadas a estados emocionais, saúde e características pessoais (IRANI F. *et al*., 2014). Dessa forma, indivíduos com maior grau de desvio vocal podem ser avaliados negativamente pelos ouvintes. O coeficiente negativo para o grau geral de desvio vocal no modelo reforça essa noção, sugerindo que quanto maior o desvio percebido, mais negativo é o julgamento da atitude do ouvinte em relação à voz do falante.

Por outro lado, a f0 mínima reflete as capacidades de variação tonal e expressividade vocal do falante (O'CONNOR *et al*., 2012; BORKOWSKA B & PAWLOWSKI B., 2011). Falantes com a f0 mais baixa ou com maior variação tonal podem ser percebidos como mais agradáveis ou confiantes, contribuindo para uma avaliação positiva (SKRINDA I. *et al*., 2015; TSANTANI M. S. *et al*., 2017). Isso pode explicar o coeficiente positivo para a f0 mínima no modelo, considerando-se que uma maior gama tonal, ou a capacidade de atingir tons mais baixos durante a fala, pode ser associada a julgamentos mais favoráveis (TSANTANI M. S. *et al*., 2017). Infere-se então que tal característica pode ser interpretada como um indicador de competência comunicativa, ambos atributos valorizados tanto em contextos sociais quanto profissionais.

O nível de ruído espectral na faixa de 100-5100 Hz, é um marcador da presença de ruídos adicionais que podem afetar a qualidade sonora percebida (PAIVA *et al*., 2024; BEN B. B. *et al*., 2018). Um nível mais baixo de ruído espectral indica uma voz com menor grau de desvio vocal e menor componente

de ruído, que é frequentemente associada a uma maior competência vocal e saúde (HODGES S., *et al*., 2010). O coeficiente positivo para o nível de ruído espectral na faixa de 100-5100 Hz no modelo sugere que vozes com menor ruído espectral são julgadas mais favoravelmente. Tal julgamento ocorre, possivelmente, porque o menor componente de ruído facilita a comunicação eficaz e aumenta a agradabilidade da voz.

A combinação dessas três medidas no modelo fornece uma abordagem holística para entender a percepção vocal e o julgamento de atitudes, capturando não apenas a qualidade vocal em si, mas também a expressividade, que são fundamentais para a comunicação humana eficaz. Essa abordagem multidimensional reflete a complexidade da percepção vocal e destaca a importância de considerar tanto aspectos acústicos quanto perceptivo-auditivos na avaliação das vozes. A eficácia do modelo em explicar uma grande parte da variabilidade nos julgamentos de atitudes reforça a ideia de que a percepção vocal é multifacetada e que uma análise abrangente pode oferecer insights valiosos para a compreensão das atitudes dos ouvintes em relação a diferentes qualidades de voz.

A capacidade do modelo de explicar 80% da variabilidade nos julgamentos de atitudes em relação à pode ter implicações na prática clínica, especialmente no tratamento de distúrbios da voz. Primeiramente, o modelo enfatiza a importância de uma avaliação abrangente, que inclua não apenas a análise acústica, mas também a percepção auditiva das vozes disfônicas. Isto é crucial para o diagnóstico preciso e o planejamento terapêutico, visto que as medidas identificadas podem servir como objetivos específicos de intervenção.

Por exemplo, a intervenção pode ser direcionada para a melhoria da qualidade vocal (com a redução do grau geral de desvio vocal e do ruído presente na voz), incluir exercícios com variação da f_0 e a prática deliberada da fala com maior variabilidade de f_0. Sendo assim, o modelo fornece insights para a terapia vocal, permitindo aos clínicos monitorarem o progresso de forma objetiva e ajustar as intervenções conforme necessário.

Do ponto de vista linguístico, o modelo oferece importantes contribuições para a compreensão da percepção vocal e da comunicação humana. Primeiramente, ele reafirma a ideia de que a voz é uma poderosa ferramenta de comunicação social, capaz de transmitir informações não apenas sobre o

conteúdo da mensagem, mas também sobre o estado emocional e físico do falante. O entendimento de como certas características vocais influenciam a percepção do ouvinte pode informar teorias linguísticas sobre pragmática e sociolinguística, particularmente aquelas relacionadas à inferência de traços de personalidade e estados emocionais através da fala.

Além disso, o modelo contribui para a área da psicolinguística, oferecendo insights sobre os processos cognitivos envolvidos na percepção e julgamento vocal. Ao identificar as medidas específicas que os ouvintes inconscientemente avaliam ao formar atitudes em relação às vozes, os pesquisadores podem explorar como essas percepções se encaixam nos modelos existentes de processamento da linguagem e interação social.

Nós obtivemos um modelo que incluiu o grau geral de desvio vocal, a f_0 na fala conectada e o *jitter* como preditores significativos do julgamento de **agradabilidade** em vozes disfônicas e não disfônicas. A significância do grau geral de desvio vocal (-0,041; t = -11,349; p < 0,001) como um preditor negativo para a agradabilidade vocal ressalta a importância da qualidade vocal na percepção do ouvinte. O grau geral do desvio vocal pode ser entendido como uma medida agregada de irregularidades e anormalidades na voz, englobando aspectos como rugosidade, soprosidade e tensão, por exemplo (Maryn & Roy, 2012). A relação inversa entre desvio vocal e agradabilidade sugere que quanto maior a percepção do grau geral de desvio vocal, menor é a probabilidade de ela ser julgada como agradável. Isso está alinhado com a literatura que aponta a preferência por vozes mais claras e estáveis, associadas a saúde vocal e facilidade de comunicação (SIMMONS L. W. *et al*., 2011; FRACCARO P. J., 2010).

O papel da f0 mínima na fala conectada (β = 0,014; t = 2,460; p = 0,019) como um preditor positivo da agradabilidade destaca a complexidade da percepção vocal. A f_0 é um componente crítico da expressão emocional e da identidade vocal (TSANTANI M. S. *et al*., 2017). Uma f_0 mínima mais alta pode ser associada a uma voz mais jovem ou mais feminina, características frequentemente valorizadas socialmente e percebidas como mais agradáveis (ZHENG Y., 2020; FRACCARO P. J., *et al*., 2014.). Esse achado sugere que, no contexto da fala conectada, pequenas variações na f0 mínima podem influenciar

significativamente a percepção de agradabilidade, reforçando a importância de considerar o contexto dinâmico da fala na avaliação vocal.

O *jitter*, uma medida da perturbação de curto termo na f0, foi identificado como um preditor negativo da agradabilidade (β = -0,454; t = -2,142; p = 0,039), o que indica que irregularidades mais acentuadas na f0 estão associadas a uma percepção menos agradável da voz. Este resultado é consistente com a ideia de que a regularidade do sinal vocal é um componente valorizado na voz, associado à qualidade e à saúde vocal. O *jitter* elevado pode ser percebido como uma irregularidade no sinal vocal, o que potencialmente reduz a agradabilidade percebida da voz.

Os resultados desse modelo oferecem importantes insights para a prática clínica na Fonoaudiologia, particularmente no tratamento de distúrbios da voz, onde a melhoria da agradabilidade vocal pode ser um objetivo terapêutico relevante. Além disso, essas descobertas têm implicações para o design de sistemas de síntese vocal e assistentes virtuais, onde a agradabilidade vocal é crucial para a aceitação e eficácia da tecnologia.

Em suma, a análise que identifica o grau geral de desvio vocal, a f_o mínima na fala conectada e o jitter como preditores significativos da agradabilidade vocal fornece uma base sólida para futuras investigações e aplicações práticas. Esses achados realçam a complexidade da percepção vocal e a importância de abordagens multidimensionais para entender e melhorar a qualidade e a agradabilidade da voz humana.

O grau geral de desvio vocal foi o único preditor da atitude de **simpatia** em relação a vozes disfônicas e não disfônicas, sendo capaz de explicar 54,2% da variabilidade no julgamento dessa atitude dos ouvintes. A relação direta entre o grau geral de desvio vocal e a percepção de simpatia sugere que a presença de componente de ruído e irregularidade na voz tem um impacto significativo sobre como as emoções e intenções são interpretadas pelos ouvintes. Isso indica que a qualidade vocal não é apenas uma característica física ou fisiológica, mas também um vetor de comunicação emocional e social.

A seleção do grau geral de desvio vocal como medida no modelo para a atitude de simpatia está intrinsecamente justificada pela natureza humana de buscar conexões emocionais e sociais através da comunicação. Vozes que se desviam menos em relação ao esperado para o gênero, idade e ocupação do

falante, podem ser percebidas como mais agradáveis, confiáveis e, por extensão, mais simpáticas. Esta percepção pode ser ancorada em vieses inconscientes, que associam vozes com menor componente de irregularidade e ruído com características positivas do falante.

Os insights gerados por esse modelo expandem nossa compreensão sobre a importância da voz na percepção social. Eles sugerem que intervenções fonoaudiológicas voltadas para a redução do desvio vocal podem não apenas melhorar a qualidade da voz do ponto de vista técnico, mas também influenciar positivamente a percepção social e emocional de indivíduos com disfonia. Além disso, esses achados podem incentivar a criação de programas de conscientização que visem reduzir o estigma associado a desvios vocais, promovendo uma sociedade mais inclusiva e empática.

A partir desse achado, podemos endossar a importância da disseminação de informações destinadas a minimizar o impacto negativo de vozes disfônicas, promovendo uma maior aceitação e compreensão da diversidade vocal. A educação e a conscientização acerca dos desvios vocais podem informar o público sobre a natureza desses desvios e como eles não refletem necessariamente a competência, inteligência ou caráter de uma pessoa, ajudando a diminuir o estigma e a discriminação. Programas de treinamento que promovem habilidades de escuta ativa e empática podem ajudar indivíduos a focar menos nas características acústicas da voz e mais no conteúdo da mensagem, enquanto organizações podem revisar suas políticas de inclusão para garantir que não haja viés contra pessoas com disfonia.

Além disso, destacar a influência do desvio vocal na percepção de simpatia sem determinar a capacidade ou valor de uma pessoa estimula uma maior aceitação da diversidade vocal. Isso contribui para a criação de ambientes mais inclusivos onde todos se sentem valorizados independentemente das características de sua voz. A pesquisa também pode inspirar o desenvolvimento de novas tecnologias assistivas para pessoas com disfonia, tornando a comunicação mais acessível e diminuindo barreiras na interação social e profissional. Ademais, o achado incentiva mais pesquisas sobre a percepção social das diferenças vocais e fomenta o debate sobre como tornar a sociedade mais inclusiva para pessoas com disfonia ou outras condições que afetam a voz.

Por fim, entender a relação entre a percepção de simpatia e o grau de desvio vocal permite o desenvolvimento de estratégias para promover a inclusão e valorizar a diversidade, reduzindo o impacto das vozes disfônicas. Isso destaca a necessidade de maior conscientização e educação sobre as disfonias, além de reforçar a importância de valorizar a diversidade em todas as suas formas. Implementando políticas inclusivas, focando em treinamentos de empatia e incentivando pesquisa, a sociedade pode avançar no sentido de celebrar, e não estigmatizar, a diversidade vocal.

O grau geral de desvio vocal (β = -0,053; t = -11,153; p < 0,001) obtido no JPA e a medida acústica da f_0 mínima na fala conectada (β = 0,020; t = 2,460; p = 0,014) foram preditores do julgamento dos ouvintes quanto à atitude **saúde** para vozes disfônicas e não disfônicas. O modelo com essas duas medidas foi capaz de explicar em 79,2% (R^2 =0,792) a variabilidade no julgamento de uma pessoa como saudável ou doente. Estes resultados lançam luz sobre como as características vocais são interpretadas pelos ouvintes e como essas percepções podem afetar o julgamento social em relação à saúde do falante. Aprofundando-se na análise desses preditores, buscamos entender melhor suas implicações e o que eles revelam sobre a natureza da comunicação humana e da percepção vocal.

O grau geral de desvio vocal, que se mostrou inversamente relacionado à percepção de saúde (β = -0,053; t = -11,153; p < 0,001), indica que quanto maior o desvio percebido na voz, menor a probabilidade de o ouvinte julgar o falante como saudável. Isso sugere que os ouvintes associam a qualidade vocal à saúde física. Vozes mais desviadas são percebidas como indicativas de doença ou mal-estar. Essa percepção pode estar fundamentada em um mecanismo evolutivo de avaliação rápida da condição de potenciais parceiros ou adversários, em que a qualidade da voz serve como um indicador acessível da saúde geral de um indivíduo (HUGHES *et al*., 2021; XU, YI *et al*., 2013; PUTS, 2011).

Por outro lado, a f0 mínima na fala conectada (β = 0,020; t = 2,460; p = 0,014) apresenta uma relação positiva com a percepção de saúde. Esse achado sugere que vozes que se situam em uma faixa mais baixa da tessitura vocal (vozes mais graves) são, em geral, associadas à saúde ou à vitalidade. Isso pode estar ligado a percepções culturais ou biológicas que associam vozes mais

graves a maior força física, estabilidade emocional ou maturidade (JONES B. C. *et al*., 2010; HODGES S., 2010) atributos geralmente considerados positivos.

A integração entre essas duas medidas em um modelo preditivo capaz de explicar 79,2% da variabilidade no julgamento da atitude saúde evidencia o potencial combinado desses fatores na formação de impressões sobre o estado de bem-estar de uma pessoa com base em sua voz. Esse achado pode ter implicações para o campo da saúde vocal, sugerindo que intervenções terapêuticas que visem a minimização dos desvios vocais e a manutenção ou recuperação da capacidade de produzir a voz em tons mais graves podem melhorar a percepção social da saúde do indivíduo. Além disso, para os usuários de tecnologia assistiva, a integração dessas descobertas pode significar vozes sintéticas que soam mais naturais e saudáveis, melhorando a interação social e a qualidade de vida.

