



Pontifícia Universidade Católica de São Paulo
Programa de Pós-graduação em Comunicação Humana e Saúde

PUC-SP

Gustavo do Amaral Ambros

Notas agudas: aspectos acústicos em tenores eruditos

Mestrado em Comunicação Humana e Saúde

São Paulo
2024



Pontifícia Universidade Católica de São Paulo
Programa de Pós-graduação em Comunicação Humana e Saúde

PUC-SP

Gustavo do Amaral Ambros

Notas agudas: aspectos acústicos em tenores eruditos

Mestrado em Comunicação Humana e Saúde

Dissertação apresentada à Banca Examinadora da Pontifícia Universidade Católica de São Paulo, como exigência parcial para obtenção do título de MESTRE em Comunicação Humana e Saúde na área de Voz, Comunicação, Saúde e Arte, sob orientação da Prof.^a Dr.^a Marta Assumpção de Andrada e Silva.

São Paulo
2024

Banca Examinadora

O presente trabalho foi realizado com apoio da **Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES)** – Código de Financiamento 001 – processo número 88887.661690/2022-00,

e do **Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)** – processo número 130341/2022-3.

Aos meus pais,

Luiz Tadeu Ambros e Elizabeth Regina do Amaral,
pelo amor e pelo apoio incondicionais, sempre, em
todas as minhas empreitadas;

e aos meus irmãos,

Tadeu, Natália e Carolina, *que, longe ou perto,*
ausentes ou presentes, são essenciais na minha vida.

AGRADECIMENTOS

Em primeiro lugar, à Prof.^a **Dr.^a Marta Assumpção de Andrada e Silva**, orientadora deste trabalho, pelo auxílio fundamental em adequar a magnitude do projeto de um pesquisador inexperiente a um estudo exequível e relevante, pela objetividade nas recomendações sobre a organização e a redação desta dissertação e, principalmente, pela confiança em mim e pelas palavras de incentivo e tranquilização nos momentos mais desafiadores desta jornada.

À Prof.^a **Dr.^a Zuleica Camargo**, pelas valiosas reflexões a respeito da voz, da fala e do canto, pelas vivências metodológicas da fonética experimental, e pela participação nas bancas de qualificação e defesa deste trabalho.

Aos professores **Dr.^a Sandra Madureira** e **Dr. Mario Fontes**, pelos conteúdos sobre física e linguística, nunca dissociados de suas implicações na comunicação e na expressividade, e sempre que possível relacionados ao canto e à música.

Aos professores **Dr.^a Lucia Santaella**, **Dr.^a Ruth Palladino** e **Dr. Luiz Augusto de Paula Souza (Tuto)**, que, em suas respectivas disciplinas de Semiótica, Linguagem e Neuropsicologia, ampliaram meus horizontes de pensamento sobre as relações humanas e a comunicação; e **Dr.^a Léslie Piccolotto Ferreira**, pelos ensinamentos sobre metodologia e prática em pesquisa, e pela confiança depositada em mim nos estudos conduzidos pelo grupo em suas disciplinas.

Às professoras **Dr.^a Ana Flavia Zuim** e **Dr.^a Joana Mariz**, por prontamente aceitarem fazer parte das bancas de qualificação e defesa desta dissertação, e por suas valiosas recomendações e importantes reflexões acerca das temáticas aqui desenvolvidas.

À Prof.^a **Dr.^a Ana Carolina Ghirardi**, por suas contribuições na pré-qualificação, imprescindíveis na definição da metodologia desta pesquisa.

Aos dois grandes amigos que o bacharelado em Música na UFRGS - Universidade Federal do Rio Grande do Sul me presenteou: o tenor **Felipe Bertol**, pelo companheirismo e pelo apoio, pelo acolhimento e pela companhia em São Paulo, pelo imprescindível auxílio em procurar sujeitos para a pesquisa, e por todas nossas conversas sobre canto e voz; e a compositora **Touanda Beal**, pelos diálogos sobre

arte e música, sobre filosofia e a vida, e pelos conselhos e apoio emocional, mesmo que em esporádicas conversas *online*, e separados por um oceano. Agradeço ainda à também colega de graduação **Ana Carolina Sant'Anna**, mezzo-soprano, pela hospedagem nas minhas primeiras semanas em São Paulo.

À amiga, cantora e colega professora de canto **Luísa Marques**, meu primeiro contato com o universo da fisiologia vocal e com o canto contemporâneo, pelas conversas sobre técnica vocal e pelo incentivo nos momentos de dúvida.

À amiga fonoaudióloga **Dr.^a Eduarda Cunha**, sempre generosa em compartilhar seus conhecimentos, pelos incontáveis diálogos sobre voz, sobre ciência e pesquisa, e pelos conselhos e encorajamento.

Aos professores de canto que tive durante a vida, em especial a soprano **Maíra Lautert**, o baixo-barítono **Ricardo Barpp** e o tenor **Flávio Leite**, que contribuíram, cada um à sua maneira, não apenas para meu desenvolvimento técnico como cantor, mas principalmente para meu pensamento crítico sobre música, performance e arte.

Aos técnicos de áudio **Ernesto Foschi**, do laboratório de rádio da Faculdade de Filosofia, Comunicação, Letras e Artes (FAFICLA) da PUC-SP, e **Luciano Vargas**, do estúdio K30 em Porto Alegre, pelo indispensável auxílio, pelo bom humor, pelas conversas e pela atenção durante os procedimentos de gravação desta pesquisa.

Aos **cantores** que aceitaram participar deste estudo, pelo interesse na pesquisa, pela disponibilidade e pela contribuição voluntária. Muito obrigado!

RESUMO

Introdução: a emissão de notas agudas impõe dificuldades técnicas para os cantores, independentemente do gênero musical, uma vez que necessitam de adaptações de estruturas do trato vocal. Tenores atuantes em óperas têm o desafio de emitir seus agudos sem o auxílio de equipamentos eletrônicos para amplificação, enquanto mantêm a qualidade vocal desejada. Ao considerar o aspecto acústico, a teoria de interação não-linear fonte-filtro sugere que, para a emissão destas notas, é preciso modificar o posicionamento dos articuladores e ajustar as ressonâncias do trato vocal em relação aos harmônicos da fonte glótica. **Objetivo:** descrever acusticamente os ajustes do trato vocal supraglótico e as características do espectro radiado das notas mais agudas da tessitura de tenores eruditos. **Método:** a amostra intencional foi composta de cinco tenores profissionais, atuantes em óperas. A gravação das emissões sonoras foi realizada em um estúdio de gravação profissional, portanto acusticamente tratado. Foram captadas as emissões: fala espontânea; vogais /a/, /e/, /i/, /o/ e /u/ faladas em sentenças-veículo; execução do trecho de uma ária de ópera; vocalização de duas escalas ascendentes de oito tons. As gravações foram analisadas nos *softwares* de livre acesso Praat versão 6.2.23 e Sopran versão 1.0.28, e extraídos: frequências dos dois primeiros formantes; amplitude dos quatro primeiros harmônicos e do pico na região do formante do cantor (entre 2,5 e 3,5 kHz), nível de pressão sonora e espectro médio de longo termo. **Resultados:** as frequências dos formantes das diferentes configurações vocálicas variaram ao longo da escala. O primeiro formante das vogais /e/ e /i/ tendeu a subir e o segundo a descer; nas vogais /o/ e /u/ ambos tenderam a subir. Os espaços vocálicos de todos os sujeitos reduziram da voz falada para a voz cantada, e das notas graves para as notas agudas. Na zona de passagem, foi mais frequente a sintonia do primeiro formante com o segundo harmônico e do segundo formante com o quarto harmônico; acima da passagem, a sintonia do segundo formante com o terceiro harmônico. Não foi identificada relação entre sintonia de formantes e o nível de pressão sonora das notas. Em todos os sujeitos, o nível de pressão sonora foi maior na voz cantada. Nos espectros médios de longo termo, todos os sujeitos exibiram o formante do cantor e apresentaram declínio espectral menos acentuado na voz cantada. **Conclusão:** foi constatada a convergência das frequências dos dois primeiros formantes em todas as vogais ao longo das escalas, o que sugere a necessidade de ajustes articulatórios para a produção das notas agudas por tenores eruditos. Não foi identificada a obrigatoriedade de uma manobra específica de ajuste dos formantes em relação aos harmônicos para a emissão das notas agudas, mas sim o emprego de diferentes soluções. A sintonia do segundo formante ao terceiro harmônico foi utilizada principalmente para manejar a zona de passagem, enquanto no topo da escala foi mais observada a distribuição equilibrada de energia entre os parciais mais graves. Os achados indicam a necessidade de diferentes ajustes de ressonâncias ao longo da tessitura aguda de tenores eruditos.

Descritores: Canto; Qualidade da Voz; Acústica; Treinamento da Voz.

ABSTRACT

Introduction: emitting high notes poses technical challenges for singers, regardless of the musical genre, as they require adaptations of the vocal tract structures. Tenors performing in operas face the challenge of producing their high notes without the aid of electronic equipment for amplification while maintaining the desired vocal quality. Considering the acoustic aspects, the theory of nonlinear source-filter interaction suggests that modifying the positioning of articulators and adjusting resonances in relation to the harmonics of the glottal source is necessary for emitting these notes. **Objective:** to describe acoustically the adjustments of the supraglottal vocal tract and the characteristics of the radiated spectrum of the highest notes in the tessitura of classical tenors. **Method:** the intentional sample consisted of five professional tenors active in operas. Sound emissions were recorded in a professionally treated recording studio. The emissions captured included: spontaneous speech; vowels /a/, /e/, /i/, /o/, and /u/ spoken in carrier sentences; performance of an excerpt of an opera aria; vocalization of two ascending eight-tone scales. Recordings were analyzed using the open-source software Praat version 6.2.23 and Sopran version 1.0.28, and the following were extracted: frequencies of the first two formants; amplitude of the first four harmonics and of the peak in the singer's formant region (between 2.5 and 3.5 kHz), sound pressure level, and long-term average spectrum. **Results:** the formants of the different vowel configurations varied along the scale. The first formant of vowels /e/ and /i/ tended to rise, and the second to fall; for vowels /o/ and /u/, both tended to rise. The vowel spaces of all subjects reduced from spoken voice to sung voice, and from low to high notes. In the passaggio zone, the tuning of the first formant with the second harmonic, and the second formant with the fourth harmonic was more frequent; above the passaggio, tuning of the second formant with the third harmonic was observed. No relationship was identified between formant tuning and the sound pressure level of the utterances. In all subjects, the sound pressure level was higher in the sung voice. In the long-term average spectra, all subjects exhibited the singer's formant and showed a less pronounced spectral slope in the sung voice. **Conclusion:** convergence of the frequencies of the first two formants in all vowels across the scale was observed, suggesting the need for articulatory adjustments for the production of high notes by classical tenors. The necessity of a specific formant adjustment maneuver in relation to the harmonics for emitting high notes was not identified; instead, different solutions were employed. Tuning of the second formant to the third harmonic was mainly used to manage the passaggio zone, while at the top of the scale, a more balanced distribution of energy among the lower partials was observed. These findings indicate the need for different resonance adjustments along the high tessitura of classical tenors.

Descriptors: Singing; Voice Quality; Acoustics; Voice Training.

LISTA DE ABREVIATURAS

CT	músculos cricotireóideos
EGG	Eletroglotografia
EMG	Eletromiografia
EMLT	espectro médio de longo termo
f0	frequência fundamental
F1, F2, ..., Fn	primeiro formante, segundo formante, ..., <i>n</i> -ésimo formante
FC	formante do cantor
<i>F_n-H_n</i>	sintonia do <i>n</i> formante com o <i>n</i> harmônico
H1, H2, ..., H_n	primeiro harmônico, segundo harmônico, ..., <i>n</i> -ésimo harmônico
H1-H2	diferença de amplitude entre o primeiro e o segundo harmônicos
M0	mecanismo vibratório laríngeo 0
M1	mecanismo vibratório laríngeo 1
M2	mecanismo vibratório laríngeo 2
M3	mecanismo vibratório laríngeo 3
<i>MFDR</i>	<i>maximum flow declination rate</i>
NPS	nível de pressão sonora
PEV	perfil de extensão vocal
Psub	pressão subglótica
TA	músculos tireoaritenóideos
TCLE	termo de consentimento livre e esclarecido
T MDF	taxa máxima do declínio de fluxo

LISTA DE FIGURAS

Figura 1 – Espectros da fonte glótica.	8
Figura 2 – Efeito do filtro sobre o espectro da fonte glótica.	9
Figura 3 – Diferenciação entre formantes e ressonâncias.	10
Figura 4 – Esquematisação da reatância no trato vocal.....	12
Figura 5 – Diferentes relações entre formantes e harmônicos em função de f_0	16
Figura 6 – Esquematisação da variação de F1 e F2 em função das vogais.	17
Figura 7 – Seleção do trecho da vogal falada na sentença-veículo.	51
Figura 8 – <i>Zoom</i> na forma de onda da vogal na imagem acima..	51
Figura 9 – Exemplo de seleção de cruzamento de zero na forma de onda.	52
Figura 10 – Seleção do trecho das notas das escalas.....	52
Figura 11 – <i>Display</i> da Filtragem Inversa no <i>software</i> Sopran.....	53
Figura 12 – Leitura da amplitude dos harmônicos no <i>software</i> Sopran.....	55
Figura 13 – Seleção do sinal para o procedimento de calibração.	56
Figura 14 – NPS da fala espontânea e da execução da ária.	69

LISTA DE QUADROS

Quadro 1 – F1 e F2 das vogais faladas e cantadas (/a/, /e/ e /i/).	61
Quadro 2 – F1 e F2 das vogais faladas e cantadas (/o/ e /u/).	62
Quadro 3 – Espaços vocálicos das vogais faladas e das notas C2 e C4.	63
Quadro 4 – Amplitudes de H1, H2, H3, H4 e do pico na região do formante do cantor das vogais cantadas (/a/, /e/ e /i/).	64
Quadro 5 – Amplitudes de H1, H2, H3, H4 e do pico na região do formante do cantor das vogais cantadas (/o/ e /u/).	65
Quadro 6 – NPS das vogais /a/ e /e/ faladas e cantadas.	67
Quadro 7 – NPS das vogais /i/, /o/ e /u/ faladas e cantadas	68
Quadro 8 – Sobreposição dos EMLT da fala espontânea e da execução da ária.	70

LISTA DE TABELAS

Tabela 1 – Identificação dos participantes.	59
Tabela 2 – Comentários dos sujeitos sobre as notas agudas.	60
Tabela 3 – Número de ocorrências de sintonia de formantes por vogal.	66
Tabela 4 – Número de ocorrências de sintonia de formantes por sujeito.	66
Tabela 5 – Número de ocorrências de sintonia de formantes por região de f0.	66
Tabela 6 – Declínio espectral da fala espontânea e da execução da ária.	69

SUMÁRIO

1. Introdução.....	1
2. Objetivo	6
2.1. Objetivos específicos	6
3. Revisão da literatura	7
3.1. A fonte e o filtro na acústica vocal	7
3.2. Interações fonte-filtro e sintonia de formantes	11
3.4. As notas agudas e a qualidade vocal na pesquisa em voz cantada	21
4. Método.....	45
4.1. Preceitos éticos	45
4.2. Casuística	45
4.3. Seleção da ária e das vogais	45
4.4. Instrumentos	47
4.5. Procedimentos	47
4.6. Análise dos dados	49
5. Resultados	59
6. Discussão	71
7. Conclusão	83
8. Referências Bibliográficas	84
Anexos	89

1. Introdução

Na minha experiência como professor de canto, percebo grande dificuldade com notas agudas no desenvolvimento dos cantores. Observo o mesmo em minha história como aluno de canto lírico na graduação, e no estudo de técnicas de canto contemporâneo. Escuto frequentemente o cantor iniciante que, ao ultrapassar notas confortáveis, recorre a compensações como, por exemplo, elevação do queixo, projeção da cabeça para frente e movimentos da língua. A voz pode adquirir uma qualidade tensa, por vezes “estridente” e sem “plasticidade”, o que limita o cantor nas suas expressividades vocal e musical. As notas mais agudas me imprimem ser um maior desafio para homens, possivelmente por se situarem em uma região de frequências fundamentais (f_0) em que são necessários ajustes fonoarticulatórios bem distintos dos da sua fala habitual.

A qualidade vocal e as dificuldades de produção das notas agudas estão, em parte, relacionadas aos registros vocais. Neumann et al. (2005) apontaram a falta de consenso sobre a classificação dos registros, historicamente tratados pela percepção do ouvinte, pela do cantor, ou pelo modo de produção do som. Segundo Roubeau et al. (2009), os registros foram descritos por profissionais de áreas diversas (fisiologistas, médicos, foneticistas, cientistas vocais, professores de canto e cantores), com diferentes interesses na voz, e portanto, diferentes abordagens. Tais descrições foram baseadas em observações laringoscópicas, eletrofisiológicas, acústicas, auditivas ou proprioceptivas. Para Titze (2000), o uso de terminologias diferentes para os mesmos fenômenos, e de terminologias iguais para fenômenos diferentes, é fonte de grande debate entre cientistas da fala e professores de canto¹.

No século XIX, o cantor, pedagogo vocal e anatomista espanhol Manuel Garcia II (1805-1906) foi o primeiro a observar a laringe durante o canto com um laringoscópio (Lamesch et al., 2007). Ele foi pioneiro em definir dois² registros vocais – peito (*poitrine*) e falsete-cabeça (*fausset-tête*) – em função de um princípio mecânico, que

¹ Para uma revisão histórica das terminologias utilizadas nas ciências vocais para os registros, conferir Roubeau et al. (2009).

² Garcia designou ainda um terceiro registro, contra-baixo (*contre-basse*), caracterizado por sons “ásperos” e correspondente às notas mais graves possíveis, porém utilizado apenas por alguns baixos-profundos russos, e portanto considerado pelo autor sem aplicação à arte do cantor (Garcia, 1847).

prevalece sobre os aspectos perceptivos (Henrich, 2006). Garcia define um registro como “uma série de sons consecutivos e homogêneos [...] produzidos pelo desenvolvimento do mesmo princípio mecânico” que difere de outros sons “igualmente consecutivos e homogêneos, produzidos por um outro princípio mecânico”³ (Garcia, 1847, p. 4). A natureza dos sons de um registro é independente de modificações de timbre ou intensidade às quais eles possam ser submetidos.

Henrich (2006) relata que, de então até meados do século XX, acreditava-se haver entre dois e cinco registros. A partir da década de 1980, a existência de ao menos dois mecanismos vibratórios laríngeos foi experimentalmente validada. Segundo a autora, a comunidade científica da voz falada chegou ao acordo implícito da existência de três registros, baseados em ajustes laríngeos específicos: pulso ou *vocal fry*, modal ou peito, e falsete. A comunidade da voz cantada, por outro lado, dada a importância da percepção para a prática do cantor, seguiu no debate de definições, número e nomes dos registros. Miller descreve um registro como uma “série de sons vocais consecutivos, de timbre igual ou similar, que pode ser distinta de uma série de sons adjacentes”⁴ (Miller, 1993, p. 158). Titze relata que o termo registro tem sido utilizado para descrever “regiões perceptualmente distintas de qualidade vocal que pode ser mantida por uma série de *pitch* e *loudness*”⁵ (Titze, 2000, p. 282). Nota-se que ambos incluem a dimensão perceptiva nas suas descrições.

Roubeau et al. (2009) descreveram quatro mecanismos vibratórios laríngeos – M0, M1, M2 e M3, do grave ao agudo – envolvidos respectivamente nos registros *vocal fry*, peito, falsete, e apito, entre outros termos historicamente utilizados. M1 e M2 são os mais frequentes na fala e no canto, e há uma faixa de frequências

³ Tradução livre do pesquisador, do original em francês : “[...] *une série de sons consécutifs et homogènes allant du grave à l'aigu, produits par le développement du même principe mécanique, et dont la nature diffère essentiellement d'une autre série de sons également consécutifs et homogènes, produits par un autre principe mécanique. Tous les sons appartenant au même registre sont, par conséquent, de la même nature, quelles que soient d'ailleurs les modifications de timbre ou de force qu'on leur fasse subir*” (Garcia, 1847, p. 4).

⁴ Tradução livre do pesquisador, do original em inglês: “A series of consecutive voice tones of equal or similar timbres, which can be distinguished from an adjoining series of tones” (Miller, 1993, p. 158).

⁵ Tradução livre do pesquisador, do original em inglês: “The term register has been used to describe perceptually distinct regions of vocal quality that can be maintained over some ranges of pitch and loudness” (Titze, 2000, p. 282).

intermediárias em que ambos são empregáveis. Os autores concluem que as concepções de registros utilizadas por cantores, baseadas em qualidades acústicas e sensações proprioceptivas, se justapõem às noções de mecanismos laríngeos. As fronteiras entre registros não correspondem necessariamente às aquelas observadas entre os mecanismos, de modo que dois registros que diferem nos aspectos acústicos e sensoriais podem ser produzidos por um mesmo mecanismo laríngeo.

Autores recentes da pedagogia vocal têm dado maior destaque a eventos acústicos para descrever os registros. Miller (2000) e Bozeman (2013) utilizam relações entre harmônicos e formantes⁶, e seu impacto na qualidade vocal, em suas classificações. Em consonância com Roubeau et al. (2009), ambos consideram a existência de padrões de vibração básicos advindos das configurações glóticas, e que os aspectos acústicos dados pela forma do trato vocal se somam a estes e resultam na sonoridade final dos registros. Neste trabalho, será adotada essa concepção, de que os registros têm um componente derivado do padrão vibratório laríngeo, que produz um sinal acústico chamado de fonte glótica, e outro determinado pela configuração do trato vocal, que atua como um filtro do sinal da fonte.

As quebras de registro, ou quebras vocais, são uma problemática do treinamento vocal crucial para as notas agudas (Miller, 2000). Caracterizadas por mudanças súbitas em f_0 e na qualidade vocal, são oriundas de alterações bruscas no comportamento laríngeo e nos sistemas de ressonância acima e abaixo das pregas vocais (Miller, 1993), e podem ou não ser intencionais (Sundberg, 2018). De grande relevância para o canto, são as quebras que tendem a ocorrer na transição entre os registros de peito e de falsete (Miller, 2000). Titze et al. (1994) referem que, mais do que outras classificações vocais⁷, tenores são notórios por quebras involuntárias em notas agudas, o que gera no público grande expectativa. O risco de instabilidade em tais notas vem da alta tensão nas pregas vocais e pressão pulmonar da sua produção,

⁶ Os conceitos da acústica citados nesse parágrafo – harmônicos, formantes, fonte e filtro – serão melhormente discutidos na revisão de literatura deste trabalho (Titze, 2008a; Sundberg, 2018)

⁷ As vozes são classificadas de acordo com sua extensão média e com seu timbre. Da mais grave à mais aguda, as vozes masculinas são divididas em baixo, barítono e tenor; e as vozes femininas em contralto, mezzo-soprano e soprano (Bennett, 1998).

o que as faz soarem vigorosas. Para os autores, reverter a tendência natural da troca para o falsete é o grande objetivo da pedagogia vocal para a voz masculina aguda.

Segundo Miller (2019), enquanto o canto na região da fala não impõe grandes desafios ao cantor não treinado, cresce nas emissões agudas a tendência à elevação laríngea, com eventual quebra para o registro de falsete, qualidade vocal indesejada no canto lírico masculino. No aspecto fisiológico, o autor destaca o controle equilibrado entre a ação dos músculos tireoaritenóideos (TA) e cricotireóideos (CT) para garantir que as mudanças de qualidade sonora sejam graduais por toda a extensão vocal.

Os aspectos acústicos igualmente apresentam desafios às notas agudas. O trato vocal funciona como um filtro, que amplifica os componentes do som gerado pela vibração das pregas vocais, os harmônicos da fonte glótica. Para tanto, os articuladores (mandíbula, língua, lábios, palato mole, faringe e laringe) precisam estar bem posicionados, para que as ondas sonoras vindas da glote “se encaixem” na configuração espacial do trato vocal. Esta tarefa é mais difícil em notas agudas, pois os harmônicos se tornam mais distantes, e precisam de ajustes finos para que sejam amplificados (Bozeman, 2013). Ao considerar ainda interações não-lineares entre fonte e filtro (Titze, 2008a), relações entre as ondas que ressoam no trato vocal e a vibração das pregas vocais podem auxiliar ou causar instabilidades na fonação. As condições para tais interações ocorrerem são mais prováveis em f_0 mais agudas.

As adaptações fisiológicas e ressonantes necessárias para manter a oscilação auto-sustentada das pregas vocais nas notas agudas têm direta influência sobre a qualidade vocal e, logo, sobre a percepção do timbre pelo ouvinte (Titze e Worley, 2009). Fica claro o desafio que a emissão dessas notas impõe para o cantor: ele deve, simultaneamente, manejar as dinâmicas musculares laríngeas; considerar possíveis interações entre o trato vocal supraglótico e a oscilação das pregas vocais; e evitar que tais ajustes resultem em uma qualidade vocal inadequada ao seu gênero musical.

O presente estudo pretende expandir a compreensão sobre as características acústicas das emissões cantadas em altas frequências fundamentais, com foco nas notas mais agudas da tessitura de tenores de ópera. Os achados desta pesquisa, além de contribuírem para os conhecimentos científicos acerca da voz cantada, podem ter aplicação prática na pedagogia vocal, na clínica fonoaudiológica e no

processo de estudo de cantores. Atualmente, estão disponíveis ao público geral ferramentas de espectrografia pagas e gratuitas, de *softwares* sofisticados a *apps* de simples utilização para celular. Esses podem informar fonoaudiólogos e professores de canto, de forma não invasiva e em tempo real, a respeito da fonação e dos ajustes de trato vocal, além de poderem ser utilizados por cantores e alunos de canto como ferramenta de *feedback* visual para treino. Aliado a esses dispositivos, o conhecimento sobre as características acústicas e articulatórias demandadas pelos diversos gêneros de canto pode possibilitar maior precisão e eficiência na intervenção em voz cantada, bem como no processo do desenvolvimento vocal, artístico e estético dos cantores.