Do ponto de vista linguístico, este modelo reforça a compreensão de que a voz é um elemento crucial na construção social de identidades, incluindo a percepção de saúde. Tal achado destaca a importância de considerar as dimensões sociais e afetivas da comunicação vocal, além das suas propriedades físicas e acústicas, em estudos linguísticos.

Finalmente, os resultados desse estudo destacam a importância de educar o público sobre as complexidades da saúde vocal e desafiam estereótipos de que desvios vocais estão diretamente relacionados à saúde debilitada. Promover uma compreensão mais matizada da relação entre voz e saúde pode contribuir para reduzir o estigma enfrentado por indivíduos com disfonia, incentivando uma sociedade mais inclusiva onde as diferenças vocais são valorizadas e não apenas toleradas.

O grau geral de desvio vocal ($\beta = -0,026$; $t = -6,276$; $p < 0,001$) obtido no julgamento perceptivo-auditivo da qualidade vocal, o declínio espectral ($\beta = 0,078$; $t = 3,102$; $p = 0,004$) e a f0 mínima na fala conectada ($\beta = 0,015$; $t = 2,196$; $p = 0,034$) foram preditores do julgamento dos ouvintes quanto à atitude **extroversão** para vozes disfônicas e não disfônicas. O modelo com essas três medidas foi capaz de explicar em 59,9% ($R^2 = 0,599$) a variabilidade no julgamento de atitude da extroversão relacionada a vozes disfônicas e não disfônicas.

Primeiramente, o grau geral de desvio vocal, que apresentou uma correlação negativa (β = -0,026) com a percepção de extroversão, sugere que quanto maior o desvio da qualidade vocal percebido, menor a probabilidade de o falante ser julgado como extrovertido. Dessa forma, esse achado pode indicar que vozes menos desviadas tendem a ser frequentemente associadas a uma comunicação mais eficaz, características que podem ser percebidas como indicativas de uma personalidade mais aberta e sociável (KISSEL *et al*., 2022; AMIR, 2013; ALLARD & WILLIANS, 2008). A extroversão, sendo uma atitude associada à sociabilidade e assertividade, poderia ser menos percebida em indivíduos cujas vozes apresentam desvios significativos. Provavelmente, essa percepção ocorre devido a uma suposição inconsciente de que tais desvios vocais poderiam impedir a comunicação efetiva, um elemento chave na interação social.

Em segundo lugar, o declínio espectral, que teve uma correlação positiva (β = 0,078) com a percepção de extroversão. Sendo assim, um declínio espectral acentuado pode ser interpretado como um indicativo de maior riqueza harmônica ou uma voz mais "cheia", potencialmente transmitindo maior assertividade do falante (TAYLOR *et al*., 2020; FEINBERG *et al*., 2005).

Por fim, a f_0 mínima na fala conectada mostrou uma correlação positiva (β = 0,015) com a percepção de extroversão. Tal achado pode sugerir que vozes mais graves podem ser associadas a características como confiança e assertividade, atitudes frequentemente atribuídas a pessoas extrovertidas. A capacidade de variar a f0, especialmente atingindo notas mais baixas na gama tonal, pode conferir ao discurso uma qualidade mais dinâmica e expressiva, elementos que reforçam a percepção de extroversão.

O modelo estatístico, com um R^2 de 0,599, indica que essas três medidas explicam aproximadamente 59,9% da variabilidade no julgamento de atitude de extroversão relacionada a vozes disfônicas e não disfônicas. Dessa forma, ressalta-se a relevância dessas características vocais na percepção da extroversão do falante, ao mesmo tempo em que contribui para enfatizar a complexidade da comunicação humana e a importância da voz como uma ferramenta poderosa na transmissão de características pessoais. Contudo, é importante notar que, apesar de significativa, uma parcela da variabilidade ainda permanece inexplicada, o que sugere a influência de outros fatores não

capturados por essas medidas na percepção da extroversão. Dessa forma, pesquisas futuras podem investigar outras características vocais ou contextuais que também afetem a percepção da extroversão do falante.

A investigação sobre a relação entre características vocais e a percepção de atitudes, como a extroversão, tem implicações significativas tanto para o campo clínico, especialmente na fonoaudiologia, quanto para os estudos em linguística. No âmbito clínico, compreender como diferentes características acústicas da voz afetam a percepção de extroversão pode informar estratégias de intervenção para indivíduos com vozes disfônicas. A presença de disfonia pode afetar negativamente a qualidade de vida do indivíduo, influenciando a autoestima, a comunicação social e, como indicado pelos resultados deste estudo, a percepção da personalidade pelo ouvinte.

Os fonoaudiólogos podem usar essas informações para desenvolver técnicas direcionadas que não apenas melhorem a qualidade vocal, mas também a forma como a voz do paciente é percebida em contextos sociais. Por exemplo, intervenções que visem a minimizar o grau de desvio vocal e otimizar o declínio espectral e a frequência fundamental podem ajudar a melhorar a percepção de extroversão, o que pode ser particularmente benéfico para indivíduos cujas profissões dependem fortemente da comunicação verbal e da expressão de certas características de personalidade.

Do ponto de vista linguístico, o modelo oferece insights valiosos tanto no contexto da fonética quanto da sociolinguística, explorando a intersecção entre a forma física da fala e a percepção social. Este estudo contribui para uma compreensão mais profunda de como elementos paralinguísticos e prosódicos da fala podem influenciar julgamentos de personalidade. Além disso, as descobertas reforçam a importância de considerar a voz como um elemento crucial na identidade pessoal e na expressão social.

Ao analisar os modelos que explicam a percepção de agradabilidade, simpatia, saúde e extroversão, todas integrantes da dimensão "avaliação", nós identificamos dois preditores comuns: o grau geral de desvio vocal e a f_0 mínima na fala conectada. Especificamente, a dimensão avaliação engloba o julgamento de valor que os ouvintes fazem sobre um falante, classificando-o como positivo ou negativo com base nas informações disponíveis (nesta pesquisa, a voz dos falantes).

A presença desses dois preditores comuns nos modelos para as atitudes da dimensão avaliação ressalta a importância da qualidade vocal e da f0 na comunicação humana. Eles indicam que, independentemente do conteúdo verbal, as características da voz têm um impacto significativo na forma como os falantes são percebidos pelos ouvintes

A análise estatística e modelagem preditiva aplicadas ao estudo das percepções da atitude **poder** em vozes disfônicas e não disfônicas revelaram uma complexa interação entre características acústicas específicas e a interpretação de poder por parte dos ouvintes. O modelo estatístico selecionou três variáveis significativas: o grau geral de desvio vocal, o nível de ruído espectral entre 100 e 5100 Hz, a f0 mínima e o primeiro quartil da f_0 na fala conectada, cada uma contribuindo de maneira única para a percepção de poder.

O grau geral de desvio vocal ($\beta = -0{,}037$; $t = -9{,}823$; $p < 0{,}001$) indica uma relação inversa significativa entre a percepção de desvios na qualidade vocal e a percepção de poder. Dessa forma, indivíduos cujas vozes são percebidas como menos desviadas são avaliados mais positivamente quanto à atitude poder. Provavelmente, vozes não disfônicas podem ser interpretadas como sinais de saúde e capacidade de controlar aspectos da comunicação verbal, transmitindo confiança e autoridade (ALLARD & WILLIANS, 2008).

O nível de ruído espectral entre 100-5100 Hz ($\beta = 0{,}067$; $t = 3{,}920$; $p = 0{,}001$) apresentou uma correlação positiva com a percepção de poder. Inicialmente, esse achado pode parecer contra-intuitivo, mas o ruído espectral pode ser associado a características vocais como a rouquidão, que em certos contextos, pode ser interpretada como um sinal de assertividade ou até mesmo de agressividade, contribuindo para uma maior percepção de poder (EVANGELISTA *et al*., 2022). Isso ressalta a importância do contexto e da interpretação cultural nas avaliações de características vocais.

A frequência fundamental mínima na fala conectada ($\beta = 0{,}017$; $t = 2{,}915$; $p = 0{,}006$) e o primeiro quartil da frequência fundamental ($\beta = -0{,}005$; $t = -2{,}152$; $p = 0{,}038$) são particularmente interessantes, pois indicam que tanto as frequências mais baixas quanto um menor alcance na variação da frequência fundamental estão associados à percepção de poder. Vozes mais graves são tradicionalmente vistas como mais autoritativas e dominantes, enquanto um alcance mais restrito de variação pode sugerir controle e estabilidade,

qualidades associadas ao poder (EVANGELISTA *et al*., 2022; ZHENG YI, *et al*., 2021).

O modelo foi capaz de explicar 81,7% da variabilidade no julgamento de atitudes de poder, o que fornece um quadro compreensivo de como diferentes aspectos da voz contribuem para a construção social do poder. A relação entre essas variáveis e a percepção de poder em vozes disfônicas e não disfônicas destaca a complexidade da voz humana como um instrumento de comunicação não-verbal. A voz não transmite apenas conteúdo explícito, mas carrega uma riqueza de informações sociais e emocionais, influenciando percepções de autoridade, confiabilidade e liderança. Este estudo reforça a ideia de que a percepção de poder é multifacetada, influenciada por uma interação de informações acústicas e perceptuais, que vão além do conteúdo verbal, moldando as dinâmicas sociais e as relações de poder.

No contexto clínico, especialmente na terapia vocal, a compreensão das implicações das características vocais na percepção de poder é fundamental. Para indivíduos que procuram melhorar a qualidade vocal para fins profissionais ou pessoais, o conhecimento gerado por este modelo pode orientar intervenções específicas. Por exemplo, tratamentos focados em reduzir o grau de desvio vocal, otimizar o nível de ruído espectral, e ajustar a f0 podem ser desenvolvidos com o objetivo de não apenas melhorar a qualidade vocal, mas também de influenciar a forma como a voz é percebida em termos de autoridade e confiança. Além disso, para pacientes disfônicos, entender as percepções sociais associadas a diferentes características vocais pode ajudar a mitigar possíveis estigmas ou julgamentos negativos, fornecendo uma base para estratégias de comunicação mais eficazes que promovam uma imagem positiva e assertiva.

Do ponto de vista linguístico, o modelo reforça a importância de se estudar a voz como um fenômeno multifacetado, que transmite não apenas conteúdo semântico, mas também informações paralinguísticas cruciais para a construção de identidades sociais e relações de poder. A pesquisa destaca como características físicas da fala podem ser interpretadas socialmente, contribuindo para uma compreensão mais profunda da sociolinguística e da pragmática.

Este estudo também abre novas perguntas de pesquisa, como a variação dessas percepções em diferentes culturas ou contextos sociais, e como outros fatores, como o sexo, a idade ou o contexto da fala, podem interagir com

as características vocais na construção da autoridade e do poder. Além disso, fornece um modelo quantitativo que pode ser usado como ponto de partida para estudos futuros, visando entender melhor a complexidade das interações entre linguagem, sociedade e cognição.

A atitude de **força**, diferentemente do conceito de poder, relaciona-se mais diretamente com a percepção de determinação, resiliência e capacidade de superação expressa através da voz (HODGES S., 2010). A análise dos dados obtidos através do modelo estatístico revela quatro variáveis explicam a variabilidade na percepção da atitude força por parte dos ouvintes: o grau geral de desvio vocal, o nível de ruído espectral entre 100 e 5100 Hz, a f_o mínima na fala conectada e o primeiro quartil da f_o na fala conectada.

O coeficiente negativo relacionado ao grau geral de desvio vocal sugere que quanto maior o desvio vocal percebido, menor a percepção da atitude de força a partir da voz do falante. Sendo assim, em contextos sociais e profissionais, vozes menos desviadas podem ser interpretadas como um indicativo de saúde física e emocional, elementos que contribuem para a ideia de força.

Similarmente, o nível de ruído espectral apresentou uma relação inversa com a percepção de força, evidenciando que indivíduos com vozes com menos ruído espectral podem ser percebidas como mais fortes. O ruído espectral pode reduzir a eficácia com que a mensagem é transmitida e, por consequência, a percepção de força relacionada ao falante.

A relação positiva entre a f_o mínima na fala conectada e o julgamento da atitude força pode indicar que o alcance de notas mais baixas pode ser percebido como manifestação de força. Tal achado pode ser atribuído à associação cultural entre vozes mais graves e características como estabilidade e solidez, que são frequentemente interpretadas como sinais de força. Nesse sentido, a relação negativa entre o primeiro quartil da f_o na fala conectada pode sugerir que uma menor variação nas notas mais baixas contribui para a percepção de força.

A relação entre essas características vocais específicas e a percepção de força pelos ouvintes reflete um entendimento complexo sobre como interpretamos sinais vocais e os associamos a atitudes humanas. A força, sendo uma qualidade desejada socialmente, é percebida através de características vocais que sugerem clareza, estabilidade e resiliência. Sendo assim, esses

achados revelam não apenas como percebemos a força na voz, mas também como valorizamos certas qualidades vocais que interpretamos como sinais de força em um falante.

Além disso, ao identificar as medidas acústicas e perceptivo-auditivas que mais influenciam a percepção de força, o modelo oferece aos fonoaudiólogos ferramentas específicas para o desenvolvimento de intervenções terapêuticas direcionadas. Essas intervenções podem ser projetadas para ajustar ou enfatizar determinados aspectos da voz, de modo a facilitar a comunicação da atitude de força, essencial em diversos contextos sociais e profissionais.

No âmbito dos estudos linguísticos, a modelagem da relação entre características vocais e a percepção de atitudes sociais como a força contribui significativamente para os estudos da pragmática e sociolinguística. Este modelo fornece evidências empíricas de como a voz atua como um veículo de comunicação não apenas de conteúdo semântico, mas também de nuances sociais e emocionais. Isso reforça a ideia de que a linguagem é multifacetada e que os sinais vocais desempenham um papel crucial na construção e interpretação de identidades e relações sociais.

Nós encontramos quatro medidas capazes de explicar a variabilidade no julgamento da atitude **autoridade** para vozes disfônicas e não disfônicas, a saber: o grau geral de desvio vocal, o nível de ruído espectral entre 100-8000 Hz, a f0 mínima e o primeiro quartil da f0 na fala conectada.