2. Objetivo

Descrever acusticamente os ajustes do trato vocal supraglótico e as características do espectro radiado nas emissões das notas mais agudas da tessitura de tenores eruditos.

2.1. Objetivos específicos

2.1.1. Identificar, nas emissões de tenores eruditos, as variações nas frequências do primeiro e do segundo formantes, das notas graves às notas agudas da sua tessitura, nas vogais /a/, /e/, /i/, /o/ e /u/.

2.1.2. Identificar, no espectro acústico radiado de tenores eruditos, as variações nos harmônicos de maior amplitude, das notas graves às notas agudas da sua tessitura, nas vogais /a/, /e/, /i/, /o/ e /u/.

2.1.3. Comparar a distribuição de energia nas regiões de frequência do espectro médio de longo termo na voz falada e na voz cantada de tenores eruditos.

3. Revisão da Literatura

Este capítulo está dividido em três seções. Na primeira, são abordadas as características fisiológicas e acústicas da produção vocal, com foco no papel da fonte e do filtro no espectro acústico radiado. São explicados termos e conceitos da acústica aplicados à produção vocal – em especial harmônicos, ressonância e formantes. A segunda seção trata das influências que as características do filtro podem ter sobre o funcionamento da fonte, de acordo com a teoria de acoplamento não-linear fonte-filtro na fonação, além das técnicas de ajustes de formantes utilizadas por cantores. A terceira parte reúne pesquisas realizadas com foco em questões subjetivas, fisiológicas e acústicas de notas agudas e da qualidade sonora na voz cantada. Nas duas primeiras seções, é feita uma revisão teórica acerca dos tópicos, enquanto na terceira seção são apresentados estudos com sujeitos humanos ou simulações computacionais. Estes estão agrupados por temáticas e, em cada temática, apresentados em ordem cronológica, na sequência: aspectos proprioceptivos dos cantores; aspectos fisiológicos das notas agudas; e aspectos acústicos e articulatórios das notas agudas e das diferentes qualidades vocais no canto.

3.1. A fonte e o filtro na acústica vocal

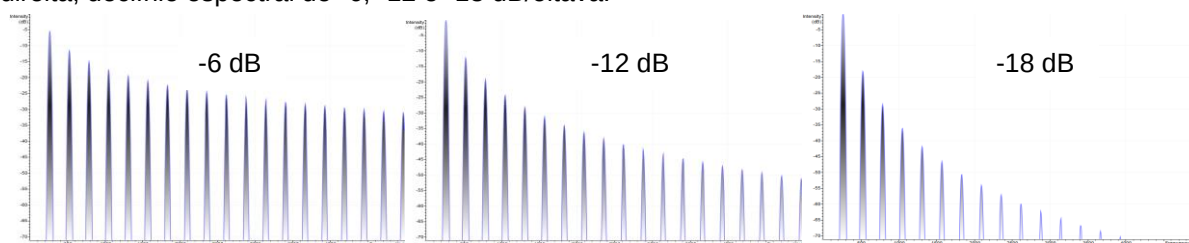
Sundberg (2018) descreve o sistema fonador como sendo composto de três partes: o sistema respiratório, as pregas vocais e as cavidades de ressonância. O fluxo gerado pela compressão do ar pelos pulmões (sistema respiratório) faz com que as pregas vocais vibrem, produzindo um som – a fonte glótica –, que é radiado pelo trato vocal (cavidades de ressonância). Este exerce a função de filtro acústico, ou seja, transforma as características do espectro da fonte (Sundberg, 2018).

A fonte glótica é o som produzido pela vibração das pregas vocais, quando entre elas passa o fluxo de ar durante a fonação. Essa vibração causa uma alternância entre a passagem e a interrupção desse fluxo, que o transforma numa série de pulsos de ar. Se a vibração for regular, periódica, será produzido um som com uma frequência definida, correspondente ao número de vezes que as pregas fecham e abrem a glote por segundo (Sundberg, 2018). Essa é a frequência fundamental (f_0), relacionada à altura ou ao *pitch* percebido auditivamente. O som da fonte glótica não é constituído apenas pela f_0 , mas é um espectro que pode ser decomposto em sons parciais. Se o

sinal acústico da fonte for periódico, ele será constituído de uma série harmônica, em que os parciais são múltiplos inteiros da f_0 ⁸, e eles serão chamados de harmônicos. O primeiro harmônico (H1) tem a mesma frequência de f_0 , e os harmônicos seguintes são numerados em ordem crescente: H2, H3, ..., Hn (Sundberg, 2018).

O espectro da fonte glótica apresenta um declínio espectral, um decréscimo sucessivo na amplitude de cada harmônico, à medida que se tornam mais agudos. O declínio espectral é caracterizado por uma perda regular de intensidade entre 6 e 18 dB por oitava, ou seja, a cada dobramento da frequência. O declínio espectral está relacionado à qualidade vocal: quanto menor a perda de energia por oitava, maior o número de frequências agudas presentes no espectro, causando uma qualidade sonora percebida como metálica (*brassy*); quanto maior a perda, menor o número de frequências agudas, e o som é percebido como semelhante ao de uma flauta (*fluty*) (Titze, 2000). A Figura 1 apresenta três exemplos de espectro da fonte glótica, com declínio espectral de -6 dB, -12 dB e -18 dB, sintetizados no gerador de som do software *VoceVista Video* versão 5.6.1.6330 (Sygyt Software, Alemanha).

Figura 1 – Espectros de fonte glótica sintetizada no software *VoceVista Video*. Da esquerda para a direita, declínio espectral de -6, -12 e -18 dB/oitava.



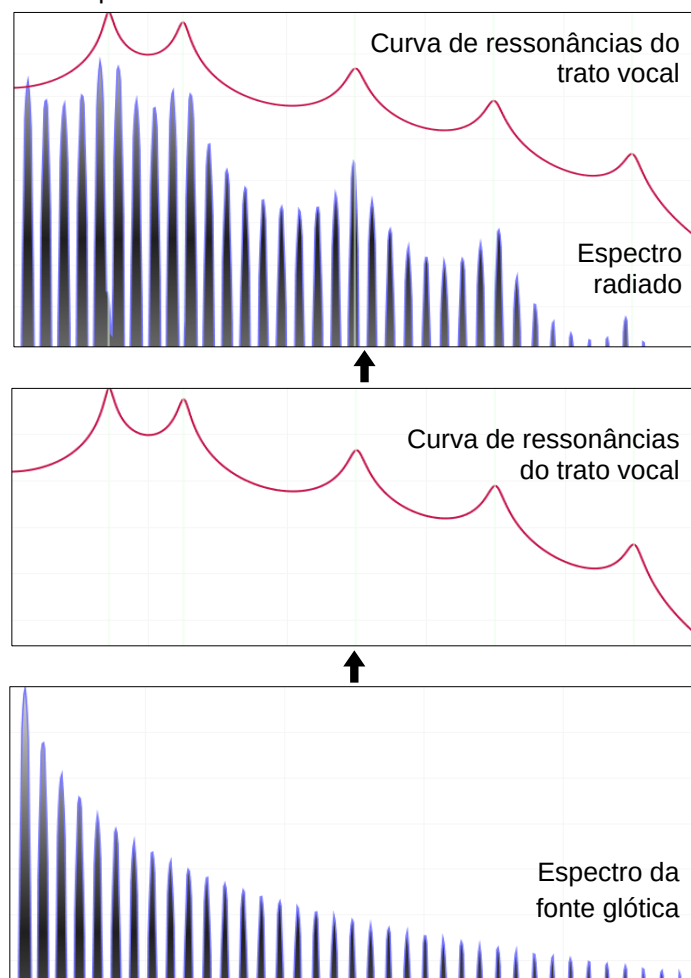
Fonte: Autor. Adaptado de Titze (2000).

Diferentemente das pregas vocais, que produzem um sinal acústico, o trato vocal supraglótico apenas filtra o som, o modifica (Titze, 2000), funcionando como um ressoador, ou seja, tem a propriedade de favorecer a transmissão de determinadas frequências e atenuar outras. As frequências favorecidas por um ressoador são chamadas frequências de ressonância, mais especificamente no caso do trato vocal, frequências de formantes (Sundberg, 2018). As frequências dos formantes são determinadas pela configuração do trato vocal, por meio do posicionamento dos

⁸ Por exemplo, se f_0 for 100 Hz, os parciais harmônicos serão 200 Hz, 300 Hz, 400 Hz, 500 Hz, ...; se f_0 for 300 Hz, os harmônicos serão 600 Hz, 900 Hz, 1200 Hz, 1500 Hz etc.

articuladores (lábios, língua, mandíbula, véu palatino, faringe e laringe) e das modificações que eles causam na seção transversal ao longo comprimento do trato vocal. Os harmônicos da fonte glótica que tiverem frequências iguais ou próximas às dos formantes serão radiados pelos lábios com maior intensidade (Sundberg, 2018). Dessa forma, o espectro da fonte é modificado pelo filtro, adquirindo um envelope com picos e vales no espectro radiado, que variam de acordo com as frequências de ressonância e com seu distanciamento em relação aos harmônicos. A Figura 2 exemplifica essa relação, em som sintetizado no *VoceVista Video*. A curva vermelha no gráfico ao centro (função de transferência do trato vocal), é determinada pelos formantes, que correspondem aos picos na curva (Sundberg, 2018). O perfil do gráfico muda de acordo com as configurações do trato vocal, ou seja, com as variações nas vogais (Titze et al., 2017). Quanto mais próximo um harmônico estiver do centro do pico da curva de ressonâncias, maior será sua amplitude no espectro radiado.

Figura 2 – Exemplo do efeito do filtro sobre o espectro da fonte glótica, em som sintetizado no software *VoceVista Video*. De baixo para cima: espectro da fonte glótica; ressonâncias do trato vocal; espectro radiado e ressonâncias sobrepostas.

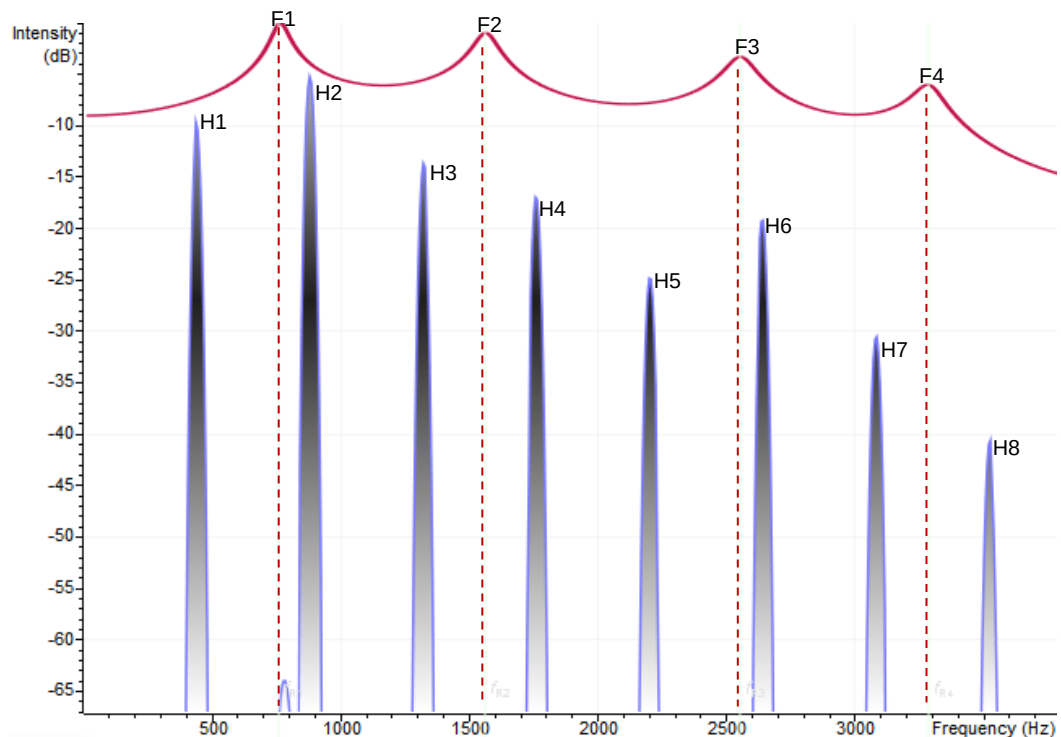


Fonte: Autor. Adaptado de Sundberg (2018).

Titze et al. (2015) relataram a falta de padronização dos termos e de suas notações na pesquisa vocal⁹. Dado que a voz é objeto de diferentes áreas do conhecimento, cada uma destas expressa os conceitos de acordo com seus pontos de vista e tradições, com suas preferências de notação. O termo formante foi utilizado para descrever tanto os picos no espectro radiado (os harmônicos de maior amplitude), quanto para se referir aos picos da curva de ressonância do filtro (as frequências favorecidas pelo trato vocal).

A Figura 3 ilustra a diferença entre os picos no espectro radiado e as ressonâncias. Considerando que as frequências de ressonância dependem exclusivamente da configuração do trato vocal, e que as frequências dos harmônicos são múltiplos inteiros da f_0 , sendo resultado somente dela, não necessariamente ocorrerão coincidências entre os harmônicos de maior amplitude e as frequências centrais das ressonâncias.

Figura 3 – Diferenciação entre os picos do espectro e os das curvas de ressonância, em som sintetizado no *software VoceVista Video*: é possível ver no espectro radiado (linhas cinzas verticais) que os harmônicos de maior amplitude não coincidem exatamente com os picos na curva de ressonância (curva no topo em vermelho). Linhas verticais tracejadas vermelhas mostram onde os harmônicos deveriam estar para estarem alinhados ao centro de frequência dos picos de ressonância.



Fonte: Autor. Baseado em Bozeman (2013) e Howell (2016).

⁹ Conferir Titze et al. (2015) para as sugestões dos autores acerca de uma padronização das notações das terminologias na pesquisa em acústica vocal.

No presente trabalho, os formantes serão considerados como a frequência central de cada ressonância do trato vocal. Assim como os harmônicos, serão numerados em ordem crescente de frequência: F_1 , F_2 , ..., F_n .

3.2. Interações fonte-filtro e sintonia de formantes

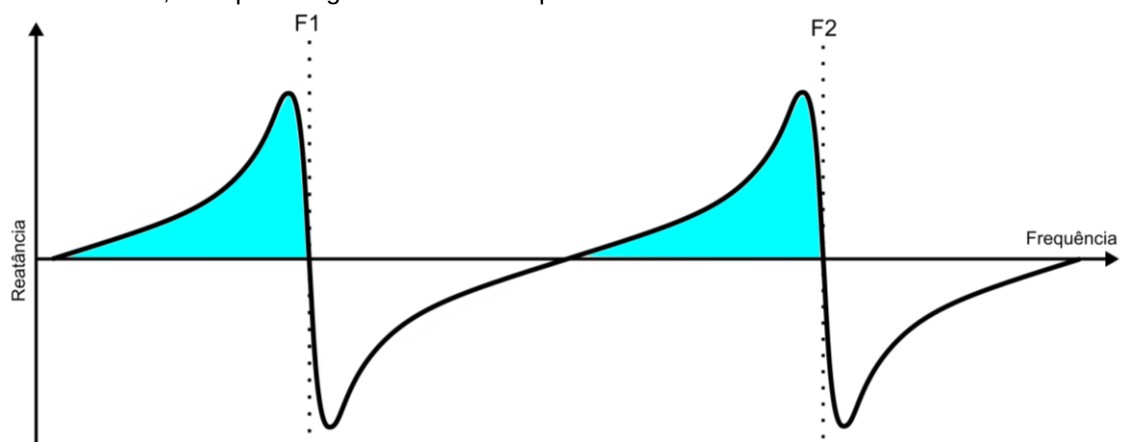
Na teoria de acoplamento não-linear fonte-filtro na fonação, Titze (2008a) explica como as configurações do trato vocal supralaríngeo e os comportamentos glóticos se influenciam. Um conceito da física essencial, porém de difícil compreensão, para explicar a não-linearidade das relações entre fonte e filtro, é a impedância acústica. Ela é uma relação entre a pressão acústica e as oscilações no fluxo de ar, sobre uma determinada área. Simplificadamente, pode-se dizer que a impedância mede a dificuldade de se produzir um fluxo oscilatório de ar através de uma dada área, com uma determinada frequência. A impedância varia com a frequência e, matematicamente, é um número complexo, composto de uma parte real e uma parte imaginária. A parte real é chamada de resistência, enquanto a parte imaginária é chamada de reatância. A resistência se relaciona às perdas de energia sonora, normalmente em forma de calor, enquanto a reatância está ligada ao armazenamento de energia acústica (Wolfe et al., 2009).

Story et al. (2000) fizeram uma analogia prática, que facilita um pouco a compreensão desses conceitos. Pensando em uma geladeira sendo empurrada sobre um chão de concreto, uma pessoa teria grande dificuldade, pois a maior parte da energia que ela utilizasse para mover a geladeira seria perdida em forma de calor, por causa do atrito desta com o chão. Esse exemplo ilustra um sistema em que a resistência é alta, havendo grande perda de energia. Porém, se fossem colocadas rodinhas embaixo da geladeira, seria necessário um maior esforço apenas no primeiro empurrão pra iniciar seu movimento, que continuaria em função da inércia. Esse é o exemplo de um sistema em que a reatância é alta: parte da energia introduzida nele é temporariamente armazenada como energia cinética, que é utilizada para auxiliar o sistema a manter-se em movimento. Valores positivos de reatância, nos quais há armazenamento de energia, caracterizam a reatância inercial, assim chamada por sua relação com a inércia do sistema.

A reatância inertiva é de especial importância para a produção da voz. Se a impedância do trato vocal for caracterizada pela reatância inertiva, parte da energia acústica introduzida nele pela oscilação das pregas vocais será armazenada na coluna de ar do trato vocal. Essa energia pode ser retroalimentada para as pregas vocais, o que auxilia na sua oscilação e diminui o esforço necessário para manter a fonação, e assim aumenta a eficiência na produção de um *output* sonoro de maior intensidade (Titze, 2008b).

Titze e Worley (2009) e Titze et al. (2011) demonstraram que os valores da reatância do trato vocal variam com sua configuração espacial, sendo determinados principalmente pelas frequências das ressonâncias (formantes) do trato vocal. Os valores de reatância são maiores (reatância inertiva) nas frequências logo abaixo dos formantes. Nas frequências imediatamente acima dos formantes, os valores da reatância caem, podendo tornar-se negativos, e voltam a subir gradativamente até atingir um novo ponto máximo, logo antes da próxima frequência de ressonância, e assim sucessivamente. Essa variação é esquematizada na Figura 4.

Figura 4 – Esquematização da reatância no trato vocal. Áreas em azul correspondem às regiões de reatância inertiva, com picos logo abaixo das frequências dos formantes.



Fonte: Autor. Adaptado de Story et al. (2000) e de Titze e Worley (2009).

A reatância inertiva no trato vocal é benéfica para a maioria das frequências de fonação humana. Os harmônicos que estiverem nas faixas de frequência de alta reatância inertiva serão mais amplificados, e quanto mais deles coincidirem com essas regiões, maior será a geração de energia acústica. Por outro lado, os harmônicos que estiverem em regiões de reatância negativa serão atenuados (Titze, 2008a). Para uma compreensão mais prática de como se dão esses fenômenos na produção da voz, será explorado o comportamento do ar dentro do trato vocal.

O trato vocal funciona como um tubo preenchido por um volume de ar, que possui massa e elasticidade, o que permite que ocorram nele oscilações mecânicas. São essas propriedades que o fazem funcionar como um ressoador. “Algo que possua massa apresenta a propriedade de, uma vez em movimento, tender a se manter em movimento (primeira lei de Newton)” e “algo que possua elasticidade apresenta a propriedade de ser comprimido e, uma vez comprimido, tender a retornar a seu volume original” (Sundberg, 2018, p. 32).

Titze (2008b) explica que, quando as pregas vocais se separam no início de uma oscilação, o ar que flui através da glote encontra a massa de ar imóvel do espaço supraglótico. Se a reatância do trato vocal for inertiva, a massa de ar, por causa de sua inércia (sua tendência a manter seu estado atual de movimento), permanece inicialmente imóvel, até que, com o aumento da pressão abaixo dela, vinda dos pulmões, ela é acelerada. As forças elásticas das pregas vocais as fazem se aproximar novamente, cortando o fluxo de ar vindo da traqueia. A coluna de ar supraglótica, mais uma vez por causa de sua inércia, tende a continuar a se mover, deixando atrás de si um vácuo parcial, na glote e logo acima dela, o que faz com que a glote se feche mais fortemente. Há assim um atraso entre o movimento da coluna de ar e os movimentos das pregas vocais, pois a massa de ar no trato vocal demora a responder às pressões sobre ela aplicada. Essa dinâmica auxilia na manutenção da oscilação das pregas vocais, aumentando a amplitude dos movimentos vibratórios e criando ondas sonoras mais fortes.

Existem dois níveis de interação não-linear fonte-filtro. A interação de Nível 1 ocorre entre o fluxo de ar transglótico e as pressões acústicas do trato vocal, sem perturbar a vibração das pregas vocais, e está presente na fala e no canto do homem e da mulher. Quando o trato vocal apresentar alta reatância inertiva, ocorrerá na glote o aumento da taxa máxima do declínio de fluxo (TMDF)¹⁰. A tendência da coluna de ar supraglótica a continuar a se mover, após a interrupção do fluxo de ar transglótico pelo início do fechamento das pregas vocais, faz com que elas se aproximem mais

¹⁰ A taxa máxima do declínio de fluxo (*maximum flow declination rate*, ou *MFDR*) é um parâmetro obtido da derivada do glotograma de fluxo. Ela corresponde à velocidade com que o fluxo de ar diminui ao final de cada ciclo glótico, ou seja, está relacionada à velocidade com que as pregas vocais fecham a cada ciclo. O fechamento da glote é o momento de maior excitação do trato vocal: quanto mais rápido ele ocorrer, mais energia será gerada pelas pregas vocais (Salomão e Sundberg, 2009).

forte e rapidamente. Uma maior TMDf está relacionada a um declínio espectral mais plano, e logo ao aumento da energia acústica por todo o espectro, com produção de novos harmônicos na fonte. Dessa forma, há um incremento da intensidade vocal, além de mudanças na qualidade vocal, uma vez que o espectro foi modificado. Para que ocorram interações do Nível 1, não é necessário que as frequências dos harmônicos coincidam com os formantes, basta estarem situadas nas regiões de reatância positiva abaixo destes. Harmônicos individuais podem ser amplificados ou atenuados, a depender se estiverem em regiões de reatância positiva ou negativa, respectivamente (Titze, 2008a).

A interação de Nível 2 ocorre mais em f_0 agudas, quando os harmônicos mais graves (especialmente H1 e H2) estão próximos aos formantes e encontram mudanças súbitas de reatância. Isso ocorre principalmente quando H1 passa por F1, ou seja, sobe de uma frequência logo abaixo F1 (reatância inercial) para uma frequência logo acima (reatância negativa). Esse tipo de interação também pode ocorrer quando H2 cruza F1. Nas interações de Nível 2, há uma interferência no próprio comportamento oscilatório das pregas vocais. Os efeitos são saltos de f_0 , produção de sub-harmônicos e ruído, e mudanças de padrão de vibração das pregas vocais, com trocas do registro modal para o registro de falsete (Titze, 2008a). Ressonâncias subglóticas, ou seja, na cavidade abaixo das pregas vocais, na traqueia, também contribuem para a interação de Nível 2, mas estão fora do escopo desse trabalho e não serão aqui discutidas.

Para que o cantor se beneficie da impedância do trato vocal, é necessário que os harmônicos da nota que ele emitir estejam dentro das zonas de frequência em que a reatância é inercial, ou seja, próximos, porém abaixo, das frequências dos formantes. Diferentes configurações do trato vocal e diferentes vogais apresentarão variações nas regiões de reatância inercial, o que significa que o cantor pode necessitar alterar a vogal que vai cantar, para que os formantes se aproximem dos harmônicos da nota cantada. Com efeito, diversos gêneros de canto, como ópera, teatro musical e *yodeling*, tem a concepção de que certas vogais ou qualidades vocais são mais adequadas para determinadas notas (Titze, 2008a). Como o cantor não pode alterar os harmônicos, que são determinados pela nota que deve ser cantada, o principal fator que pode ser alterado é a localização das formantes (Bozeman, 2013).

Fazer com que os harmônicos se situem em regiões de reatância inertiva torna-se uma tarefa progressivamente mais difícil à medida que f_0 aumenta. Os harmônicos, por serem múltiplos inteiros de f_0 , ficam cada vez mais espaçados (Bozeman, 2013). Exemplificando, se a nota entoada for A2¹¹, f_0 será 220 Hz, logo H2=440 Hz, H3=660 Hz, H4=880 Hz, H5=1100 Hz, ..., H10=2200 Hz etc., e haverá dez harmônicos entre 220-2200 Hz. Subindo a nota uma oitava para A3, ou seja, dobrando f_0 pra 440 Hz, a sequência será H2=880 Hz, H3=1320 Hz, H4=1760 Hz, H5=2200 Hz etc. Na mesma região de frequência, entre 220-2200 Hz, haverá apenas cinco harmônicos.