O grau geral de desvio vocal, com um coeficiente β de -0,023, sugere que, quanto menor o desvio vocal percebido, maior a percepção de autoridade. Sendo assim, indivíduos com vozes menos desviadas podem ser percebidos com uma maior autoridade. Esta descoberta ressalta a preferência por uma voz menor componente de perturbação e ruído, o que é frequentemente interpretado como um sinal de confiança e competência, atributos estreitamente ligados à autoridade (PETINOU e PHINIKETTOS, 2018).

Por outro lado, a análise revelou que o nível de ruído espectral na faixa de 100–8000 Hz exibiu um coeficiente β positivo de 0,058, o que implica que incrementos neste parâmetro acústico estão correlacionados com uma ampliação na percepção de autoridade. Tal observação pode inicialmente parecer paradoxal. No entanto, ela sugere que um determinado nível de

irregularidade na qualidade vocal pode ser interpretado como indicativo de assertividade ou determinação, atributos comumente associados ao perfil de uma figura autoritária (EVANGELISTA, 2022).

No modelo desenvolvido, nós encontramos uma correlação positiva entre a f_0 mínima e a percepção de autoridade, assim como uma correlação negativa entre o primeiro quartil da f_0 e a percepção de autoridade. A f_0 mínima que representa o limite inferior da variação tonal de um falante e parece estar associada a uma maior percepção de autoridade. Isso pode ser explicado pelo fato de que vozes mais graves, que naturalmente teriam uma f_0 mínima mais baixa, são frequentemente associadas a qualidades como força, confiança e competência. Estudos anteriores (ANDERSON *et al.,* 2014, SKRINDA *et al.,* 2014, MCALEER, TODOROV; BELIN, 2014) corroboram essa associação, mostrando que vozes mais graves podem ser percebidas como mais dominantes e autoritárias. Portanto, a correlação positiva encontrada sugere que, mesmo nos pontos mais baixos de sua variação tonal, falantes com uma f_0 mínima mais baixa podem evocar uma maior percepção de autoridade.

Por outro lado, a descoberta de uma correlação negativa entre o primeiro quartil da f_0 e a percepção de autoridade pode ser entendida por meio da distribuição e da média da f_0 em que a voz do falante está durante a fala conectada. O primeiro quartil da f_0 indica o limite abaixo do qual apenas 25% das frequências da fala se encontram, funcionando como um indicador do "piso" habitual da frequência vocal. Uma correlação negativa sugere que, quanto mais alta a frequência que compõe esse "piso", menor a percepção de autoridade. Isso está alinhado com a ideia de que vozes com mais agudas podem ser percebidas como menos assertivas ou imponentes, influenciando a percepção de autoridade de maneira inversa à f_0 mínima.

É importante notar que a percepção de autoridade não é determinada exclusivamente por esses elementos acústicos, mas é influenciada por uma série de fatores, incluindo contextos sociais e culturais, além de características individuais dos ouvintes (ENGLERT, 2016; ZHENG YI, *et al*., 2020; RAYMOND M., *et al*., 2019; SUIRE A., *et al*., 2018; SHIRAZI T. N., PUTS D. A., ESCASA M. J., 2018; KYRIAKOU, PETINOU e PHINIKETTOS 2018; VILLAR, KORN e AZEVEDO, 2016; RUMBACH *et al*., 2015). No entanto, a correlação entre essas medidas específicas e a percepção de autoridade sublinha a importância da voz

como ferramenta de comunicação social e como veículo de expressão de poder e status.

No contexto clínico ou de treinamento para desenvolvimento da comunicação profissional, podem ser desenvolvidas intervenções personalizadas que visam não apenas a saúde vocal, mas também a otimização da comunicação segundo as necessidades individuais dos pacientes. Isso é especialmente relevante para indivíduos cujas profissões dependem fortemente da projeção de autoridade e confiança por meio da voz, como líderes, educadores e profissionais da área jurídica.

Além disso, ao elucidar as bases acústicas da percepção de autoridade, o modelo abre caminho para investigações interculturais sobre como essas percepções são moldadas por fatores culturais e linguísticos. Isso pode levar a um entendimento mais aprofundado dos universais e variações culturais na comunicação de poder e status social através da voz. Finalmente, a modelagem de atitudes em relação à voz disfônica com base na valência positiva ou negativa pode fornecer insights sobre o estigma social associado a certos padrões vocais, contribuindo para uma maior compreensão de como as atitudes em relação à voz impactam a interação social e a comunicação eficaz.

A modelagem preditiva realizada neste estudo selecionou duas medidas capazes de explicar em 70,2% a variabilidade no julgamento da atitude **competência** para vozes disfônicas e não disfônicas, incluindo o grau geral de desvio vocal e a f_0 mínima na fala conectada.

Os achados quanto ao grau geral de desvio vocal, demonstraram que, quanto maior o desvio vocal percebido, menor é a percepção de competência atribuída ao falante, a partir da sua voz. Este resultado sugere que vozes percebidas como desviadas ou disfônicas podem levar a avaliações negativas em termos de confiança ou competência (EVANGELISTA *et al*., 2022, AMIR e LEVINE-YUNDOF, 2013). O desvio vocal pode ser associado, consciente ou inconscientemente, a uma menor saúde vocal ou a uma dificuldade de controle vocal, o que, por sua vez, pode ser interpretado como uma falta de habilidade ou confiabilidade.

Por outro lado, indivíduos cujas vozes apresentam uma f_0 mínima mais baixa são percebidos como mais competentes. Tal achado pode estar relacionado a percepções socioculturais associadas a vozes mais graves, que

frequentemente são interpretadas como mais autoritativas ou confiáveis (ZHENG YI, *et al*., 2020; RAYMOND M., *et al*., 2019; SUIRE A., *et al*., 2018; KYRIAKOU, PETINOU e PHINIKETTOS, 2018; VILLAR, KORN e AZEVEDO, 2016). Este achado reforça a importância da f_0 na construção social da autoridade e competência.

A relação entre o desvio vocal, a f_0 mínima e a percepção de competência pode ser entendida considerando-se a importância da congruência entre a expectativa do ouvinte e as características vocais do falante. Vozes que se desviam significativamente do esperado podem violar as expectativas do ouvinte e, consequentemente, prejudicar a percepção de competência. Em contraste, uma f_0 mínima mais baixa, dentro de um contexto de fala conectada, pode alinhar-se melhor com as expectativas de vozes competentes e confiáveis, especialmente em contextos profissionais ou formais.

O alto valor de R^2 (70,2%) reforça a informação de que as características acústicas da voz têm um impacto significativo na forma como a competência é julgada, destacando a relevância de considerar esses aspectos em avaliações clínicas e intervenções focadas na comunicação vocal.

Do ponto de vista clínico, os achados deste estudo têm implicações diretas para a avaliação e o tratamento de indivíduos com disfonia ou outros distúrbios de voz. A identificação de medidas acústicas específicas que influenciam a percepção de competência pode orientar os fonoaudiólogos na elaboração de objetivos terapêuticos mais precisos e na seleção de intervenções que visem não apenas a melhoria da qualidade vocal, mas também a otimização da comunicação social e profissional dos pacientes. Por exemplo, intervenções focadas em reduzir o desvio vocal e em ajustar a f_0 para níveis que sejam percebidos como mais competentes podem ter um impacto significativo na autoestima e na eficácia comunicativa dos indivíduos.

Além disso, esses resultados reforçam a necessidade de uma abordagem holística na avaliação vocal, que considere não apenas os aspectos fisiológicos da produção vocal, mas também as consequências sociais das características vocais. Isso implica uma maior integração entre a fonoaudiologia e áreas como a psicologia e a sociolinguística, promovendo uma compreensão mais ampla dos efeitos da voz na vida dos indivíduos.

No âmbito dos estudos em linguística, os resultados deste estudo destacam a importância da prosódia e de outras características acústicas na construção social de significados e identidades. A capacidade de predizer a percepção de competência com base em medidas acústicas específicas sublinha o papel da voz não apenas como um veículo para a linguagem, mas também como uma ferramenta poderosa na negociação de status, poder e identidade social. Isso reforça a relevância dos estudos prosódicos e da sociolinguística na compreensão das dinâmicas de interação humana e na identificação de como variações sutis na fala podem ter implicações significativas na percepção social.

Ademais, a aplicação de modelos preditivos e análises estatísticas avançadas em estudos linguísticos representa um avanço metodológico importante, permitindo uma análise mais precisa e detalhada das relações entre a linguagem e a sociedade. Essa abordagem quantitativa complementa os métodos qualitativos tradicionalmente empregados na linguística, abrindo novas possibilidades de pesquisa e contribuindo para uma compreensão mais profunda dos fenômenos linguísticos.

Os modelos preditivos para as atitudes que fazem parte da dimensão potência (poder, força, autoridade e competência) apresentam características em comum que são fundamentais para entender a percepção dessas atitudes em vozes disfônicas e não disfônicas. Todos os modelos incluíram o grau geral de desvio vocal como um preditor significativo, com coeficientes negativos, indicando que quanto maior o desvio vocal percebido, menor a percepção de poder, força, autoridade e competência. Isso sugere uma preferência por vozes que são percebidas como mais claras ou menos desviadas, possivelmente associadas a uma maior confiabilidade e estabilidade vocal, qualidades valorizadas nessas quatro atitudes.

A f_o mínima foi um preditor positivo significativo em todos os modelos, implicando que vozes com capacidade de alcançar tons mais baixos são associadas a uma maior percepção de potência. Isso está alinhado com a literatura que associa frequências fundamentais mais baixas com qualidades como dominância e autoridade (RAYMOND M., *et al*., 2019; SUIRE A., *et al*., 2018; SHIRAZI T. N., PUTS D. A., ESCASA M. J., 2018; BABEL M., MCGUIR G., KIN J., 2014; XU, *et al*., 2013).

A preferência por vozes com menor desvio vocal e a capacidade de produzir tons mais baixos pode refletir uma busca por sinais de confiança, estabilidade e seriedade. Em contextos sociais, essas qualidades vocais podem ser interpretadas como sinais de competência, confiabilidade e capacidade de liderança. A associação de f_0 mais baixas com a percepção de dominância e autoridade é bem documentada. Essas características podem ser interpretadas como indicativos de força física ou assertividade, qualidades valorizadas nas atitudes de poder e autoridade. A habilidade de controlar a variação da f_0, especialmente evitando variações extremas no primeiro quartil, pode ser vista como uma forma de autocontrole ou modulação intencional da expressão vocal para adequar-se a contextos específicos, reforçando a percepção de competência e autoridade.

Quando ao julgamento de **resistência**, quatro medidas foram capazes de explicar em 80,1% a variabilidade no julgamento dessa atitude para vozes disfônicas e não disfônicas: o grau geral de desvio vocal, o nível de ruido espectral entre 100-5100 Hz, a f_0 mínima e o primeiro quartil da f_0 na fala conectada. O conceito de resistência, neste contexto, pode ser entendido como a qualidade ou característica da voz que transmite robustez, firmeza ou a capacidade de um falante de sustentar a fala sob condições adversas, o que é especialmente relevante na avaliação de vozes disfônicas.

Nós encontramos que indivíduos com vozes mais desviadas tendem a ser percebidos pelos ouvintes como menos resistentes. Provavelmente, o desvio da qualidade vocal pode ser, inconscientemente, associado à fraqueza ou vulnerabilidade.

Observou-se que um aumento no nível de ruído espectral na faixa de frequência de 100-5100 Hz está associado a uma maior percepção de resistência. Dessa forma, a presença de certo grau de ruído pode ser percebida como uma característica de pessoas "rústicas" ou "fortes", em contraste com vozes com reduzido componente de ruído, que podem parecer mais frágeis para o ouvinte (MCGUIRE G., KIN J., 2014).

Vozes com f_0 mínimas mais altas na fala conectada tenderam a ser percebidas como mais resistentes no presente estudo. Tal achado pode refletir a associação entre uma voz capaz de alcançar tons mais altos, mesmo que mínimos, e a capacidade de manter uma projeção vocal forte, sugerindo

robustez. Por sua vez, um aumento no primeiro quartil da f_o está negativamente relacionado com a percepção de resistência. Dessa forma, vozes cujas f_o está predominantemente mais baixa (indicadas por valores mais altos no primeiro quartil) são percebidas como menos resistentes, possivelmente devido à associação de tonalidades mais graves com um maior esforço vocal ou menor projeção da voz.

Os resultados obtidos em relação à f_o mínima e ao primeiro quartil da f_o na fala conectada oferecem insights valiosos sobre como nuances acústicas específicas da voz influenciam a percepção de resistência pelos ouvintes. A análise conjunta desses resultados revela uma relação complexa entre a altura tonal da voz e a percepção de resistência.

O coeficiente positivo associado à f_o mínima indica que, quanto mais alta for a f_o mínima de um indivíduo, mais resistente ele é percebido. Dessa forma, indivíduos capazes de alcançar notas mais altas em sua extensão mínima são associadas a uma maior energia e capacidade de projeção, o que pode ser interpretado como um sinal de resistência. A capacidade de elevar a f_o, mesmo que no limite inferior da extensão vocal, pode refletir uma maior flexibilidade e controle das pregas vocais, qualidades que são percebidas como indicativas da percepção de resistência.

Por outro lado, o coeficiente negativo para o primeiro quartil da f_o sugere que um valor mais alto neste parâmetro, que indica uma tendência geral para frequências mais baixas, está associado a uma menor percepção de resistência. Isso pode ser interpretado à luz de como tonalidades mais graves são percebidas. Indivíduos cujas vozes apresentam uma gama predominante de frequências mais baixas podem ser percebidos como menos capazes de projeção, talvez devido a associações com um maior esforço vocal. Tal fato, poderia comprometer a percepção de resistência.