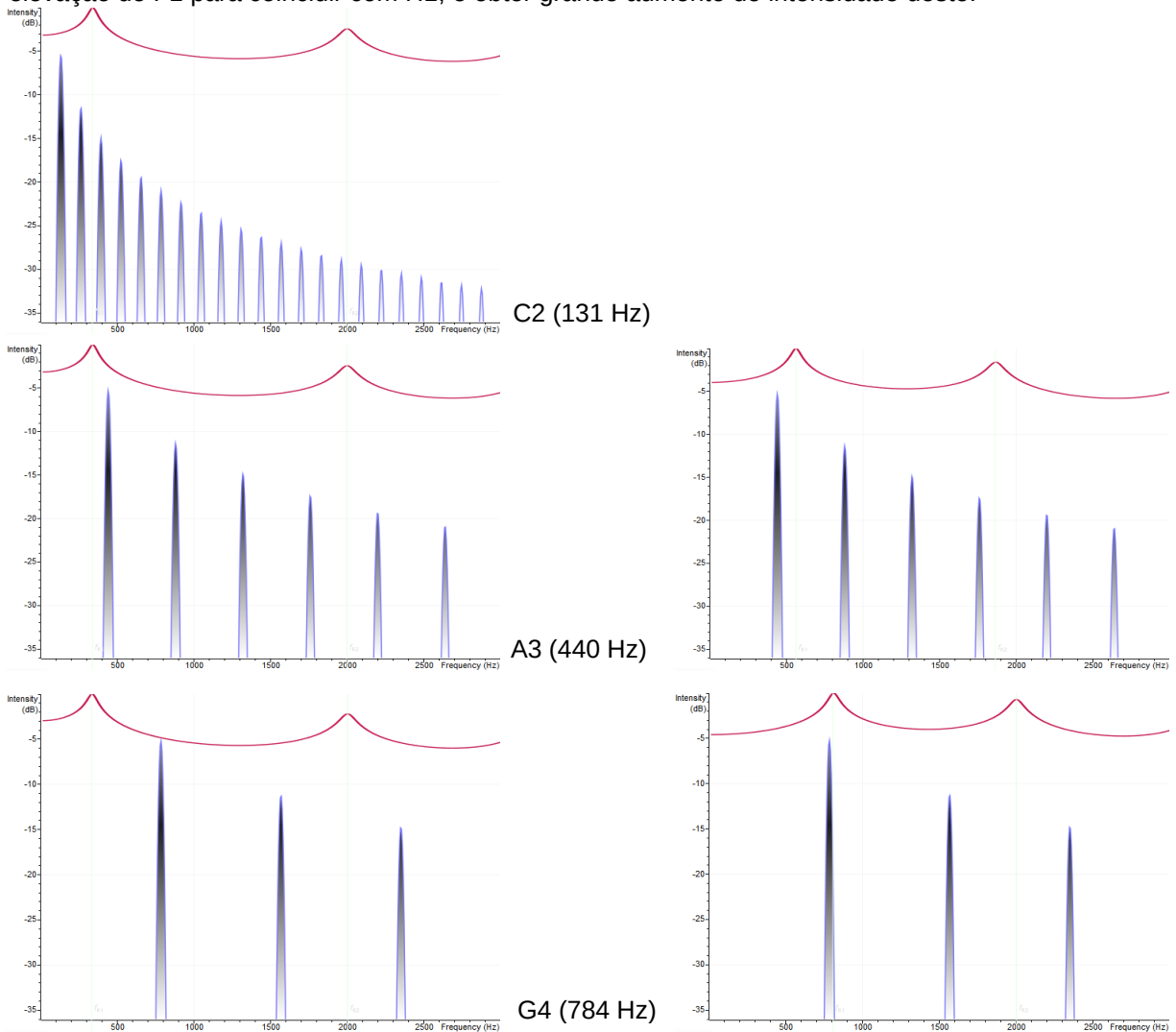
Segundo Sundberg (2018), as frequências de ressonância de um homem adulto variam entre 150-1000 Hz para F1 e de 600-2500 Hz para F2, dependendo da vogal. Considerado o exemplo acima, a nota A2 tem até cinco harmônicos que poderiam se beneficiar da reatância inertiva de F1, e oito que poderiam se beneficiar de F2, enquanto a nota A4 teria no máximo dois na região positiva de F1 e quatro na de F2, ao levar em conta os valores máximos possíveis para F1 e F2. Assim, ao emitir notas mais graves, o cantor não precisa se preocupar com as ressonâncias, pois há diversos harmônicos, próximos um dos outros, para serem ressoados por F1, F2 e pelos formantes mais altos. Porém, em notas mais agudas, o vocalista precisará de mais atenção às modificações das ressonâncias, para poder selecionar os harmônicos que serão amplificados, uma vez que estes estão mais espaçados (Bozeman, 2013).

A Figura 5 ilustra essa relação. Nas imagens do lado esquerdo, F1 e F2, correspondentes aproximadamente à vogal /e/, foram mantidos iguais, como pode ser visto pela localização dos dois picos na curva vermelha no topo de cada espectro. De cima para baixo, estão representados os espectros de séries harmônicas com f_0 nas notas C2 (131 Hz), A3 (440 Hz) e G4 (784 Hz). Fica evidente como os harmônicos ficam mais espaçados, quanto mais aguda for a nota. Apenas a nota C2 apresenta harmônicos na região de reatância inertiva abaixo de F1, e muitos harmônicos

¹¹ A escala musical utilizada no ocidente é composta de sete notas, Dó, Ré, Mi, Fá, Sol, Lá e Si, com respectiva notação alfabética C, D, E, F, G, A e B. A letra b (bemol) após a nota indica uma nota intermediária meio tom abaixo; o sinal # (sustenido), uma nota intermediária meio tom acima. A nota Dó central do piano, 261,63 Hz, será considerada como C3. Ao dividir sua frequência pela metade para 130,81 Hz, obtém-se a nota uma oitava abaixo, C2; ao dobrá-la para 523,25 Hz, a nota uma oitava acima, C4. O mesmo se aplica às outras notas (Med, 1996). No presente trabalho será utilizada a notação alfabética, sempre seguida da respectiva frequência, arredondada para um número inteiro.

próximos na região favorável de F2. As notas A3 e G4 não têm nenhum harmônico bem posicionado em relação aos formantes.

Figura 5 – Diferentes relações entre formantes e harmônicos em função de f_0 : à esquerda, a mesma função de transferência do trato vocal (curva vermelha) é comparada a séries harmônicas nas notas C2 (131 Hz), A3 (440 Hz) e G4 (784 Hz), ilustrando o aumento do espaçamento dos harmônicos e o mau posicionamento deles em relação aos formantes nas notas agudas. À direita, soluções possíveis de alteração dos formantes para se adequarem aos harmônicos: na nota A3, elevação de F1 para que H1 esteja na sua região de reatância inertiva, e diminuição de F2, para amplificar H4; na nota G4, elevação de F1 para coincidir com H1, e obter grande aumento de intensidade deste.

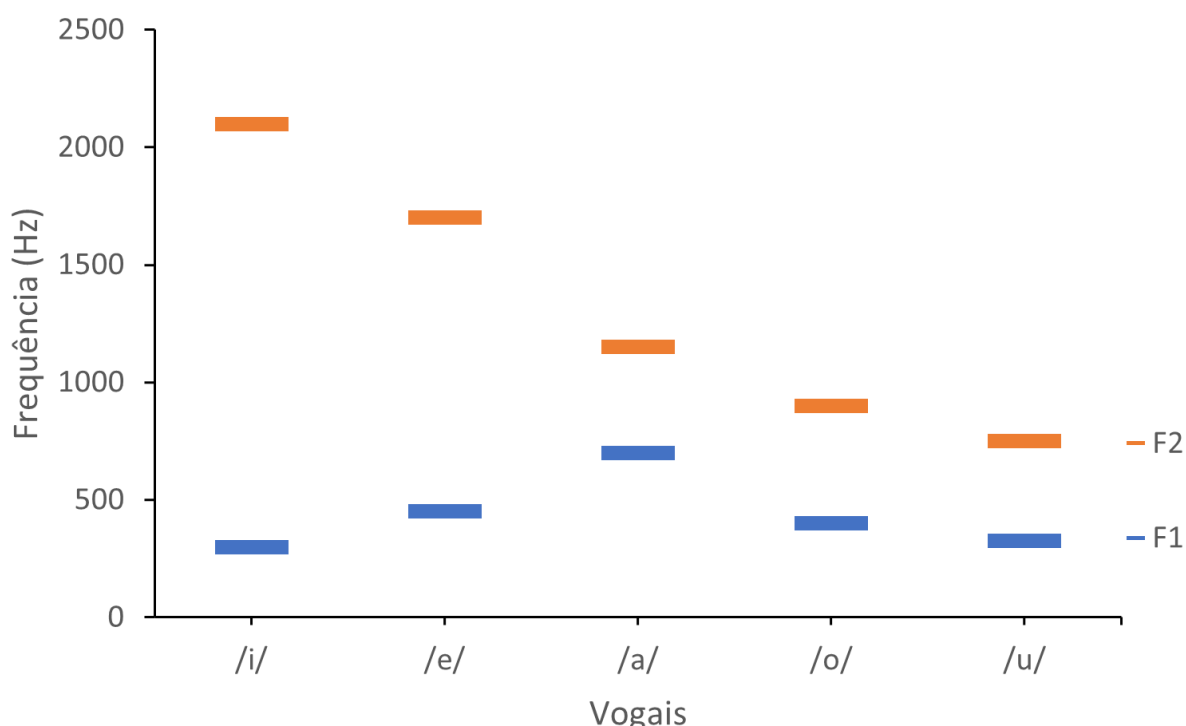


Fonte: Autor. Baseado em Howell (2016).

É importante salientar que alterar as frequências de F1 e F2 pode ter como consequência a modificação da vogal percebida. A percepção das diferentes vogais é definida principalmente pelas frequências dos dois primeiros formantes (Figura 6). Os formantes F3, F4, e F5 têm maior relevância na determinação das qualidades individuais da voz (Sundberg, 2018). Os diferentes gêneros de canto podem ser diferenciados pelo timbre, seja ele uma escolha estética do artista ou uma

necessidade das condições de performance, como o canto não amplificado em um grande teatro (Titze et al., 2017). Dessa forma, em certos estilos de canto pode ser favorecido, em regiões de f_0 específicas, utilizar determinadas vogais ou configurações de trato vocal (Titze, 2008a). Mesmo que ocorram alterações nas vogais do texto, a prioridade é manter uma qualidade vocal apropriada ao gênero musical (Bozeman, 2013; Miller, 2019).

Figura 6 – Esquema da variação média das frequências de F1 e F2 em função das vogais, como proposto por Bozeman (2013) para o canto erudito.



Fonte: Autor. Adaptado de Bozeman (2013).

Bozeman (2013) relaciona as frequências em que começa a ser necessário modificar os formantes com as tradicionais zonas de passagem do canto erudito. Segundo Miller (2019), a zona de passagem (*zona di passaggio*), na técnica italiana do canto lírico para vozes masculinas, se refere à região de frequências em que se torna crítico fazer as modificações graduais na ativação dos músculos TA e CT. O limite inferior dessa região, a primeira passagem (*primo passaggio*), corresponde às notas em que se inicia a tendência à elevação laríngea e do queixo, sendo o limite da região confortável da fala. Um cantor sem treino consegue cantar mais agudo, porém as compensações articulatórias comprometem cada vez mais a qualidade vocal, até ser atingida uma determinada f_0 em que a voz irá quebrar ou trocar para o registro de

falsete. Esse é o limite superior da zona de passagem masculina, chamada de segunda passagem (*secondo passaggio*).

Miller (2019) situa a primeira passagem masculina entre as notas B2 (247 Hz) e D3 (294 Hz), e a segunda passagem variando entre E3 (330 Hz) e F#3 (370 Hz), dependendo da classificação vocal do cantor. Em f0 próximas a segunda passagem masculina, está a primeira passagem das vozes femininas, quando a cantora lírica deve fazer a troca da voz de peito para a voz mista de cabeça. O autor cita ainda a segunda passagem feminina, situada em notas agudas, acima das quais a cantora deve utilizar a voz de cabeça, o que ocorre entre D4 (587 Hz) e F#4 (740 Hz), dependendo da classificação vocal.

Bozeman (2013) propõe que, além das questões musculares, as passagens são também influenciadas pelas relações entre formantes e harmônicos, especialmente F1, e H1 e H2. Dessa forma, as notas exatas das passagens variam de acordo com as vogais. As notas nas quais H2 ultrapassa F1 nas vogais semifechadas /e/ e /o/ são aproximadamente as mesmas da primeira passagem masculina. A nota em que H2 ultrapassa F1 na vogal aberta /a/, que tem o valor mais alto de F1, corresponde à segunda passagem masculina e à primeira passagem feminina. As notas da segunda passagem feminina correspondem às frequências em que H1 se iguala a F1 na vogal /a/.

Bozeman (2013) faz uma classificação da voz em timbres, denominados registros acústicos, determinados pela relação entre F1 e os harmônicos, com destaque para H1 e H2. O autor relaciona os eventos acústicos a modificações auditivamente perceptíveis na qualidade vocal. Na região mais grave, quando há dois ou mais harmônicos abaixo de F1, ocorre o *open timbre*, ou “timbre aberto”. Quando há apenas um harmônico abaixo de F1, ou seja, quando H2 ultrapassa F1, tem-se o *close timbre*, ou timbre fechado. Essa qualidade vocal é citada pelo autor como o timbre característico do cantor erudito após efetuar a manobra chamada de cobertura, quando a voz adquire uma qualidade que pode ser descrita como mais escura, arredondada. O autor cita duas situações especiais, instintivamente utilizadas por humanos quando querem ou precisam vocalizar com muita intensidade. A primeira ocorre quando H2 se iguala a F1, chamada de *yell coupling*, ou “acoplamento de grito ou berro”, em que há forte domínio de H2 no espectro radiado. Esse é citado pelo

autor, além de outros pesquisadores (Titze e Worley, 2009), como o som característico do *belting*. O segundo acoplamento especial acontece quando H1 se alinha a F1, chamado de *whoop coupling*, com domínio de H1 no espectro. A palavra *whoop* não tem correspondência direta em português que seja adequada ao uso aqui empregado, com tradução “gritar, chamar, fazer algazarra” (Michaelis On-line, 2023). Durante uma *live* no canal do YouTube *Dr Dan’s Voice Essentials* (2018), Bozeman exemplificou a qualidade sonora típica do registro de *whoop* pelo som alto que se faz para comemorar em eventos esportivos, em f0 agudas no registro de falsete, na vogal /u/. Essa é a qualidade vocal característica das notas agudas femininas e dos contratenores¹² no canto erudito. O autor cita ainda a estratégia de modificar F2 para amplificar H3 ou H4, frequentemente utilizada por tenores de ópera acima da zona de passagem.

Quando o cantor intencionalmente altera F1 ou F2 da vogal que ele está cantando, para que estes amplifiquem algum harmônico específico, diz-se que ele está fazendo uma “sintonia de formantes” (*formant* ou *resonance tuning*). E a estratégia em que o cantor segue alterando as configurações do trato vocal supraglótico para modificar os formantes ao longo de uma passagem musical, a fim de manter um harmônico específico com maior intensidade, chama-se “rastreo de formantes” (*formant* ou *resonance tracking*). As estratégias são empregadas por cantores em diversos gêneros. Tenores eruditos, após perder a amplificação de H2, quando este ultrapassa F1, sintonizam F2 para amplificar H3 ou H4 nas notas agudas. Cantoras líricas, nas notas agudas de sua tessitura, acima das localizações normais de F1 das vogais faladas, aumentam intencionalmente a frequência de F1 para que ele continue alinhado ao H1, de modo a garantir intensidade sonora. Cantores de *belting* aumentam F1 em notas médio-agudas, para manter a forte amplificação de H2 na região acima da nota em que ele ultrapassaria F1 (Bozeman, 2013).

Um caso especial do uso dos formantes no canto é o formante do cantor (*Singer’s Formant Cluster*), descrito por Sundberg (2018). A estratégia, utilizada por vozes masculinas no canto erudito, especialmente tenores, consiste normalmente na elevação da frequência de F3 e no abaixamento de F5, para que eles se aproximem de F4. A aproximação de formantes faz com que as suas amplitudes e a das regiões

¹² Voz masculina caracterizada pelo uso do registro de falsete (Sundberg, 2018).

entre eles aumentem. Dessa forma, os três formantes juntos criam um forte pico de energia por volta de 2500 a 3000 Hz, uma região de frequências em que o som da orquestra não tem muita intensidade, garantindo que a plateia escute o cantor mesmo contra os fortes sons dos instrumentos. Titze e Worley (2009) associam a estratégia a um estreitamento da epilaringe, que corresponde aos primeiros 2 a 3 cm do trato vocal acima das pregas vocais, terminando nas pregas ariepiglóticas.

Leino (2009) descreveu ainda o “formante do falante”, um discreto pico de energia entre 3000 e 4000 Hz, próximo ao F4, encontrado em vozes masculinas consideradas de boa qualidade. Cleveland et al. (2001) identificaram o mesmo pico em cantores de música *country*. Os autores especulam que esta pode ser também uma estratégia utilizada na música popular, ou ainda resultado da qualidade de voz falada utilizada no gênero *country*. Master et al. (2008) também encontraram um pico de energia nessa região de frequência em atores, mais forte que em falantes não-atores, que denominaram de “formante do ator”.

A fim de controlar as frequências dos formantes para produzir a qualidade vocal que deseja, o cantor deve modificar a configuração do trato vocal por meio da manipulação dos seus articuladores. Segundo Sundberg (2018), o comprimento do trato vocal (distância entre a glote e os lábios) e a sua forma (variação da área transversal ao longo do comprimento) definem as frequências das ressonâncias. Essas dimensões dependem da anatomia e morfologia de cada pessoa, mas podem ser variadas. O abaixamento da laringe e a protrusão dos lábios alongam o trato vocal, diminuindo as frequências de todos os formantes, enquanto o estiramento dos lábios encurta o trato vocal, as aumentando. Mudanças na área transversal podem elevar as frequências de determinadas ressonâncias ao mesmo tempo que abaixam outras. F1 é particularmente sensível à abertura da mandíbula, aumentando com a maior abertura desta. F2 é definida principalmente pela forma do corpo da língua: quando esta comprime a parte anterior do trato vocal, F2 sobe; quando comprime a parte posterior, F2 desce. A distância entre a ponta da língua e os dentes incisivos inferiores é inversamente proporcional à frequência de F3. Os formantes mais agudos parecem estar mais relacionados ao comprimento do trato vocal.

Sundberg (2018) alerta que a movimentação dos articuladores não é isolada. Uma movimentação da língua, por exemplo, que é ligada ao osso hioide, por sua vez

ligado à laringe, pode causar uma movimentação desta, e a sua posição pode afetar o modo de fonação e as características da fonte glótica.

Titze e Worley (2009) discutem se o termo “sintonia de formantes” é adequado. Os autores dizem que, mesmo nos resultados publicados de pesquisadores que descrevem a estratégia, os harmônicos não costumam coincidir com a frequência central dos formantes. Para eles, é mais provável que os cantores aproximem os formantes dos harmônicos, de modo que aqueles fiquem ligeiramente acima destes, garantindo sua amplificação pelas regiões de reatância inercial.

3.4. As notas agudas e a qualidade vocal na pesquisa em voz cantada

Sousa et al. (2010) conduziram um estudo acerca do uso de metáforas como recurso didático na pedagogia vocal. A pesquisa consistiu na aplicação de um questionário em entrevistas abertas com vinte professores de quatro diferentes abordagens de ensino de canto (cinco de cada): erudito, canto popular brasileiro, canto comercial contemporâneo americano, e abordagem holística. Apesar do estudo não ser focado na produção de notas agudas, serão destacados aqui aspectos relevantes ao presente trabalho. Perguntados sobre os conceitos que aplicam em aula, relacionados à ressonância vocal, 60% dos entrevistados ressaltaram a importância do apoio; 35% deram ênfase à qualidade da emissão, citando o equilíbrio da vibração das pregas vocais e dos registros; e 25% apontaram o filtro como definidor das sonoridades. Sobre as metáforas utilizadas, foram mencionadas pelos professores: vogais mais claras/escuras, dar leveza à voz, trazer harmônicos, ressonância frontal/focar a voz, bocejo, valorização do espaço, imaginar uma cúpula de igreja dentro da boca, sentir o perfume de uma flor, relaxar o aparelho fonador. Tais metáforas foram correspondidas a expectativas acerca de ajustes musculares objetivados pelos professores, como elevação do palato, expansão faríngea, abaixamento da laringe, abaixamento da língua, fonação equilibrada/alternância dos músculos TA e CT, aumento da adução glótica/diminuição do escape de ar, respiração tranquila, controle do excesso de Psub, elevação dos músculos zigomáticos e relaxamento da mandíbula.

O estudo de Fernandes e Andrada e Silva (2020a) teve como objetivo analisar as emissões superagudas de cinco sopranos eruditas profissionais, duas lírico-ligeiro

e três lírico-coloratura. O trabalho baseou-se nos autorrelatos das cantoras e na avaliação percepto-auditiva de três professores de canto e três fonoaudiólogos. Foi analisada a nota D5 (1175 Hz) na vogal /e/, extraída da execução de um trecho de uma ária de ópera. De maior interesse para este trabalho, são as respostas das cantoras às questões da entrevista semidirigida, que considerou sensações corporais, sensações de movimentação laríngea, e uma terceira pergunta mais aberta acerca de ajustes necessários à emissão da nota superaguda. As cinco cantoras relataram a necessidade de aumentar o apoio ou a sustentação; quatro relataram sentir uma certa elevação laríngea, enquanto uma disse perceber um pouco de movimentação laríngea; duas relataram sensações no corpo inteiro; uma identificou a necessidade de elevar o músculo zigomático a abaixar a parte inferior da boca; e uma descreveu como se a nota se “encaixasse em outro lugar”. Os autores apontaram para a falta de pesquisas que investigam a questão perceptiva dos cantores durante suas emissões agudas, com a maioria dos trabalhos focados apenas nos aspectos acústicos e nos ajustes de trato vocal.

Outra pesquisa conduzida pelos mesmos autores (Fernandes, Andrada e Silva, 2020b), ainda a respeito da produção de notas superagudas, propôs-se a comparar as opiniões de professores de canto às de sopranos profissionais, cujos repertórios demandassem a emissão de tais notas. Foi aplicado um questionário formulado em cima das entrevistas do estudo anterior (Fernandes, Andrada e Silva, 2020a). A pesquisa contou com 35 participantes: 24 cantoras profissionais, as quais foi solicitado que respondessem o questionário com base em suas sensações; e 11 professores experientes, que responderam de acordo com suas recomendações pedagógicas. Foram abordados apoio respiratório, movimentação vertical laríngea, posicionamento da língua, padrões articulatórios, ressonância, e aspectos auditivos e corporais. Cantoras e professores concordaram na importância de adaptar o apoio respiratório, na posição da língua “descansada” no assoalho da boca, e nos ajustes corporais (liberdade de movimento e participação do corpo todo na produção). Houve discordância sobre o movimento laríngeo, com instrução de abaixamento pelos professores, enquanto cantoras relataram elevação; e sobre o padrão de movimento articulatório da boca, com ambos considerando a necessidade da abertura vertical, mas com a sensação sorriso e abertura horizontal sendo também consideradas, apenas pelas cantoras. Sobre os aspectos auditivos, somente as cantoras percebem

diminuição da intensidade, que a voz se torna “menor”, e que há modificações no brilho e no timbre. Os autores explicam esse achado, pelo fato de a audição que um indivíduo tem da sua própria voz é mediada não apenas pela propagação das ondas sonoras no ar, como ocorre para o ouvinte externo, mas também pela condução óssea. Os resultados sugerem a necessidade de ajustes e intervenções pedagógicas individualizadas para cada cantora produzir suas notas hiperagudas.

Serão destacados dois estudos que apontaram para as necessidades de adaptações fisiológicas e musculares laríngeas em função das variações de f_0 : as pesquisas de Roubeau et al. (2009) e Kochis-Jennings et al. (2014).

Roubeau et al. (2009) empregaram a eletroglotografia (EGG), em que dois eletrodos circulares são posicionados sobre o pescoço do sujeito, cada um em um lado da cartilagem tireoide. Esse método mede a área de contato entre as pregas vocais, por meio de uma corrente elétrica de baixa intensidade que passa entre os eletrodos. Quanto maior o contato entre as pregas vocais, mais facilmente a corrente passa através delas, de um eletrodo para o outro, pois os tecidos são bons condutores de eletricidade. Quando as pregas vocais se afastam, a corrente encontra maior resistência para passar, uma vez que o ar entre as pregas vocais é um mau condutor de eletricidade. Dessa forma, o método permite avaliar de forma indireta a amplitude do contato das pregas vocais durante ciclos oscilatórios sucessivos, além das principais fases do contato.

O estudo (Roubeau et al., 2009) teve como objetivo introduzir o conceito de mecanismos vibratórios laríngeos, além de relacioná-los às concepções de registros vocais. Os autores utilizaram três bases de dados de pesquisas anteriores, das quais duas são relevantes para o presente trabalho. A primeira era composta de sinais de EGG e de áudio de 10 cantores e 9 cantoras, com e sem treinamento vocal, com o objetivo de analisar a troca entre mecanismos laríngeos. Os sujeitos produziram glissandos ascendentes e descendentes¹³ de forma espontânea, e vogais /a/, /o/ e /i/ sustentadas, em dois mecanismos laríngeos, em diferentes notas. A segunda pesquisa consistiu na elaboração do perfil de extensão vocal (PEV) de 21 cantores e

¹³ No glissando ascendente, o cantor vocaliza de suas notas mais graves até as mais agudas, subindo progressivamente f_0 . No glissando descendente, é feito o movimento contrário (Roubeau et al., 2009).

21 cantoras, entre profissionais, amadores e pessoas sem treinamento vocal. Foi utilizado um medidor de nível de pressão sonora (NPS) para fazer a construção dos PEV com medidas de emissões da vogal /a/ em semitons, para as faixas de frequências de dois mecanismos laríngenos.