A interpretação conjunta desses resultados enfatiza a importância da variação e flexibilidade da frequência fundamental na percepção de resistência vocal. Enquanto a habilidade de alcançar notas mais altas, mesmo que mínimas, é valorizada e associada o julgamento da atitude de resistência, uma tendência geral para manter frequências mais baixas (indicada por um valor mais alto no primeiro quartil) pode ser interpretada como uma limitação na projeção vocal.

Estes achados destacam um aspecto importante da percepção vocal: a resistência é percebida relacionada à capacidade de modulação e ajuste dinâmico da altura tonal. A capacidade de modular a voz, alcançando tanto notas altas quanto mantendo um timbre rico e pleno, parece ser um elemento chave na construção da percepção de uma pessoa como resistente.

Essa análise aponta para a complexidade da percepção humana da voz e como características específicas da frequência fundamental podem influenciar a imagem que um falante projeta. No contexto clínico e terapêutico, esses insights podem orientar intervenções focadas não apenas na recuperação dos distúrbios da voz, mas também na otimização da percepção de resistência e autoridade vocal, ajustando-se às necessidades comunicativas e profissionais dos indivíduos.

O ruído espectral entre 100-5100 Hz e a f_0 mínima na fala conectada foram preditores do julgamento dos ouvintes em relação à atitude de **segurança** em vozes disfônicas e não disfônicas. A capacidade do modelo de explicar 68,5% da variabilidade no julgamento de segurança enfatiza a relevância dessas medidas acústicas na formação de impressões atitude segurança relacionada aos falantes.

O nível de ruído espectral refere-se à quantidade de energia acústica distribuída ao longo de um espectro de frequência que não é harmonicamente relacionada à f_0 da voz. Um coeficiente positivo ($\beta = 0,044$) para esta medida sugere que um maior nível de ruído espectral entre 100-5100Hz está associado a uma percepção aumentada de segurança do falante. Mais uma vez, nós observamos que, no contexto da comunicação humana, um certo componente de ruído espectral pode contribuir para a percepção de uma voz mais expressiva ou emocionalmente envolvente. Consequentemente, essa característica acústica pode ser percebida como um indicativo de autenticidade e confiança, contribuindo para a sensação de segurança que o ouvinte recebe do falante. Além disso, deve-se considerar que há uma variabilidade nos parâmetros vocais vista como normal, ou seja, prequenas modificações no padrão de produção vocal tem relação com a naturalidade da voz, e por isso estas são impressas nas vozes sintetizadas, por exemplo (ENGLERT, 2016).

A f_0 mínima, com um coeficiente positivo ($\beta = 0,015$), indica que, indivíduos cujas vozes conseguem alcançar notas mais baixas na fala

conectada, tendem a ser percebidos como mais seguros. Tal achado alinha-se com a percepção comum de que vozes mais graves são associadas a características de liderança, confiança e estabilidade (EVANGELISTA *et al*, 2020). A capacidade de um falante de modular sua voz para notas mais baixas pode ser interpretada como um sinal de controle vocal e emocional, elementos que são intrinsicamente ligados à ideia de segurança e confiança.

A relação entre o nível de ruído espectral, a f_o mínima e a percepção de segurança pode ser entendida através da maneira como essas características acústicas influenciam a interpretação do ouvinte sobre a competência, a confiança e a capacidade emocional do falante. Enquanto o nível de ruído espectral contribui para a percepção de uma voz autêntica e expressiva, a capacidade de alcançar f_o mínimas mais baixas reforça a ideia de estabilidade e autoridade. Juntas, estas características acústicas formam uma base para a percepção de segurança em vozes disfônicas e não disfônicas.

Para pacientes com disfonia ou outras condições que afetam a voz, a capacidade de comunicar-se de maneira que seja percebida como segura pode influenciar significativamente a qualidade de vida, as relações sociais e profissionais. Assim, uma abordagem terapêutica que inclua técnicas para melhorar o nível de ruído espectral e a capacidade de modulação da frequência fundamental pode ajudar os pacientes a alcançarem não apenas uma voz saudável, mas também uma voz que reflete confiança e estabilidade emocional.

O nível de ruído espectral na faixa de frequência entre 100-2600 Hz e a medida H1A1 foram selecionados como preditores para explicar a variabilidade no julgamento da atitude **calma** para vozes de falantes disfônicos e não disfônicos.

Um nível mais baixo de ruído espectral sugere uma qualidade vocal menos desviada, o que pode ser, intuitivamente, associado à calma. A presença de ruído em excesso na voz pode ser percebida como tensão, elementos contrários à noção de calma. Portanto, mesmo que o valor de β seja relativamente baixo ($\beta = 0{,}044$), a significância estatística ($p < 0{,}001$) indica que reduções no ruído espectral podem influenciar a percepção de calma, sugerindo que vozes com menor componente de ruído nessa faixa de frequência são associadas a uma maior tranquilidade ou serenidade.

A medida H1A, ao destacar a predominância do primeiro harmônico sobre o segundo, serve como um marcador acústico de timbre vocal. Valores elevados de H1A sugerem uma característica de voz com componente de soprosidade, que é comumente associada à suavidade e, portanto, à calma. Isso se alinha com a noção de que sinais vocais com maior presença de ar e menor tensão nas pregas vocais podem ser percebidos como mais relaxados e tranquilos, evocando uma sensação de calma nos ouvintes.

A partir do ponto de vista do modelo estatístico (R2 =0,436), é importante notar que, embora as duas medidas sejam preditores significativos, elas explicam menos da metade da variabilidade na percepção da atitude calma. Isso implica que existem outros fatores, possivelmente incluindo características acústicas não examinadas, aspectos linguísticos ou até mesmo fatores psicossociais, que também influenciam essa percepção. A complexidade da comunicação humana e a riqueza da expressão vocal significam que a percepção da calma é influenciada por uma gama ampla de variáveis, muitas das quais podem ser sutis ou subjetivas.

O grau geral de desvio vocal e a f0 máxima na fala conectada foram preditores significativos do julgamento de ouvintes em relação à atitude de independência. Primeiramente, o grau geral de desvio vocal, que demonstrou um coeficiente beta negativo, sugere que quanto menor o desvio vocal percebido, maior é a percepção da atitude de independência pelo ouvinte. Dessa forma, vozes menos desviadas podem ser, inconscientemente, associadas a uma maior estabilidade emocional e psicológica, atributos frequentemente relacionados à independência.

Por outro lado, a f0 máxima na fala conectada, com um coeficiente beta positivo, indica que um aumento nesta medida está associado ao aumento da percepção de independência. Sendo assim, vozes que alcançam f0 máximas mais elevadas podem ser percebidas como mais dinâmicas e potencialmente mais expressivas de autonomia e determinação. Este achado apoia a ideia de que características vocais associadas à força e à assertividade podem influenciar a percepção social de traços de personalidade como a independência.

O modelo estatístico, com um R^2 de 0,682, indica que estas duas medidas juntas explicam significativamente uma grande parte da variabilidade

na percepção da atitude de independência em relação a vozes disfônicas e não disfônicas.

Os modelos preditivos para as atitudes de resistência, segurança, calma e independência, que compõem a dimensão de atividade, apresentam algumas características em comum, revelando aspectos fundamentais sobre como as vozes disfônicas e não disfônicas são percebidas em relação a essas atitudes.

O nível de ruído espectral aparece como um preditor significativo em três dos quatro modelos (resistência, segurança e calma), com valores de β positivos. Isso sugere que um maior nível de ruído espectral, dentro de determinadas faixas de frequência, pode estar associado à percepção de maior resistência, segurança e calma. Isso pode indicar que um certo grau de rugosidade ou soprosidade na voz, interpretado como ruído espectral, é valorizado nessas dimensões de atividade, talvez sinalizando autenticidade ou uma forma de presença vocal que transmite confiança e estabilidade.

A f0 (mínima para resistência e segurança; máxima para independência) é um preditor significativo em três modelos. Isso indica que características específicas da f0, seja em seu extremo inferior ou superior, influenciam a percepção de atitudes dentro da dimensão de atividade. Para resistência e segurança f0 mínimas mais altas podem indicar uma capacidade de vocalização firme e controlada, enquanto para independência, uma frequência fundamental máxima mais alta pode sugerir vigor ou energia.

A associação do nível de ruído espectral com atitudes positivas na dimensão de atividade sugere que uma voz que não é puramente "limpa" ou isenta de ruído pode ser percebida como mais autêntica ou genuína, qualidades valorizadas nas interações humanas. O ruído espectral pode contribuir para uma sensação de "corpo" ou "textura" vocal, que em contextos adequados, transmite confiança e serenidade.

A importância das extremidades da f_o na percepção das atitudes sugere que a capacidade de expressar uma gama dinâmica de tons vocais é valorizada. f_o mínimas mais altas podem ser interpretadas como uma voz mais firme e estável, enquanto máximas mais altas podem refletir vigor ou intensidade, qualidades associadas a uma personalidade independente e assertiva.

Os modelos preditivos desenvolvidos para avaliar a percepção das atitudes de agradabilidade, simpatia, saúde, extroversão, poder, força,

autoridade, competência, resistência, segurança, calma e independência a partir de vozes disfônicas e não disfônicas representam uma contribuição substancial para os estudos de linguística, especialmente no campo da sociolinguística e da psicolinguística. Eles fornecem uma base empírica para entender como certas características vocais influenciam a percepção social dos falantes. Ao quantificar a relação entre medidas acústicas e perceptivas e a atribuição de certas atitudes sociais, esses modelos aprofundam nossa compreensão de como a voz atua como um veículo de identidade social e pessoal. Eles destacam a complexidade da comunicação humana, mostrando que a voz carrega informações significativas além do conteúdo verbal, capazes de influenciar a percepção de características pessoais e sociais, como confiabilidade, autoridade, saúde e simpatia. Isso abre novas possibilidades de pesquisa sobre a influência da variação vocal na interação social e no estabelecimento de relações interpessoais.

Para a prática clínica, especialmente na fonoaudiologia, a contribuição desses modelos é igualmente valiosa. Eles oferecem insights sobre como intervenções vocais podem não apenas restaurar ou melhorar aspectos funcionais da voz, mas também como podem influenciar a percepção social dos pacientes. Na reabilitação vocal, entender essas dinâmicas pode levar a abordagens mais holísticas, considerando não apenas a qualidade vocal em termos acústicos, mas também os efeitos psicossociais da voz. Isso é particularmente relevante para pacientes com disfonia, para quem as percepções de competência, autoridade e agradabilidade podem ser prejudicadas. Ao integrar esses modelos preditivos na avaliação e no tratamento, os clínicos podem visar melhorias que transcendem a funcionalidade vocal, abordando também o impacto da voz na qualidade de vida e nas interações sociais dos pacientes, promovendo assim uma recuperação mais abrangente e significativa.

6.5 MODELO BASEADO EM ML PARA PREDIÇÃO DA VALÊNCIA DAS ATITUDES DOS OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS E NÃO DISFÔNICAS DE FALANTES NATIVOS DO PORTUGUÊS BRASILEIRO, A PARTIR DAS MEDIDAS ACÚSTICAS

A implementação de modelos baseados em ML para prever as atitudes dos ouvintes a partir de medidas acústicas de vozes disfônicas e não disfônicas representa um avanço significativo na interseção entre a tecnologia, a linguística, a fonoaudiologia e a psicologia social. Os resultados obtidos pelos modelos utilizando os métodos Kernel SVM e Naive Bayes, ambos com acurácia e AUC superiores a 0.90, evidenciam a alta eficácia dessas abordagens computacionais na análise e interpretação das nuances acústicas que influenciam a percepção humana. Tal nível de precisão sugere não apenas a validade dos métodos escolhidos, mas também a relevância das características acústicas e da variável “sexo” na modulação das atitudes dos ouvintes em relação à disfonia.

O modelo Naive Bayes, ao selecionar 22 medidas acústicas dentre as 47 analisadas, além da variável “sexo”, demonstra uma abordagem abrangente, capturando um espectro amplo de características vocais que influenciam a percepção das atitudes. A alta consistência observada nas medianas de acurácia para todas essas variáveis, superiores a 0.98, indica uma seleção criteriosa de fatores que são robustos preditores da valência das atitudes dos ouvintes. Isso sublinha a complexidade da voz humana como um instrumento de comunicação, capaz de transmitir uma variedade de sinais sociais e emocionais por meio de características acústicas sutis. A análise detalhada dessas variáveis fornece uma compreensão mais profunda de como cada aspecto da voz contribui para a percepção social, um insight valioso tanto para a teoria linguística quanto para a prática clínica.

Por outro lado, a parcimônia do modelo Kernel SVM, que selecionou apenas quatro medidas acústicas, reflete uma estratégia focada que enfatiza a importância de características específicas na determinação das atitudes dos ouvintes. A eficácia de um modelo mais simplificado pode ser particularmente atraente para aplicações práticas, onde a simplicidade e a facilidade de interpretação são cruciais. A contribuição estável do shimmer, a alta confiabilidade do CPP como preditor, a sensibilidade do modelo à f_o máxima e a

variabilidade observada com a f_0 mínima, juntos, fornecem uma visão compreensiva dos elementos-chave que moldam a percepção das atitudes para vozes disfônicas e não disfônicas. A variabilidade na contribuição da f_0 mínima, em particular, destaca a complexidade da percepção vocal e a necessidade de considerar como diferentes aspectos da voz interagem na formação de juízos de atitude.

Os resultados alcançados por esses modelos, portanto, não apenas confirmam a aplicabilidade de técnicas de ML na análise da percepção vocal, mas também enriquecem nossa compreensão sobre a interação entre características acústicas específicas e a percepção social da voz. Esse conhecimento tem implicações significativas tanto para a teoria linguística, ao iluminar os mecanismos subjacentes à comunicação vocal e suas percepções associadas, quanto para a prática clínica, ao oferecer diretrizes para intervenções focadas na melhoria da qualidade vocal e na modulação de suas percepções sociais.