Na análise dos espectrogramas de banda estreita dos glissandos espontâneos, foram identificados três pontos de descontinuidade, que correspondem aos limites entre os mecanismos laríngenos M0, M1, M2 e M3. A quebra de M1 para M2 está associada à redução da amplitude e modificação do formato do sinal do EGG. Essas mudanças, por sua vez, são relacionadas pelos autores à diminuição da espessura das pregas vocais, causada pelo desacoplamento entre as camadas de tecido mais superficiais da cobertura e aquelas mais profundas do corpo das pregas vocais. A oscilação fica limitada à cobertura, com consequente redução da massa vibrante. A troca de M1 para M2 ocorreu em média em 238 Hz (Bb2) nos homens e 312 Hz (Eb3) nas mulheres, nos glissandos ascendentes. As mudanças repentinas na amplitude e na forma do sinal de EGG foram igualmente observadas nas tarefas de vogais sustentadas, com troca intencional de mecanismo durante a emissão. Os PEV do segundo estudo mostram que há uma zona de sobreposição entre M1 e M2, em que é possível escolher quais dos mecanismos será utilizado, correspondente a uma faixa de frequência de aproximadamente uma oitava – de F2 (175 Hz) a F#3 (370 Hz) nos homens e G#2 (208 Hz) a G#3 (415 Hz) nas mulheres, em média (Roubeau et al., 2009).

O momento de desacoplamento súbito das camadas profundas das pregas vocais é relacionado pelos pesquisadores ao conceito de massa crítica para o registro de peito. Com o aumento da f_0 , a tensão e a rigidez não são igualmente distribuídas entre as diferentes camadas de tecidos, em função das suas particularidades histológicas. Quando é atingido um ponto crítico, tais particularidades acabam induzindo o desacoplamento entre as camadas das pregas vocais. A repentina alteração das características biomecânicas da massa vibrante é a provável causa dos saltos de f_0 característicos das quebras vocais. Mesmo naqueles sujeitos treinados, que conseguiram fazer a troca de mecanismo nas notas sustentadas sem quebra audível, ocorreram as mudanças drásticas no sinal do EGG. Tal feito mostra ser

possível o manejo das trocas de mecanismos, de modo a fazer parecer que a voz é auditivamente homogênea (Roubeau et al., 2009).

O estudo de Kochis-Jennings et al. (2014) utilizou medidas de eletromiografia (EMG), em que eletrodos metálicos em forma de gancho são inseridos diretamente nos músculos para verificar sua atividade elétrica durante diferentes tarefas. O objetivo foi testar a hipótese de que os registros de peito e mix de peito teriam predomínio de atividade de músculos TA, enquanto os de cabeça e mix de cabeça teriam predomínio de atividade dos músculos CT. Os autores apontaram o uso dos termos “domínio de TA” e “domínio de CT” na pedagogia vocal para descrever diferentes qualidades vocais, respectivamente mais semelhantes ao registro de peito (*chest-like*) e ao de cabeça (*head-like*). Os pesquisadores relataram, porém, não haver ainda estudos que comparassem diretamente a proporção de atividade entre os dois grupos musculares em diferentes registros, apenas o acionamento isolado de cada um. Eles alertaram para que os achados de maior ou menor atividade de um determinado músculo em um registro do que em outro não fossem generalizados para a conclusão de que esse músculo esteja mais ou menos ativo que outro músculo, com cuja ação não foi comparado.

Participaram da pesquisa sete sujeitos (duas cantoras treinadas, e quatro cantoras e um cantor não treinados). Os eletrodos foram inseridos nos músculos TA e CT dos sujeitos, que também tiveram o sinal de áudio captado por microfones. As cantoras com experiência fizeram emissões sustentadas em notas entre Eb3 (311 Hz) e Bb3 (466 Hz), cada nota repetida nos registros de peito, mix de peito, mix de cabeça e cabeça; e cantaram um excerto de uma canção, também repetida nos quatro registros. Todos os sujeitos vocalizaram glissandos ascendentes na vogal /i/ (Kochis-Jennings et al., 2014).

Os resultados revelaram que, na maioria dos sujeitos, independentemente de registro, as fonações abaixo de 200-300 Hz (aproximadamente G2-D3) apresentaram predomínio de atividade de TA, enquanto aquelas acima apresentaram predomínio de CT. Mesmo nas notas mais agudas do registro de peito, entre 400-430 Hz (G3-A3), foi observada fonação com domínio de CT; e uma cantora apresentou domínio de TA nas fonações graves, abaixo de 350Hz (F4), mesmo em registro de cabeça e mix de cabeça. Duas cantoras apresentaram domínio de CT em todos os registros, mesmo

nas notas graves em registro de peito. Todos, à exceção de uma cantora, apresentaram pequenas mudanças de proporção CT:TA logo antes ou durante as transições de registro nos glissandos. Os comportamentos específicos variaram nas notas mais agudas, com diferentes variações na proporção CT:TA em função de registro de f0, mas todos exibiram nelas maior ou menor predomínio de CT. Quebras abruptas de registro foram identificadas em alguns sujeitos nos glissandos ascendentes mesmo quando já havia predomínio de CT desde notas mais graves. Os autores levantaram a hipótese de que a co-contracção de outras musculaturas laríngeas, das quais pouco então se sabia, poderia estar envolvida nas trocas de registros. Os autores concluíram que a proporção de atividade entre TA e CT estaria geralmente mais relacionada à frequência de fonação do que ao registro, e que ambos os músculos seriam progressivamente mais recrutados no continuum voz de cabeça – mix de cabeça – mix de peito – voz de peito (Kochis-Jennings et al., 2014).

Vale notar que a faixa de frequências em que Roubeau et al. (2009) identificaram a tendência a ocorrer as mudanças de M1 para M2 (238 a 312 Hz), e Kochis-Jennings et al. (2014) a troca de predomínio de atividade entre os músculos TA e CT (300 a 350 Hz), corresponde aproximadamente àquela identificada por Miller (2019) como a zona de passagem para as vozes masculinas e à primeira passagem das vozes femininas.

As características espectrais e as estratégias de ajustes de ressonâncias foram estudadas por diversos autores, prioritariamente no canto erudito, mas também em alguns outros gêneros de canto. O estudo de Schutte e Miller (1993) teve como objetivo descrever aspectos acústicos e fisiológicos de estilos de canto não-eruditos na extensão média da voz feminina. Foram captados sinais de EGG e de áudio de uma única mezzo-soprano, que produziu emissões em três diferentes estilos, lírico, *pop* e *belting*, nas vogais /a/ e /ɔ/, em notas médias e agudas, F#3 (370 Hz), Bb3 (466 Hz) e Eb4 (622 Hz). Logo após cada emissão das vogais, a cantora produziu um som em *vocal fry*, ao mesmo tempo que tentava manter a configuração do trato vocal. A fonte do *vocal fry* é não-harmônica e resulta em um espectro contínuo, de forma que os seus picos correspondem às frequências centrais dos formantes.

Os autores constataram o forte domínio do H2 sobre H1 no *belting*, em todas as f0 analisadas, enquanto no canto lírico encontraram proeminência de H3 na nota

F#3 e intensidades semelhantes em H1 e H2 na nota Bb3, com uma qualidade sonora resultante mais “redonda” e “escura”. O gênero *pop*, na nota Bb3, exibiu também predomínio do H2, porém não tão intenso quanto no *belting*, e apresentou H1 com maior amplitude que este. Outra diferença foi a distribuição da energia nas frequências mais agudas do espectro: enquanto o *belting* apresentou energia mais homogeneamente distribuída entre os harmônicos mais agudos, o lírico e o *pop* apresentaram maior concentração de energia entre H6 e H8, com os outros harmônicos agudos mais atenuados. Essa alta energia nos harmônicos agudos foi relacionada ao resultado perceptivo intenso e claro do *belting*. As qualidades vocais características de cada estilo foram associadas às frequências de F1 e F2, mais baixas no canto lírico, intermediárias no *pop* e mais altas no *belting*, este último com aumento progressivo de F1 para rastrear H2. A laringe alta e as vogais mais abertas foram apontadas pelos autores como a provável estratégia utilizada pelas *belters* para elevar os formantes (Schutte e Miller, 1993).

Titze et al. (1994) tiveram como objetivo explorar as propriedades acústicas da voz aguda do tenor erudito. Foram analisadas, por meio de análise espectrográfica, notas entre F3 (349 Hz) e B3 (494 Hz). Seis tenores não profissionais, mas em treinamento para ingressar na carreira artística, tiveram suas vozes gravadas por um microfone, em frente a uma pequena audiência. Os cantores repetiram sequências de vogais encadeadas, em notas entre C3 (262 HZ) e B3 (494 HZ). A sequência também foi repetida em uma nota grave, C2 (131 HZ), com qualidade de fala e de canto. A nota mais grave foi escolhida por permitir melhor estimativa das frequências dos formantes: uma vez que os harmônicos estão mais próximos e a informação está mais compactada no espectro, o perfil deste corresponde razoavelmente ao da função de transferência do trato vocal. Os formantes das notas agudas foram estimados por meio de modelos matemáticos teóricos do trato vocal e de técnicas de otimização de parâmetros operadas por *software*. Em comparação com a voz falada, os valores de F1 das vogais cantadas foram significativamente mais altos, provavelmente pela maior abertura de boca no canto. Todos os cantores elevaram F1, possivelmente para mantê-lo acima de H1 e impedir que eles coincidissem. Posto que essa sintonização é a estratégia utilizada por sopranos, o timbre resultante seria provavelmente pouco masculino. A estratégia mais comum foi distribuir a energia entre H2, H3 e H4 coletivamente, obtendo regiões de grande amplitude mais largas no espectro, ao invés

de harmônicos isolados fortes. Entretanto, o tenor mais experiente exibiu a tendência de reforçar H3 nas vogais /a/ e /o/.

Hollien et al. (2000) analisaram 18 cantores eruditos profissionais (13 mulheres e 5 homens), para verificar a inteligibilidade das vogais, a extensão da mudança nos formantes, e a influência dessas mudanças na identificação dessas vogais em f_0 agudas. Foram gravadas as emissões dos sujeitos nas vogais /a/, /i/ e /u/, sustentadas em duas notas dentro da suas extensões vocais, uma moderadamente grave e uma extremamente aguda, nas intensidades fraca e forte. Foram reproduzidas notas de referência para os cantores, mas foi permitido que eles emitissem a nota mais confortável próxima à nota ouvida. As notas de referência graves foram G1 (98 Hz) ou A1 (110 Hz) para os homens, e G2 (196 Hz) ou A2 (220 Hz) para as mulheres; e as agudas ao menos D3 (294 Hz) para os homens e A4 (880 Hz) para as mulheres. Serão destacados os resultados acerca das frequências dos formantes. Tanto os homens quanto as mulheres apresentaram, nas notas graves, uma diminuição do F2 da vogal /i/ em relação ao esperado na fala para uma f_0 semelhante. Nas notas agudas, o F1 das vogais /i/ e /u/ subiram quase uma oitava e o da vogal /a/ aproximadamente meia oitava. O F2 da vogal /i/, por já ter uma frequência bastante aguda, não foi muito afetado pela subida de f_0 , enquanto o das vogais /a/ e /u/ aumentaram substancialmente. Os autores concluíram que o aumento de f_0 cria o que pode ser considerado “uma nova vogal”, uma vez que os F1 das três vogais tendem a se encontrar numa frequência ligeiramente acima do F1 esperado para a vogal /a/, e o F2 em uma frequência próxima ao esperado para a vogal /i/.

O estudo de Sundberg (2001) investigou alguns aspectos acerca do formante do cantor, o pico de energia por volta de 3000 Hz identificado em cantores, causado pela aproximação de F3, F4 e F5. A pesquisa teve dois objetivos. O primeiro foi investigar as variações de amplitude do formante do cantor em função do NPS e das vogais, e discutir como definir uma amplitude mínima para se considerar que um cantor apresenta o formante do cantor. Para tanto, foram comparadas simulações computacionais a emissões de cantores. O segundo objetivo foi identificar a frequência central do pico do formante do cantor por classificação vocal, por meio da análise de espectros médios de longo termo (EMLT) de gravações comerciais de cantores de ópera. Foram analisados excertos de aproximadamente 30 s de sopranos,

contraltos, tenores, barítonos e baixos. Foram selecionados quatro cantores de cada classificação vocal, e analisados quatro excertos de cada um. As frequências centrais foram medidas para todas as vozes, com exceção das sopranos, que apresentaram muito menor energia na região de frequências do formante do cantor, e dois picos identificáveis ao invés de um, o que sugere não ter havido aproximação dos formantes. Houve variações sistemáticas da medida entre os tipos vocais. Nas contraltos, com um pico mais expressivo que o das sopranos, porém ainda bem menor que o dos homens, foi identificada a frequência central em 3000 Hz. Os homens apresentaram um pico amplo e muito melhor definido, os baixos com frequência central média em 2420 Hz, barítonos em 2550 Hz, e tenores em 2840 Hz. É relevante para o presente trabalho destacar que houve considerável variação na frequência central do formante do cantor dos tenores, entre aproximadamente 2600 Hz e 3100 Hz. As vogais /o/ e /u/ apresentaram maior amplitude no formante do cantor do que as outras. O autor hipotetiza que a diferença se dá com o objetivo de equalizar o timbre dessas vogais com o das outras, que tem mais energia na região aguda, mas alerta que é necessário levar em consideração o NPS da emissão vocal, informação indisponível, uma vez que os dados foram extraídos de gravações comerciais.