Modelos baseados em IA apresentam uma habilidade notável em perceber, definir e prever elementos que se destacam em diversos contextos relacionados ao som da voz humana. Isso inclui a identificação de falantes, ativação de sistemas por meio do sinal de voz, percepção de emoções por meio de amostras vocais, como em sistemas de assistentes virtuais, capazes de detectar se um usuário está frustrado, feliz ou irritado, com a finalidade de ajustar as respostas de acordo com essas emoções, oferecendo um atendimento mais personalizado, por exemplo (SHARMA *et al*., 2021; BOTA *et al*., 2019). Dessa forma, observa-se também sua capacidade em detectar ou prever para além da impressão causada por uma voz, mas as atitudes dos ouvintes ao escutarem-na, a partir de medidas perceptivo-auditivas e/ou medidas acústicas do sinal vocal.

A f0 é um dos principais parâmetros acústicos utilizado na prática clínica (RODRIGUES, 2007; LAUKKA, 2004; SCHERER, 1996; SCHERER, 1989), determinada por uma complexa interação entre comprimento, massa e tensão das pregas vocais e referida pela sigla f0, é o menor componente periódico resultante da vibração das pregas vocais, tornando-se a primeira frequência produzida na glote.

A f0 tornou-se objeto de muitos estudos para analisar a sua relação com o julgamento de atitudes por meio da voz. Como já comentado, estudos relacionam esta medida a atratividade, inclusive relacionada ao gênero dos falantes, em que ouvintes do sexo masculino apresentam maior atratividade para vozes femininas com a f0 elevada, isto é, mais aguda (ZHENG YI, *et al*., 2020; RAYMOND M., *et al*., 2019; SUIRE A., *et al*., 2018; SHIRAZI T. N., PUTS D. A., ESCASA M. J., 2018; BABEL M., MCGUIRE G, KIN J., 2014; XU, *et al*., 2013; P. J. FRACARRO *et al*., 2013; O'CONNOR R. D., *et al*., 2012; DEFRADAS M. B., *et al*., 2012; FRACCARO *et al*., 2011; XUAN, 2011; JONES *et al*., 2010; SAXTON T. K., *et al*., 2009), e os ouvintes do sexo feminino, julgam as vozes masculinas com a fo reduzida, como mais atraentes (ZHENG YI, *et al*., 2020; RAYMOND M., *et al*., 2019; SEBESTA P., *et al*., 2017; TSANTANI M. S. , *et al*., 2017; XU, *et al*., 2013; O'CONNOR R. D., *et al*., 2012; DEFRADAS M. B., *et al*., 2012; SIMMONS L. W., PETERS M., RHODES G., 2011; JONES *et al*., 2010; SAXTON T. K., *et al*., 2009).

No contexto comunicativo, a voz tem o poder de fazer com que o falante faça modificações de acordo com a intenção e situação da transmissão da mensagem, e o ouvinte gera uma resposta/ reação em relação ao que escuta. Esse processo está diretamente relacionado a prosódia do comunicador, sobretudo à entonação. A modulação vocal, fazendo-se utilização de variações de frequência durante uma conversa, relaciona-se às medidas de f0 máxima e mínima, em conjunto com a qualidade vocal, podem modificar o discurso e consequentemente a percepção do ouvinte (MEDRADO R., FERREIRA L. P., BELHAU, M., 2005).

A presença de disfonia pode repercutir em variações na f0 (OLIVEIRA R. C., *et al*., 2021; SAMPAIO M., *et al*., 2020; KUANG E LIBERMAN, 2016), sobretudo quando há lesões em pregas vocais. Tais variações podem modificar o padrão vocal esperado para determinado gênero, profissão e estrutura física, por exemplo. Além disso, pode restringir a capacidade de modulação vocal durante o discurso, afetando a expressividade do falante, as ênfases, a psicodinâmica, a transmissão de emoções, muitas vezes, tornando o discurso monótono e pouco atrativo por não ser capaz de transmitir a mensagem de forma efetiva. Este fato pode ter relação com o julgamento de atitudes com valências mais negativas por parte dos ouvintes.

A medida acústica *shimmer* foi preditora das atitudes dos ouvintes diante de vozes disfônicas. Esta medida está associada à variação de amplitude de vibração das pregas vocais, a partir da instabilidade ou irregularidade do sinal sonoro (CHRISTOPHER Y., *et al*., 2023; LOPES *et al*., 2012), relacionando-se com a percepção dos parâmetros rugosidade e soprosidade, que tendem a estar alterados em vozes com valores elevados de *shimmer.*

Pesquisas internacionais mencionaram que essa medida acústica não se correlacionou significativamente com o julgamento de atitudes para vozes saudáveis e disfônicas (LATOSZEK B. *et al*., 2019; HUGHES S. M., MOGILSKI J. K. E HARRISON M. A., 2014), no entanto, deve-se observar aspectos sociais como a cultura e língua do falante que repercutem na percepção dos ouvintes. Além disso, os estudos descrevem aspectos relacionas apenas à f0, e não informam com clareza se as vozes utilizadas nos experimentos eram saudáveis ou se também foram apresentadas vozes disfônicas aos ouvintes.

Para falantes do português brasileiro, amostra utilizada no presente estudo, o *shimmer* se apresentou como um preditor relacionado ao julgamento positivo e negativo das atitudes de ouvintes também brasileiros, o que pode ter relação com o fato de vozes mais desviadas apresentar julgamento com valência mais negativa, tanto para o grau geral, como para rugosidade e soprosidade.

O CPP também se destacou como preditora da valência da atitude dos ouvintes diante de vozes disfônicas. O CPP está relacionado à regularidade e intensidade dos harmônicos na avaliação acústica, isto é, representa a relação entre o espectro harmônico, o som de qualidade, e o ruído presente na emissão (SAMLAN *et al*., 2013). É uma medida cepstral e não depende dos ciclos da f0 para ser extraída, utiliza tarefas de voz e fala, e por isso é mais sensível na análise de vozes mais desviadas. Por isso, tem sido indicado mais recentemente, entre as muitas opções, como um dos parâmetros mais precisos para avaliação da qualidade vocal (PATEL *et al*, 2018; MARYN Y. *et al*., 2009).

O *CPP* é objeto de diversos estudos, a partir da sua capacidade nas análises acústicas da disfonia (LOPES *et al*., 2019; WATTS C. R., AWAN S. N., 2011; AWAN S. N. *et al*., 2010; AWAN S. N., ROY N., DROMEY C., 2009; HEMANACKAH Y. D., MICHAEL D. D., HEMAN-ACKAH Y. D. *et al*., 2003; GODING G. S. JR., 2002) e entre a percepção do ouvinte e o componente

prosódico da fala (YANUSHEVSKAYA *et al*., 2017; S. KAKOUROS *et al*., 2017; M. GARELLEK, J. WHITE, 2015).

Diante da efetividade dessa medida e a sua forte relação com a qualidade vocal, diversos estudos com o objetivo de relacionar medidas acústicas e julgamento que utilizaram algoritmos da ML utilizaram o CPP em seus métodos com a finalidade de rastreio de alterações vocais secundárias à alterações funcionais ou neurológicas (ZHAOYAN, 2023; AMATO F., 2023; BYEON, H. 2021; GEORGOPOULOS V. C., 2020; HUIYI WI *et al*., 2018, MEHTA D. D., 2015), mas o CPP não fez parte dos modelos estatísticos utilizados para os resultados, visto que as pesquisas se concentram em medidas acústicas tradicionais, restringindo a possiblidade de comparações e análises conjuntas (HUGHES, 2014).

No presente estudo o CPP apresentou, para além de uma relação significante, como preditivo para o julgamento da valência positiva e negativa de atitudes diante de vozes disfônicas, por meio do método ML. Este achado deve-se ao fato de vozes alteradas apresentarem um julgamento mais negativo por parte do ouvinte e do CPP ter alta sensibilidade e poder discriminante para vozes disfônicas. Assim, quando menor o valor do pico cepstral, menor a qualidade do sinal vocal e mais negativo é o julgamento do ouvinte diante do discurso.

Dessa forma, encontrou-se no presente estudo resultados promissores e pioneiros, estimulando reflexões científicas e clínicas para sua aplicabilidade. Sabe-se que é necessária a continuidade de experimentos com objetivos semelhantes, que permitam maiores discussões e comparações. No entanto destaca-se a grande contribuição a partir da utilização de métodos atuais e robustos de IA para predição de medidas para o julgamento de atitudes de ouvintes, por meio da ML, tornando as avaliações vocais mais práticas, sobretudo relacionadas à comunicação.

Os achados possibilitam a investigação de apenas quatro medidas específicas, determinantes para avaliação da percepção de ouvintes, sendo possível padronização de protocolo, inclusive a delimitação das tarefas de fala que devem ser gravadas, tais como: vogal sustentada e fala encadeada/contagem, que atendem tanto as medidas de frequência e perturbação, quanto às medidas cepstrais.

No contexto clínico, os modelos de IA possibilitam que os programas de reabilitação sejam mais eficazes e direcionados, acelerando o processo de avaliação, aperfeiçoamento e recuperação diante da capacidade de analisar uma dados específicos da voz, permitindo uma avaliação mais detalhada e precisa dos distúrbios.

Portanto, a capacidade de prever a valência das atitudes dos ouvintes em relação a vozes disfônicas e não disfônicas tem implicações importantes para a prática clínica. Compreender quais características acústicas influenciam a percepção das vozes pode ajudar os profissionais de saúde a desenvolver intervenções mais eficazes para melhorar a comunicação vocal de pacientes disfônicos. Do ponto de vista linguístico, esses achados podem contribuir para uma melhor compreensão de como a prosódia e outros aspectos acústicos da fala afetam a percepção social dos falantes. Isso pode ser útil para o desenvolvimento de tecnologias de processamento de fala e para aprimorar o ensino de línguas, onde a pronúncia e a entonação são componentes críticos.

7 CONCLUSÃO

Há diferenças significativas e a presença de correlações nas medidas perceptivo-auditivas e acústicas entre vozes julgadas com valência positiva e negativa, através que cada atributo investigado.

As vozes com maior intensidade de desvio na qualidade vocal foram julgadas mais negativamente, sobretudo as rugosas e soprosas. As medidas acústicas ceptrais e espectrais estão associadas ao julgamento positivo e negativo dos ouvintes para vozes disfônicas e não disfônicas. O maior componente de ruído e irregularidade vocal influencia na percepção e julgamento do ouvinte. Observou-se que quanto mais distante dos valores preconizados como normais, mais negativo é o julgamento.

As medidas acústicas e perceptuais que justificaram o julgamento de atitudes para o julgamento da valência geral – positiva ou negativa foram: EAV - GG, fomin-CS, SNL3-CS. O grau geral foi preditivo para todas as atitudes, indicando que o desvio vocal influencia na percepção do ouvinte sobre vozes disfônicas. A frequência fundamental mínima se destacou como preditiva para as atitudes das dimensões avaliação, potência e atividade. Quartil (F_0), e o nível de ruído espectral (SNL), foram parâmetros preditivos para potência e atividade.

As medidas acústicas de fo máxima e fo mínima, CCP e s*himmer* foram capazes de predizer a valência das atitudes dos ouvintes em relação às vozes dos falantes disfônicos nativos do português brasileiro.

Diante dos avanços tecnológicos, é fundamental a utilização de modelos baseados em inteligência artificial para melhorar a qualidade do diagnóstico e tratamento da disfonia. Tais tecnologias não apenas aprimoram a precisão diagnóstica, mas também oferecem terapias mais personalizadas e eficazes, promovendo melhores resultados e uma qualidade de vida significativamente melhor para os pacientes, aliado à implicações para os estudos em sociolinguística, psicolinguística e fonética, através de análises robustas dos impactos sociais da fala.

REFERÊNCIAS

A. SLUIJTE; V. Van Heuven. Acoustic correlates of linguistic stress and accent in Dutch and American English,"in **Proc. of ICSLP**.1996. 2, 1996, pp. 630–633.

AGHEYISI, R.; FISHMAN, J. The study of language atitudes. ***International Journal of the Sociology of Language,*** v.3, p.5-19, 1980.

AGUILERA, V. de A. Crenças e atitudes linguísticas: o que dizem os falantes das capitais brasileiras. **Estudos Linguísticos**, São Paulo, v. 2, n. 37, p. 105-112, 2008.

ALTENBERG, E. P.; FERRAND, C. T. Perception of individuals with voice disorders by monolingual English, bilingual Cantonese−English, and bilin- gual Russian-English women. **J Speech Lang Hear Res.** 2022; 49:879–887.

_____. Perception of individuals with voice disorders by monolingual English, bilingual Cantonese−English, and bilingual Russian-English women. **J Speech Lang Hear Res.** 2006; 879– 887.

ALLARD, E. R.; WILLIAMS D. F. Listeners' perceptions of speech and language disorders. **J Commun Disord.** 2008: 1(41) 108–123.

ALPAYDIN, E. **Introduction to Machine Learning**. MIT Press; 2014.

ALLPORT, G. W. Attitudes. In: Fishbein, M. **Readings in atitude theory and measurement.** New York: Wiley, 1967.

ALKMIM, T. M. Sociolingüística: Parte I. In: Mussalim, F.; Bentes, A. C. (Orgs.). **Introdução à lingüística:** domínios e fronteiras. 4. ed. São Paulo: Cortez, 2004, p. 21-44.

ALTENBERG, E. P.; FERRAND, C. T. Perception of individuals with voice disorders by monolingual English, bilingual Cantonese−English, and bilin- gual Russian-English women. **J Speech Lang Hear Res.** 2022; 49:879– 887.

AMATO, F. *et al*. **Análise de voz baseada em aprendizado de máquina e estatística de pacientes com doença de Parkinson:** uma pesquisa. Sistemas Especialistas com Aplicativos, 119651. 2023.

AMIR, O.; LEVINE-YUNDOF, R. Listeners' attitude toward people with dysphonia. **J Voice**. 2013; 27(4):524-532.

ANDERSON, R. C. *et al*. Vocal Fry May Undermine the Success of Young Women in the Labor Market. ***PLoS ONE.*** 2014: 9(5): 975-981.

ANGELILLO, M.; DI MAIO, G.; COSTA, G. Prevalence of occupational voice disorders in teachers. **J Prev Med Hyg.** 2009; 50:26–32.

ARMSTRONG, M. M.; LEE, A. J.; & FEINBERG, D. R. A house of cards: Bias in perception of body size mediates the relationship between voice pitch and perceptions of dominance. **Anim. Behav**. 2019; 147: 43–51.

AUNG, T. & PUTS, D. Voice pitch: A window into the communication of social power. **Curr. Opin. Psychol.** 2020; 33: 54–161.

AWAN, S. N. *et al*. **Quantifying dysphonia severity using a spectral/cepstral-based acoustic index:** comparisons with auditory-perceptual judgements from the CAPE-V. Clin Linguist Phon. 2010;
24(9):742-58.

AWAN, S. N.; ROY, N.; DROMEY, C. **Estimating dysphonia severity in continuous speech: application of a multi-parameter spectral/cepstral model.** Clin Linguist Phon. 2009; 23(11):825-41.

BABEL, M.; MCGUIRE, G.; KING, J. Towards a more nuanced view of vocal attractiveness. ***PLoS One*;** 2014; *9*(2): 443-449.

BAINBRIDGE, K. E.; ROY, N.; LOSONCZY, K. G. **Voice disorders and associated risk markers among young adults in the united states**. laryngo- scope. 2017; 127:2093–2099.

BARKAT-DEFRADAS, M. *et al*. Dimension esthétique des voix normales et dysphoniques: Approches perceptive et acoustique. **TIPA. Travaux interdisciplinaires sur la parole et le langage**, 2012; 28.

BEHROOZMAND, R.; ALMASGANJ, F. Optimal selection of wavelet-packet-based features using genetic algorithm in pathological assessment of patients' speech signal with unilateral vocal fold paralysis. **Comput Biol Med.** 2007; 37:474–485.

BEN, B. B. *et al*. A meta-analysis: acoustic measurement of roughness and breathiness. **J Speech Lang Hear Res.** 2018; 61(2):298-323.

BERARDI, M. L.; HUNTER, E. J.; FERGUSON, S. H. **Talker age estimation using machine learning.** Proc Meet Acoust. 2017; 30:040014.

BERTSIMAS, Dimitris; WIBERG, Holly. Machine learning in oncology: methods, applications, and challenges. **JCO Clinical Cancer Informatics,** v. 4, 2020.

BESTELMEYER, P., ROUGER J., DEBRUINE L. Auditory adaptation in vocal affect perception. **Cognition.** 2010; 117:217–223.

BERGNER, Anouk S.; HILDEBRAND Christian; HÄUBL, Gerald.
"Machine Talk: How Verbal Embodiment in Conversational AI Shapes Consumer–BrandRelationships". **Journal of Consumer Research**. 2023.

BERTELSEN C.; ZHOU, S.; HAPNER, E. R. Sociodemographic characteris- tics and treatment response among aging adults with voice disorders in the United States. **JAMA Otolaryngol Head Neck Surg**. 2018; 144:719– 726.

BLOOD, G. W.; MAHAN, B. W.; HYMAN, M. Judging personality and appearancefrom voice disorders. **J Commun Disord**. 1979; 12:63–67.

BORKOWSKA, B.; PAWLOWSKI, B. Female voice frequency in the context of dominance and attractiveness perception. **Animal Behaviour.** 2011; 82(1): 55–59.

BRUCE, P.; BRUCE, A. **Practical Statistics for Data Scientists:** 50 essential concepts. [S.l.]:O'reilly, 2019.

BRUCKERT, L.; BESTELMEYER P.; LATINUS M. biology JRC, 2010 undefined. Vocal attractiveness increases by averaging. Disponível em: https://www.sciencedirect. com/science/article/pii/S0960982209020491. Acesso em: 23 out. 2023. Elsevier [Internet].

BRUCKERT, L. *et al.* Vocal attractiveness increases by averaging. ***Curr Biol***. 2010, 20(2):116-120.

BYEON, Haewon. Comparing ensemble-based machine learning classifiers developed for distinguishing hypokinetic dysarthria from presbyphonia. **Applied Sciences**, v. 11, n. 5, p. 2235, 2021.

CHOUDHURY, Ambika; GUPTA, Deepak. A survey on medical diagnosis of diabetes using machine learning techniques. In: Recent Developments in Machine Learning and Data Analytics: IC3 2018. **Springer Singapore**, 2019. p. 67-78.

COHEN, S. M. *et al*. Direct health care costs of laryngeal diseases and disorders. **Laryngoscope.** 2012; 122(7):1582-1588.

COHEN, S. M. **Self-reported impact of dysphonia in a primary care popu- lation: an epidemiological study.** Laryngoscope. 2010; 120:2022–2032.

CONSERVA, Késsia Cecília Fernandes; FRANÇA, Fernanda Pereira; LOPES, Leonardo Wanderley. Auditory-perceptual and acoustic measures in women with and without vocal nodules. **Audiology-Communication Research**, v. 27, 2022.

CORBARI, C. C. **Atitudes Linguísticas:** Um estudo nas localidades paranaenses de Irati e Santo Antônio do Sudoesta {Tese}. Universidade Federal da Bahia. Salvador, 2013.

CUOCOLO, Renato Caruso *et al*. Learning in oncology: a clinical appraisal. **Cancer letters**, v. 481, p. 55-62, 2020.

DA COSTA, V.; PRADA, E.; ROBERTS, A. Voice disorders in primary school teachers and barriers to care. **J Voice***.* 2012; 26: 69–76.

DAVENPORT, T. *et al.* How artificial intelligence will change the future of marketing. **Journal of the Academy of Marketing Science.** 2020; 48: 24-42.

DE BRUIJNE, Marleen. Machine learning approaches in medical image analysis: From detection to diagnosis. **Medical image analysis**, v. 33, p. 94-97, 2016.

DUTTA, Mohan. Communication, power and resistance in the Arab Spring: Voices of social change from the South. **Journal of creative communications**, v. 8, n. 2-3, p. 139-155, 2013.

EADIE, T. L.; RAJABZADEH, R.; ISETTI, D. D. The effect of information and severity on perception of speakers with adductor spasmodic dysphonia. **Am J Speech Lang Pathol.** 2017; 26:327–341.

EGELMAN, Serge; PEER, Eyal. Predicting privacy and security attitudes. **Acm Sigcas Computers and Society**, v. 45, n. 1, p. 22-28, 2015.

ENGLERT, Marina Taborda. **Erro Perceptivoauditivo de Vozes Humanas e Sintetizadas.** Dissertação apresentada à Universidade Federal de São Paulo – Escola Paulista de Medicina, para obtenção do título de Mestre em Ciências. 2016.

EKBERG, M. *et al*. Acoustic Features Distinguishing Emotions in Swedish Speech. **Journal of Voice.** 2023.

EVANGELISTA, D. D. S. *et al.* Predictive Factors of Listeners' Attitudes Related to Dysphonic Voices in Native Brazilian Portuguese. **J Voice.** 2022; (22).

FEINBERG, D. R. *et al*. Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. ***Animal behaviour*****,** 2005, *69*(3), 561-568.

_____. *et al.* Correlated preferences for men's facial and vocal masculinity. **Evolution and Human Behavior**, 2008; *29*(4), 233-241.

FRACCARO, P. J. *et al.* Faking it: deliberately altered voice pitch and vocal attractiveness. **Animal Behaviour,** 2013; *85*(1), 127-136.

_____. *et al*. Correlated male preferences for femininity in female faces and voices. **Evolutionary Psychology.** 2010: *8*(3).

FROSI, V. M.; MANTOVANI, G. O.; FAGGION, C. M. Da estigmatização à solidariedade: atitudes linguísticas na RCI. In: FROSI *et. al*. **Cultura e atitudes linguísticas.** Editora EDUCS. Caxias do Sul: Rio Grande do Sul, 2010.

GARRET, P. **Attitudes to language**. Cambridge: Cambridge University

Press, 2010.

GAVIDIA-CEBALLOS, L.; HANSEN, J. H. **Direct speech feature estimation using an iterative EM algorithm for vocal fold pathology detection.** IEEE Trans Biomed Eng. 1996; 43:373–383.

GELFER, M. P.; BENNET Q. E. Speaking Fundamental Frequency and Vowel Formant Frequencies: Effects on Perception of Gender. **J Voice***.* 2013; 27(5): 556-566.

GEORGOPOULOS, Voula C. "Análise avançada de tempo-frequência e aprendizado de máquina para detecção de voz patológica." ***2020 12º Simpósio Internacional de Sistemas de Comunicação, Redes e Processamento Digital de Sinais (CSNDSP)*.** IEEE, 2020.
GOBL, Christer; AILBHE Ni Chasaide.The role of voice quality in communicating emotion, mood and attitude. **Speech communic ation***,* 2003. 40(1): 189-212.

GOVINDU, Aditi; PALWE, Sushila. Early detection of Parkinson's disease using machine learning. **Procedia Computer Science**, v. 218, p. 249-261, 2023.

HEGDE, S.; SHETTY, S.; RAI, S. A survey on machine learning approaches for automatic detection of voice disorders. **J Voice**. 2019; 33(6):947.e11–947.e33.

HEMAN-ACKAH, Y. D. Cepstral peak prominence: a more reliable measure of dysphonia. **Ann Otol Rhinol Laryngol**. 2003, Apr; 112(4):324-33.

HEMAN-ACKAH, Y. D.; MICHAEL, D. D.; GODING, G. S. JR. The relationship between cepstral peak prominence and selected parameters of dysphonia. **J Voice.** 2002 Mar; 16(1):20-7.

HEMMERLING, D.; SKALSKI, A.; GAJDA, J. Voice data mining for laryngeal pathology assessment. **Comput Biol Med.** 2016; 69:270–276.

HODGES-SIMEON, Carolyn R.; STEVEN, J. C. Gaulin; PUTS, David A. Different vocal parameters predict perceptions of dominance and attractiveness*.* ***Human Nature,*** 2010; 21(4): 406-427.

HENTON C. G. Creak as a sociophonetic marker. **J Acoust Soc Am.** 20176; 80 (suppl 1):S50. Disponível em: https://doi.org/10.1121/1.2023837. (S1)-S50. Acesso em: 23 out. 2023.

HOLMES, W.; BIALIK, M.; FADEL, C. **Artificial Intelligence in Education – Promises and Implications for Teaching and Learning.** Boston: The Center for Curriculum Redesign, 2019.

HUGHES, S. M.; MOGILSKI, J. K.; HARRISON, M. A. The perception and parameters of intentional voice manipulation. **Journal of Nonverbal Behavior,** 2014; *38*(1), 107–127.

HUGHES, Susan M.; PUTS, David A. Vocal modulation in human mating and competition. **Philosophical Transactions of the Royal Society B**, v. 376, n. 1840, p. 20200388, 2021.

HUGHES, S. M.; PASTIZZO, M. J.; GALLUP, G. G. The sound of symmetry revisited: Subjective and objective analyses of voice. ***Journal of Nonverbal Behavior*;** 2008; *32*(2), 95–108.

IDRISOGLU, A. *et al.* Applied Machine Learning Techniques to Diagnose Voice-Affecting Conditions and Disorders: Systematic Literature Review. **J Med Internet Res.** 2023; 19(25):461-471.

IRANI, F.; ABDALLA, F.; HUGHES, S. Perceptions of voice disorders: a survey of Arab adults. ***Logoped Phoniatr Vocol*.** 2014; 39:87–97.

ISETTI, D.; XUEREB L.; EADIE T. L. Inferring speaker attributes in adductor spasmodic dysphonia: ratings from unfamiliar listeners. **Am J Speech Lang Pathol.** 2014; 23:134–145.

IYER, R.; HOSMER, D. W.; LEMESHOW, S. Applied Logistic Regression. **The Statistician**. 1991; 40:458.

I YANUSHEVSKAYA, A.; NI, Chasaide; C. Gobl. Cross-speaker ´ variation in voice source correlates of focus and deaccentuation, in **Proc. of Interspeech**, 2017, pp. 1034–1038.

JAREK, K.; MAZUREK, G. Marketing and Artificial Intelligence. **Central European Business Review.** 2019; 8(2): 130-141.

JO, T. **Machine learning foundations**. Available at: Disponível em: https://link. springer.com/content/pdf/10.1007/978-3-030-65900-4.pdf. Acesso em: xx xxx xxxx. 2021.

JONES, B. C. *et al.* A domain-specific opposite-sex bias in human preferences for manipulated voice pitch. **Animal Behaviour,** 2010; *79*(1), 57-62.

KAPSNER-SMITH, M. R. *et al.* Clinical cutoff scores for acoustic indices of vocal hyperfunction that combine relative fundamental frequency and cepstral peak prominence. **Journal of Speech, Language, and Hearing Research,** 2022. *65*(4), 1349-1369.

KISSEL, Imke *et al*. Listeners' attitudes towards voice disorders: An interaction between auditory and visual stimuli. **Journal of Communication Disorders**, v. 99, p. 106241, 2022.

KLOFSTAD, C. A, ANDERSON, R. C, NOWICKI, S. Perceptions of competence, strength, and age influence voters to select leaders with lower-pitched voices. ***PLoS ONE*.** 2015; 10(8):133-139.

KOJIMA, T.; FUJIMURA, S.; HASEBE, K. **Avaliação objetiva da voz patológica utilizando inteligência artificial baseada na escala GRBAS.** Disponível em: http://www.jvoice.org/article/ _S0892199721004094/texto completo. Acesso em: 17 abr. 2022.

KREIMAN, J.; SIDTIS, D. Voices and listeners: toward a model of voice perception. **Acoustics Today***.* 2011; 7(4): 188-197.

KREIMAN, J. Today DSA. 2011 undefined. Voices and listeners: toward a model of voice perception. acousticstoday.org [Internet]. Disponível em: https:// acousticstoday.org/wp-content/uploads/2017/09/Article_1of4_fro m_ATCODK_7_4.pdf. Acesso em: 00 dez. 2023.

KRISCHKE, S. *et al.* Quality of life in dysphonic patients**. Journal of Voice,** *19*(1), 132-137. 2005.

KYRIAKOU, K., PETINOU K., PHINIKETTOS, I. Risk factors for voice disorders in university professors in Cyprus. **J Voice***.* 2018; 32:643-652.

KÜHNE, Katharina; FISCHER, Martin H.; ZHOU, Yuefang. The human takes it all: Humanlike synthesized voices are perceived as less eerie and more likable. evidence from a subjective ratings study. **Frontiers in neurorobotics**, v. 14, p. 105, 2020.

LAMBERT, W. E. *et al*. Attitudinal and cognitive aspects of intensive study of a second language. **Journal of Abnormal and Social Psychology,** v.66, n.4, p.358-68, 1963.

LATINUS, M.; CRABBE, F.; BELIN, P. **Learning-induced changes in the cerebral processing of voice identity.** Cereb Cortex. 2011. Dec; 21(12):2820-2828.

LALLH, A. K.; PUTNAM, Rochet A. The effect of information on listener's attitudes toward speakers with voice or resonance disorders. **J Speech Lang Hear Res.** 2000; 43:782–795.

LANDIS, J. R.; KOCH, G. G. The measurement of observer agreement for categorical data. **Biometrics.** 1977; 33:159-165.

LASS, N. J. Adolescents' perceptions of normal and voice-disordered children. **J Commun Disord.** 1991; 24:267–274.

LATOSZEK, B. *et al*. The Acoustic Breathiness Index (ABI): A Multivariate Acoustic Model for Breathiness*.* **J Voice.** 2017; 31(4):511-527.

LE GR´EZAUSE, E. **Um and Uh, and the Expression of Stance in Conversational Speech.** 2017. Tese. Universidade de Washington.

LEE, Seo-Young *et al*. Expressing Personalities of Conversational Agents Through Visual and Verbal Feedback. **Electronics,** 2019; 8 (7), 794.

LEWIS, J. W.; TALKINGTON, W. J.; WALKER, N. A. Human cortical organization for processing vocalizations indicates representation of harmonic structure as a signal attribute. **J Neurosci**. 2022; 29:2283–2296.

LIN, H. *et al*. Identification of digital voice biomarkers for cognitive health. **Explor Med.** 2020; 1:406-417.

LIU, Xuan; YI, Xu. What makes a female voice attractive?. **ICPhS.** 2011.

LIU, X. ICPhS Yx. undefined. What Makes a Female Voice Attrac- tive? researchgate.net [Internet] Cited November 3 2022. Disponível em: https://www. researchgate.net/profile/Yi-Xu-20/publication/267219102_WHAT_MA KES_A_FEMALE_VOICE_ATTRACTIVE/links/575a935b08ae414 b8e464baa/WHAT-MAKES-A-FEMALE-VOICE-ATTRACTIVE. Acesso em: 04 dez 2023. pdf; 2011; 17-21.

LÓPEZ-DE-IPIÑA, K. *et al*. On the selection of non-invasive methods based on speech analysis oriented to automatic Alzheimer disease diagnosis. **Sensors** (Basel). 2013; 13(5):6730-6745.

LOPES, L. W. *et al*. Cepstral measures in the assessment of severity of voice disorders. **CoDAS.** 2019; 31(4):255-283.

_____. *et al.* Acurácia das medidas acústicas tradicionais e formânticas na avaliação da qualidade vocal. **CoDAS.** 2018; 30(5).

LOPES, L. W.; CAVALCANTE, D. P.; COSTA, P. O. Intensidade do desvio vocal: integração de dados perceptivo-auditivos e acústicos em pacientes disfônicos. **CoDAS.** 2014; 26(5):382-388.

_____. *et al*. Gravidade dos distúrbios vocais em crianças: correlações entre dados perceptivos e acústicos. **Jornal da Voz.** 2012. *26* (6), 819-e7.

LYBERG-AHLANDER V, ... MH... journal of speech. 2015 undefined. **Does the Speaker's Voice Quality Influence Children's Performance on a Language Comprehension Test?.** Cited November 3 2022. Disponível em: https:/ www.tandfonline.com/doi/abs/10.3109/17549507.2014.898098; 2015 February 1. Taylor & Francis [Internet]:17(1):63. Acesso em: xx xxxx xxxx.

M. GARELLEK; J. White. Phonetics of Tongan stress. **Journal of the International Phonetic Association.** 2015; 45(1). 13– 34.

MA, Liye.; SUN, Baohong. Machine learning and AI in marketing–Connecting computing power to human insights. **International Journal of Research in Marketing.** 2020; 37(3): 481-504.

MA, E. P.; YU, C. H. Listeners' attitudes toward children with voice problems. **J Speech Lang Hear Res.** 2013; 56:1409–1415.

MAIDMENT, J. A. The phonetic description of voice quality.
J Int Phonet Asso. 1981; 11:78–84.

MALI, R. *et al.* Parkinson's disease data analysis and prediction using ensemble machine learning techniques. In: **Mobile Computing and Sustainable Informatics:** Proceedings of ICMCSI 2021.

MANDOUST, Sadegh Bafandeh; BOLANDRAFTAR, Mohammad. Application of k-nearest neighbor (knn) approach for predicting economic events: Theoretical background. **International Journal of Engineering Research and Applications,** v. 3, n. 5, p. 605-610, 2013.

MCALEER, P.; TODOROV, A. E. BELIN, P. How Do You Say "Hello"? Personality Impressions from Brief Novel Voices. **Psychology**. 2014; 9,1-9.

McDERMOTT, J. H. Auditory Preferences and Aesthetics: Music, voices, and everyday sounds. Neuroscience of Preference and Choice. In: **McDERMOTT J**. **H.** New York: p 227-255. 2012.

MCLEAN, Graeme; OSEI-FRIMPONG, Kofi; BARHORST, Jennifer. Alexa, do voice assistants influence consumer brand engagement?–Examining the role of AI powered voice assistants in influencing consumer brand engagement. **Journal of Business Research**, v. 124, p. 312-328, 2021.

MCKINNON, S. L.; HESS, C. W.; LANDRY, R. G. Reactions of college students to speech disorders. **J Commun Disord.** 1986; 19:75–82.

MEDRADO, Reny; FERREIRA, Leslie Piccolotto; BEHLAU, Mara. Voice-over: perceptual and acoustic analysis of vocal features. **Journal of voice**, v. 19, n. 3, p. 340-349, 2005.

MEHTA, Daryush D. Usando monitoramento ambulatorial de voz para investigar distúrbios vocais comuns: atualização de pesquisa. **Fronteiras em bioengenharia e biotecnologia,** v. 3, p. 155, 2015.

MELLEY, Le; SATALOFF, R. T. Beyond the Buzzwords: artificial Intelligence in Laryngology. **J Voice.** 2022; 36:2–3.

MORSOMME, D.; MINEL, L.; VERDUYCKT I. **Impact of teachers'voice quality on children's language.** 2011.

N. CAMPBELL; M. Beckman. "Stress, prominence, and spectral tilt," in Intonation: **Theory, Models and Applications**, 1997, pp. 67–70.

NEMR, K. *et al*. Voice deviation, dysphonia risk screening and quality of life in individuals with various laryngeal diagnoses. ***Clinics*****,** *73*. 2018.

OLIVEIRA, Rafaella Cristina; ANA, C. C. Gama; MAX, D. C. Magalhães. "Fundamental voice frequency: acoustic, electroglottographic, and

accelerometer measurement in individuals with and without vocal alteration." **Journal of Voice**, 35.2 (2021): 174-180.

OSGOOD, Charles E. *et al*. **The Measurement of Meaning**. Urbana, Illinois: University of Illinois Press, pp.342, 1957.

P. PRIET; M. Ortega-Llebaria. Stress and accent in Catalan and Spanish: Patterns of duration, vowel quality, overall intensity, and spectral balance, in **Proc. of Speech Prosody**, 2006, pp. 337–340.

PASCHEN, J.; JAN, K.; T. C. Kietzmann. Artificial intelligence (AI) and its implications for market knowledge in B2B marketing. **Journal of business & industrial marketing.** 2019; 34(7): 1410-1419.

PEARSELL, S.; PAPE, D. The effects of different voice qualities on the perceived personality of a speaker. **Frontiers in Communication**, v. 7, p. 909427, 2023.

PERLOFF, R. **The dynamics of persuasion:** communication and attitudes in the 21st Century. New York: Academic Press, 2008.

PETTY, Brian E. Beauty and Attractiveness in the Human Voice.
Journal of Voice. 36(4), 507-514, 2022.

PISANSKI, K.; RENDALL, D. The prioritization of voice fundamental frequency or formants in listeners' assessments of speaker size,masculinity and attractiveness. **The Journal of theAmerican Society of America,** 2011; *129*(4), 2201–2212.

PISANSKI, K. perception DFO handbook of voice, 2019 undefined. vocal attractiveness. researchgate.net [Internet]. Disponível em: https://www.researchgate. net/profile/Katarzyna-Pisanski/publication/334151544_Vocal_attractiveness/links/5d1a6461a6fdcc2462b728ea/Vocal-attractiveness.pdf. 2018. Acesso em: 23 dez. 2023.

PORCARO C.; EVITTS P.; KING N. Effect of dysphonia and cognitive- perceptual listener strategies on speech intelligibility. **J Voice.** 2020; 34:806e7–806e18.

PUTS, D. A. *et al.* Women's attractiveness changes with estradiol and progesterone across the ovulatory cycle. **Hormones and behavior.** 2013; *3*(1): 13-19.

_____. *et al*. Intrasexual competition among women: Vocal femininity affects perceptions of attractiveness and flirtatiousness. **Personality and Individual Differences,** 2011; *50*(1): 111-115.

_____. A. *et al.* **Vozes masculinas como sinais de dominância:** frequências vocais fundamentais e formantes influenciam as atribuições de dominância entre os homens. Evolução e Comportamento Humano. 2007; 28: 340–344.

RAYMOND, M.; SUIRE, A.; BARKAT-DEFRALDAS, M. **Male vocal quality and its relation to females' preferences.** Evolutionary Psychology. 2019; 17(3): 147-152.

RE, D. E. *et al.* Preferences for very low and very high voice pitch in humans. **PloS one**, 2012; *7*(3): 327-334.

RICHENS, Jonathan G.; LEE, Ciarán M.; JOHRI, Saurabh. Improving the accuracy of medical diagnosis with causal machine learning. **Nature communications.** v. 11, n. 1, p. 3923, 2020.

RITCHINGS, R. T.; Mcgillion, M.; Moore, C. J. Pathological voice quality assessment using artificial neural networks. **Med Eng Phys.** 2002; 24:561–564.

ROBIN, J. *et al*. Evaluation of Speech-Based Digital Biomarkers: Review and Recommendations. **Digit Biomark**. 2020; 4(3):99-108.

ROGERSON, J. **Voice BDJ of. Undefined. Is there an effect of dysphonic teachers' voices on children's processing of spoken language?.** 2005

ROY, M. M. Artificial Intelligence in Pharmaceutical Sales & Marketing–A Conceptual Overview. **IJIRT**.2022; 8(11): 897-902

RUMBACH, A.; KHAN A.; BROWN M. Voice problems in the fitness industry: factors associated with chronic hoarseness. **Int J Speech Lang Pathol.** 2015; 17: 441–450.

SAMPAIO, M. *et al*. Effects of fundamental frequency, vocal intensity, sample duration, and vowel context in cepstral and spectral measures of dysphonic voices. **Journal of Speech, Language, and Hearing Research,** *63*(5), 1326-1339. 2020.

SANTOS, M.; KOENIGKAM, M. *et al*. Inteligência artificial, aprendizado de máquina, diagnóstico auxiliado por computador e radiômica: avanços da imagem rumo à medicina de precisão. **Radiologia brasileira**; 2019; 52: 387-396.

SARA, M.; SAEED, S.; RABIEE A. Speech Emotion Recognition Based on a Modified Brain Emotional Learning Model. Biologically inspired cognitive architectures. **Elsevier.** 2017; 19:32-38.

SAVILLE-TROIKE, M. **The ethnography of communication:** an introduction. 2nd ed. Oxford: Blackwell, 2003.

SAXTON, T. K. *et al*. Face and voice attractiveness judgments change during adolescence. **Evolution and Human Behavior,** 2006; *30*(6), 398-408.

SAYED, Md Abu. Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms. **Journal of Computer Science and Technology Studies**, v. 5, n. 4, p. 142-149, 2023.

SCHERER, K. R. (1977). Affektlaute and vokale Embleme. In R. Posner H. P. Reinecke (Eds.) Zeichenprozesse — **Semiotische Forschung in den Einzelwissenschaften** (pp. 199–214). Wiesbaden: Athenaion.

SEBESTA, P. Voices of Africa: acoustic predictors of human male vocal attractiveness. **Animal Behaviour,** 2017; 127, 205-211.

SHAILAJA, K.; SEETHARAMULU, B.; JABBAR, M. Machine Learning in Healthcare: A Review. In: **Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)**; Março 29-31, 2018; Coimbatore, Tamil Nadu, India p. 910-914.

SHARMA, T. *et al*. Emotion Analysis for predicting the emotion labels using Machine Learning approaches. In 2021 **IEEE 8th Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering** (UPCON) (pp. 1-6).

SHIRAMIZU, Victor Kenji M. *et al*. "The role of valence, dominance, and pitch in social perceptions of artificial intelligence (AI) conversational agents' voices." **Scientific Reports**. (2022); 12(1): 224 – 231.

SHIRAZI, T. N.; PUTS, D. A.; ESCASA-DORNE, M. J. Filipino women's preferences for male voice pitch: Intra-individual, life history, and hormonal predictors. **Adaptive Human Behavior and Physiology,** 2018; *4*(2): 188-206.

SIMMONS, L. W.; PETERS, M.; RHODES, G. Low pitched voices are perceived as masculine and attractive but do they predict semen quality in men?. **PloS one,** 2011; *6*(12).

SKRINDA, I. *et al.* Body height, immunity, facial and vocal attractiveness in young men. **Naturwissenschaften**, 2014; 101(12), 1017-1025.

SMITH, E.; VERDOLINI K.; GRAY S. **Effects of voice disorders on quality of life.** 2016, December 13 Cited August 5 2022; 4:223-244.

SOROKOWSKI, P. *et al.* Voice of authority: Professionals lower their vocal frequencie when giving expert advice. **Journal of Nonverbal Behavior***, 2019; 43(2), 257–269.*

SRIRAM, T. V. Intelligent Parkinson disease prediction using machine learning algorithms. **Int. J. Eng. Innov. Technol**, v. 3, p. 212-215, 2013.

STOKLASA, J.; TALÁŠEK T.; STOKLASOVÁ J. Semantic differential for the twenty-first century: scale relevance and uncertainty entering the semantic space. **Qual Quant.** 2019; 53:435–448.

STUART, R.; PETER, N. **Artificial intelligence:** a modern approach. 3rd ed. Available at: Disponível em: https://www.academia.edu/download/61853459/Artificial-Intelligence-A-Modern-Approach-3rd-Edition-by- Stuart-Russell-Peter-Norvig20200121-107745-13gd7bj.pdf. Acesso em: 21 nov. 2023.

STYLER, W. **On the Acoustical and Perceptual Features of Vowel Nasality.** 2015. Tese. Universidade do Colorado.

SUIRE, Alexandre; RAYMOND, Michel; BARKAT-DEFRADAS, Melissa. "Human vocal behavior within competitive and courtship contexts and its relation to mating success." **Evolution and Human Behavior**, 39.6 (2018): 684-691.

S. KAKOUROS, O.; Ras¨ Anen; P. Alku. Evaluation of spectral ¨ tilt measures for sentence prominence under different noise con- ditions, in **Proc. of Interspeech**, 2017, pp. 3211–3215.

TARUNIKA, K.; R. B. Pradeeba and P. Aruna. "Applying machine learning techniques for speech emotion recognition." **9th International Conference on Computing, Communication and Networking Technologies (ICCCNT).** IEEE, 2018.

TAYLOR, Sammi *et al*. Age-related changes in speech and voice: spectral and cepstral measures. **Journal of Speech, Language, and Hearing Research**, v. 63, n. 3, p. 647-660, 2020.

TITZE, I. R.; LEMKE J., MONTEQUIN D. Populations in the U.S. workforce who rely on voice as a primary tool of trade: a preliminary report. **J Voice.** 2022; 11:254–259.

TOLMEIJER, Suzanne *et al.* "Female by Default? Exploring the Effect of Voice Assistant Gender and Pitch on Trait and Trust Attribution". **Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems**, 1–7, Crossref. 2021.

TRACY, J. M. *et al.* Investigating voice as a biomarker: Deep phenotyping methods for early detection of Parkinson's disease. **J Biomed Inform.** 2020; 104:1362-1371.

TSANTANI, M. S. *et al.* Low vocal pitch preference drives first impressions irrespective of context in male voices but not in female voices. **Perception.** 2017; *45*(8), 946-963.

UMENO, H. *et al*. A summary of the Clinical Practice Guideline for the Diagnosis and Management of Voice Disorders. **Auris Nasus Larynx.** 2020; 47(1):7-17.

VAN, Borsel J.; JANSSENS, J.; DE BODT, M. Breathiness as a feminine voice characteristic: A perceptual approach. **Journal of Voice,** *23*, 291-294. 2009.

VAN HOUTTE, E.; CLAEYS, S.; WUYTS, F. The impact of voice disorders among teachers: vocal complaints, treatment-seeking behavior, knowledge of vocal care, and voice-related absenteeism. **J Voice.** 2011; 25:570–575.

VAN, Stan J.; MEHTA, D. D.; HILLMAN, R. E. Recent Innovations in Voice Assessment Expected to Impact the Clinical Management of Voice Disorders. **Perspect ASHA SIGs.** 2017; 2(3): 4-13.

VERDOLINI, K.; RAMIG L. O. **Review:** occupational risks for voice problems. **Logoped Phoniatr Vocol.** 2001; 26: 37–46.

VILLAR, Acnwb; KORN G. P.; AZEVEDO, R. R. Perceptual-auditory and acoustic analysis of air traffic controllers' voices pre- and postshift. **J Voice***.* 2016; 3: 768-775.

VUKOVIC, J. *et al.* Self-rated attractiveness predicts individual differences in women's preferences for masculine men's voices. **Personality and Individual Differences.** 2008; *45*(6), 451-456.

WAARAMAA, T. *et al*. Impressions of Personality from Intentional Voice Quality in Arabic-Speaking and Native Finnish-Speaking Listeners*.* **J Voice.** 2021; 35(2); 326-328.

WAARAMAA, T.; LUKKARILA, P.; JARVINEN, K. Voice AGJ of, 2021 unde € fined. Impressions of personality from intentional voice quality in Arabic- speaking and native Finnish-speaking listeners. Disponível em: https://www.sciencedirect.com/science/article/pii/S0892199719302462. Accesso em: xx xxx xxx November 3 2022. Elsevier [Internet].

WANG, C.; FRED, A. L.; DA SILVA, H. P. **A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals.** IEEE. 2019; 7:140990-141020.

WATTS, C. R.; AWAN, S. N. Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. **J Speech Lang Hear Res.** 2011; 54(6):1525-37.

WILLIAMS, N. R. Occupational groups at risk of voice disorders: a review of the literature. **Occu Med.** 2003; 53:456–460.

WILSON, J. A. *et al.* The quality of life impact of dysphonia. **Clinical Otolaryngology & Allied Sciences**, 2021. *27*(3), 179-182.

WILSON, Deirdre; TIM, Wharton. Relevance and prosody. **Journal of pragmatics.** 2006: 38(10) 1559-1579.

WU, Huiyi, **Um método de aprendizagem profunda para detecção de voz patológica usando redes convolucionais de crenças profundas**. Interfala, 2018.

XU, Yi, *et al*. "Human vocal attractiveness as signaled by body size projection." ***PloS one*** **8.4 (**2013): e62397.

Y. MARYN *et al*. **Acoustic measurement of overall voice quality: A meta-analysis,**" **The Journal of the Acoustical Society of America,** v. 126, n. 5, pp. 2619–2634, 2009.

YAM, Christopher *et al.* Prevalência do uso de medidas objetivas de voz em práticas laringológicas nos Estados Unidos. **Jornal da Voz**, 2023.

YAMASAKI, R. *et al.* Auditory-perceptual Evaluation of Normal and Dysphonic Voices Using the Voice Deviation Scale. **J Voice.** 2017; 31(1):67-71.

YU, G. *et al*. **Speech emotion recognition using voiced segment selection algorithm.** ECAI. 2016; 28(1): 682-1683.

ZHANG, Zhaoyan. Voice Feature Selection to Improve Performance of Machine Learning Models for Voice Production Inversion. **Jornal de Voz,** 37.4 (2023): 479-485.

ZHENG, Yi, *et al*. Vocal attractiveness and voluntarily pitch-shifted voices. **Evolution and Human Behavior**, v. 41, n. 2, p. 170-175, 2020.

ZHAO, C. *et al*. Is infant neural sensitivity to vocal emotion associated with mother-infant relational experience?. ***PLoS One****.* 2019;14(2): 299-314.

ZUCKERMAN, M.; MIYAKE, K. The attractive voice:Whatmakes it so? **Journal of Nonverbal Behavior,** 1993; *17*(2), 119–135.

ANEXOS

ANEXO A

ESCALA DE JULGAMENTO DE ATITUDES

ASSOCIADAS A VOZES DISFÔNICAS

Prezado colaborador, o objetivo desta pesquisa é avaliar as impressões (atitudes) ocasionadas nos ouvintes ao escutarem diferentes tipos de vozes. Dessa forma, você escutará frases produzidas por diferentes falantes e deverá marcar o quanto cada uma dessas vozes transmite as impressões descritas abaixo. Para cada impressão (atributo) os números variam de 1 (mais negativo) a 6 (mais positivo), ou seja, você deverá marcar o número que mais representa a impressão transmitida pela voz dos falantes. Quanto mais negativa positiva a impressão, marque mais próximo ao número **"1"**. Quanto mais positiva a impressão, marque mais próximo ao número **"6"**.

Falante 01:

Desagradável	① ② ③ ④ ⑤ ⑥	**Agradável**
Fraco	① ② ③ ④ ⑤ ⑥	**Poderoso**
Antipático	① ② ③ ④ ⑤ ⑥	**Simpático**
Fraco	① ② ③ ④ ⑤ ⑥	**Forte**
Frágil	① ② ③ ④ ⑤ ⑥	**Resistente**
Introvertido	① ② ③ ④ ⑤ ⑥	**Extrovertido**
Doente	① ② ③ ④ ⑤ ⑥	**Saudável**
Submisso	① ② ③ ④ ⑤ ⑥	**Autoritário**
Agitado	① ② ③ ④ ⑤ ⑥	**Calmo**
Inseguro	① ② ③ ④ ⑤ ⑥	**Seguro**
Incompetente	① ② ③ ④ ⑤ ⑥	**Competente**

Dependente	① ② ③ ④ ⑤ ⑥	Independente

ANEXO B

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado (a) Senhor (a)

Esta pesquisa é sobre **PREFERÊNCIAS E ATITUDES DOS OUVINTES EM RELAÇÃO A VOZES NORMAIS E ALTERADAS** e está sendo desenvolvida pelo pesquisador Deyverson da Silva Evangelista aluno do Curso de Pós Graduação em Linguística da Universidade Federal da Paraíba, sob a orientação do Prof. Dr. Leonardo Wanderley Lopes. O objetivo do estudo é avaliar as impressões (atitudes) ocasionadas nos ouvintes ao escutarem diferentes tipos de vozes.

Existem diversas técnicas que avaliam a atitude linguística. Nesta pesquisa será utilizada a escala de diferencial semântico, que reúne diversos atributos e escores menores que quatro indicam uma atitude positiva e maiores que quatro, uma atitude negativa. Solicitamos a sua colaboração para o preenchimento da Escala de julgamento de atitudes associadas a vozes disfônicas, como também sua autorização para apresentar os resultados deste estudo em eventos da área de saúde e publicar em revista científica. Por ocasião da publicação dos resultados, seu nome será mantido em sigilo. Informamos que essa pesquisa não oferece riscos, previsíveis, para a sua saúde.

Esclarecemos que sua participação no estudo é voluntária e, portanto, o(a) senhor(a) não é obrigado(a) a fornecer as informações e/ou colaborar com as atividades solicitadas pelo Pesquisador(a). Caso decida não participar do estudo, ou resolver a qualquer momento desistir do mesmo, não sofrerá nenhum dano, nem haverá modificação na assistência que vem recebendo na Instituição.Os pesquisadores estarão a sua disposição para qualquer esclarecimento que considere necessário em qualquer etapa da pesquisa.

Diante do exposto, declaro que fui devidamente esclarecido(a) e dou o meu consentimento para participar da pesquisa e para publicação dos resultados. Estou ciente que receberei uma cópia desse documento.

_____

Assinatura do Participante da Pesquisa ou Responsável Legal

Contato do Pesquisador (a) Responsável:

Caso necessite de maiores informações sobre o presente estudo, favor ligar para o pesquisador:

Deyverson da Silva Evangelista.

Telefone: (083) 998875257

Ou

Comitê de Ética em Pesquisa do Centro de Ciências da Saúde da Universidade Federal da Paraíba Campus I - Cidade Universitária - 1º Andar – CEP 58051-900 – João Pessoa/PB

 (83) 3216-7791 – E-mail: eticaccsufpb@hotmail.com

Atenciosamente,

_____

Assinatura do Pesquisador Responsável

_____

Assinatura do Pesquisador Participante

ANEXO C

Autorização do Laboratório Integrado de Estudos da Voz

LIEV
Laboratório Integrado de Estudos da Voz

AUTORIZAÇÃO

Eu, Profa. Dra. Anna Alice Figueiredo de Almeida (SIAPE:1634755), docente do Departamento de Fonoaudiologia da Universidade Federal da Paraíba/UFPB e líder do Laboratório Integrado de Estudos da Voz (LIEV), autorizo o pesquisador Prof. Dr. Leonardo Wanderley Lopes a utilizar as instalações e o banco de dados do referido Laboratório para sua pesquisa de dissertação de mestrado intitulada " **PREDITORES ACÚSTICOS DAS ATITUDES DE OUVINTES EM RELAÇÃO ÀS VOZES DISFÔNICAS E NÃO DISFÔNICAS**".

João Pessoa-PB, 29 de abril de 2022.

aafalmeida

Dra. ANNA ALICE FIGUEIREDO DE ALMEIDA

SIAPE: 16685454

Departamento de Fonoaudiologia Centro de Ciências da Saúde / Universidade Federal da Paraíba
Campus I – João Pessoa / PB E-mail: depfono@ccs.ufpb.br (83) – 3216-7831

ANEXO D

CENTRO DE CIÊNCIAS DA SAÚDE DA UNIVERSIDADE FEDERAL DA PARAÍBA - CCS/UFPB

PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: POTENCIAIS EVOCADOS PARA VOZES DISFÔNICAS E NÃO DISFÔNICAS
Pesquisador: DEYVERSON DA SILVA EVANGELISTA
Área Temática:
Versão: 2
CAAE: 26353419.0.0000.5188
Instituição Proponente: Centro de Ciências Humanas, Letras e Artes
Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 5.249.734