A pesquisa de Cleveland et al. (2001) testou duas hipóteses, levantadas com base em estudos anteriores, acerca de cantores de música *country* norte-americana: que o EMLT da voz cantada desses artistas seria semelhante ao da voz falada; e que eles não apresentariam no mesmo espectro o formante do cantor, pico de energia por volta de 2800 Hz presente em cantores eruditos. Participaram da pesquisa cinco cantores de música *country*, experientes e razoavelmente conhecidos. Todos tinham o sotaque típico do sul dos Estados Unidos, onde a música *country* é característica. Para comparação, participou também do estudo um cantor erudito profissional, como exemplo de EMLT que apresenta o formante do cantor. Os cantores de *country* cantaram uma canção de sua escolha do repertório *country*, e o hino nacional dos Estados Unidos. Os cantores também leram as letras das duas músicas de maneira conversacional. O cantor erudito cantou o hino norte-americano, e leu a letra do mesmo nas formas conversacional e declamatória. De todas as gravações foram gerados EMLT.

Como esperado, o cantor erudito exibiu, na voz cantada, o pico característico do formante do cantor em 2800 Hz, seguido de uma queda brusca de energia na região de frequências logo acima dele. Os cantores de *country* não exibiram o formante do cantor em nenhuma tarefa. Três dos sujeitos da música *country* apresentaram EMLT muito semelhantes entre voz falada e cantada. Um outro exibiu a curva da voz falada toda com menor amplitude que a da voz cantada. A curva da leitura do hino americano de outro deles apresentou uma queda mais brusca de energia em direção às frequências mais agudas, o que os pesquisadores associaram ao fato de ele ter lido a letra bem mais baixo do que executou as outras tarefas. O cantor erudito não apresentou o formante do cantor na fala conversacional, nem na fala declamatória. Os EMLT dos cantores de *country* apresentaram um pico de energia por volta de 3500 Hz, tanto na voz falada quanto na voz cantada (vale destacar que esse pico não apresentou amplitude comparável àquela do pico do formante do cantor no canto erudito). Essa região de frequências é a mesma do “formante do falante” identificado em estudos anteriores. Um pico de destaque nessa região exige que o espectro da fonte tenha uma boa quantidade de energia nas frequências agudas, e que a largura de banda dos formantes seja estreita. Ambos os requisitos são resultado de um bom fechamento glótico, sem escape de ar. Com efeito, estudos anteriores com os mesmos cantores revelaram o emprego de uma longa fase fechada. (Cleveland et al., 2001).

Os autores (Cleveland et al., 2001) justificam a falta do formante do cantor nos cantores de música *country*, primeiramente por um fator técnico: diferentemente dos cantores eruditos, eles contam com um sistema de amplificação eletrônica, ou seja, não precisam projetar a voz em um teatro, sobre o som de uma orquestra, sem microfone. O formante do cantor necessita de algumas modificações às vogais para ser produzido, e parece contribuir para a sonoridade característica do canto erudito; utilizá-lo na música *country* iria potencialmente descaracterizá-la. No canto erudito, é essencial manter uma qualidade vocal homogênea ao longo de toda extensão vocal; na música *country*, é muito importante a inteligibilidade das letras das canções, uma vez que elas normalmente narram uma história. Essa qualidade de narrativa também é associada pelos autores à manutenção da qualidade da voz falada na música *country*.

O estudo de Schutte et al. (2005) buscou identificar estratégias de ressonância em notas agudas de tenores eruditos, por meio da examinação de gravações comerciais de cantores de grande renome internacional do século XX. Foram analisadas 60 gravações de performances de 34 tenores da ária *Celeste Aida*, do personagem Radamès na ópera *Aida*, do compositor Giuseppe Verdi. Foi analisada a nota aguda Bb3 (466 HZ), sustentada ao final da ária, na vogal semifechada [o] da palavra *sol*. É importante ressaltar que, nesse estudo, os autores não executaram medidas das frequências dos formantes, mas sim inferiram suas distâncias em relação aos harmônicos, com base nas amplitudes destes e no perfil geral dos espectros.

Dos 34 tenores, 24 apresentaram, em pelo menos uma de suas performances, grande amplitude no H3, provavelmente pela sua sintonia por F2. Tal estratégia foi considerada pelos autores como o padrão para tenores de ópera em notas agudas (tendo como base também resultados de medições de formantes feitas em pesquisas anteriores). Os autores sugerem que a estratégia seja conscientemente empregada, uma vez que ela acarreta, no caso da vogal [o], grande aumento da frequência de F2, o que causa considerável modificação à qualidade da vogal. O som potente produzido por tenores, advindo do quociente fechado (porcentagem do ciclo glótico em que as pregas vocais ficam em contato) relativamente alto de suas produções vocais, gera um espectro glótico com grande energia nos parciais agudos, o que é favorável para a amplificação de harmônicos acima do H2. A segunda estratégia mais identificada foi a utilização do formante do cantor, que amplificou, nesse caso, os harmônicos H5 a H7, entre 2300Hz e 3000 Hz, pela aproximação de F3, F4 e F5. Os autores destacaram essa ser a opção de tenores muito conhecidos por cantar o papel de Radamès, um marco no repertório do *tenore robusto*, o que sugere ser o formante do cantor uma estratégia típica de vozes dramáticas. Associado ao formante do cantor, alguns cantores exibiram importante amplitude no H1, provavelmente pela aproximação de F1. A utilização do formante do cantor comumente aparece com menores alterações das frequências de F1 e F2 esperados para a vogal. Alguns poucos tenores apresentaram ainda, quando executaram a nota toda ou parte dela em *pianissimo* (baixa intensidade), o domínio de H2 no espectro (porém não muito mais forte que H1), provavelmente pela proximidade de F2. Três tenores exibiram semelhante domínio de H2 em emissões de maior intensidade, o que os autores sugerem ter sido alcançado pelo agrupamento de F1 e F2, respectivamente abaixo e

acima de H2. Foram identificados também casos de uso combinado do H1 amplificado por F1, ou do formante do cantor, simultaneamente à amplificação do H3 por F2 (Schutte et al., 2005).

Neumann et al. (2005) focaram sua pesquisa na zona de passagem masculina em cantores líricos. O estudo se propôs a fornecer uma definição para a transição de registro de peito para o de cabeça (aqui considerado como “voz aguda plena”, e não como falsete), baseada na detecção de parâmetros acústicos objetivos, característicos da zona de passagem. Participaram da pesquisa onze cantores profissionais de um coro de ópera (tenores, barítonos e baixos). Foram registrados sinais de áudio e de EGG, captados durante a execução de escalas nas vogais abertas /a/ e /ɔ/, através da região de notas em que ocorre a transição do registro de peito para o de cabeça. Alguns cantores também fizeram as produções em registro de peito durante toda a escala. Das gravações de áudio foram gerados espectros de banda estreita, e no sinal de EGG foi calculado o quociente fechado (proporção de tempo que a glote permanece fechada durante um ciclo glótico).

Todos os sujeitos exibiram os mesmos padrões espectrais na transição de peito para cabeça, relacionados aos formantes F1 e F2, e aos harmônios H2, H3 e H4. As notas em que ocorreram as transições variaram de acordo com a classificação vocal. Os autores presumiram as movimentações dos formantes com base nos desenvolvimentos dos espectros, uma vez que não foram efetuadas medidas formânticas, à exceção de um sujeito, no qual foram medidos os formantes durante fonação ingressiva. No limite superior do registro de peito, foi identificada a sintonia de F1 com H2, e F2 com H4, com progressivo aumento das frequências de F1 e F2 para rastrear os harmônicos citados. A transição para o registro de cabeça foi caracterizada pela diminuição da frequência de F1, resultante na perda de amplitude de H2, e pela mudança da sintonia de F2 de H4 para H3. Já nas escalas em que foi mantido o registro de peito até o agudo, os cantores seguiram elevando F1 e F2 para manter a amplitude de H2 e H4. Dessa forma, os autores descreveram, para as vogais abertas /a/ e /ɔ/, o registro de peito como sendo caracterizado pelo domínio de H2 e H4, e a transição entre os registros marcada pela diminuição na frequência de F1, com a registro de cabeça caracterizado pela consequente perda de amplitude de H2, e pela grande amplitude de H3, agora amplificado por F2. Os autores alertam que

essas descrições não podem ser generalizadas para outras vogais, pela grande diferença dos valores de F1 e F2. Não foram observados padrões consistentes nas medidas de quociente fechado extraídas dos sinais de EGG: em algumas gravações não houve mudanças, em outras permaneceu igual, porém foi detectada uma tendência a subir na maioria. Para os autores, as marcantes diferenças inter-sujeitos indicam que os ajustes glóticos às novas ressonâncias podem ser relacionados a técnicas individualmente empregadas. Os pesquisadores concluíram que o padrão da fonte glótica do registro de peito permanece o mesmo na transição para o registro de cabeça masculino, e que as mudanças entre os registros são causadas principalmente por ajustes ressonanciais (Neumann et al., 2005).

Com o objetivo de mostrar como as relações não-lineares entre glote e trato vocal podem ser utilizadas em notas agudas no canto, Titze e Worley (2009) apresentaram um estudo baseado em modelos de trato vocal. Foram contrastadas as configurações de trato vocal de dois gêneros de canto na voz masculina, ópera e *belting*. Os autores se perguntam se a formato de boca do canto operístico, descrito como um megafone invertido (abertura dos lábios menor que a da faringe), favoreceria o reforço dos harmônicos acima de f_0 , nas regiões de reatância inertiva supraglótica acima de F1; e se o formato de megafone ou trompete (lábios abertos e faringe mais estreita), característico do *belting*, favoreceria H1 e H2, na região de reatância inertiva abaixo de F1. O artigo é iniciado com uma discussão teórica e matemática da impedância do trato vocal, e apresenta, para os modelos de trato vocal, gráficos chamados inertogramas, que plotam o valor de inertância em função da nota cantada. A inertância é um parâmetro calculado a partir dos valores da reatância inertiva, e como esta, também tem picos de valores máximo logo abaixo das frequências dos formantes, que favorecem a amplificação dos harmônicos. Considera-se aqui os termos “inertância”, “reatância inertiva” e “reatância positiva” semelhantes: referem-se às regiões de frequência benéficas à amplificação dos harmônicos, abaixo dos formantes. Os inertogramas demonstram o efeito do estreitamento na epilaringe, típico do formante do cantor, sobre a inertância, mantendo seus valores favoráveis ao longo de uma maior faixa de frequências e amplificando harmônicos entre 2000 e 3000 Hz.

Foram modelados para o estudo quatro tratos vocais: dois baseados em imagens de ressonância magnética de um barítono erudito, durante a fala e durante o

canto; um baseado no tenor operístico Luciano Pavarotti; e um baseado no cantor de jazz Cab Calloway. O barítono era especialista em *Lieder* (canções eruditas alemãs) e em música antiga, e não apresentou o estreitamento da epilaríngia característico de cantores de óperas. Os tratos vocais de Pavarotti e Calloway foram modelados a partir de imagens de vídeo dos cantores, disponíveis no YouTube, e de suposições baseadas em conhecimentos científicos acerca dos respectivos gêneros de canto. Os autores reconhecem que os modelos não são reproduções exatas dos respectivos tratos vocais, mas servem como exemplos característicos dos dois estilos de canto: o megafone invertido do canto operístico, com laringe baixa, faringe ampla e estreitamento epilaríngio; e o megafone do *belting*, com boca mais aberta, faringe estreita e laringe mais alta (Titze e Worley, 2009).

Foram analisadas imagens do rosto dos cantores durante a emissão da nota A3 (440 Hz). A abertura de boca de Calloway é muito maior que a de Pavarotti. Sobrepostos os harmônicos da nota A3 (440 Hz) aos inertogramas dos modelos, todos os tratos vocais teriam H1 dentro da região de inertância abaixo de F1, enquanto apenas no modelo de Calloway H2 estaria nessa região. O modelo de Pavarotti apresentaria H2 e H3 dentro da zona de inertância de F2, com H3 no ponto máximo de inertância, logo abaixo do F2. A inertância aumentada na região de 2500 Hz pelo estreitamento epilaríngio no modelo de Pavarotti amplificaria H6. No trato vocal do barítono, por não ocorrer esse estreitamento, não haveria a região de inertância nessa faixa de frequências. O H2 no modelo de Calloway ganharia grande amplificação na região de alta inertância logo abaixo de F1, enquanto seu H3 não estaria em região favorável (Titze e Worley, 2009).

As análises dos espectros de banda estreita dos áudios correspondentes às imagens satisfizeram às predições das modelagens: Pavarotti apresentou H1, H2 e H3 fortes, particularmente H2 e H3. Sua região do formante do cantor não beneficiou apenas H6, mas também H7 e H8 foram amplificados. Os autores justificam o fato pela modelagem não corresponder exatamente ao trato vocal do cantor, e que o reforço de H7 e H8 poderia ser obtido na nota emitida com estreitamento epilaríngio. O espectro de Calloway foi dominado por H2, muito mais forte que H1 e H3, também em concordância com a modelagem. Foram também analisados espectros e imagens do tenor operístico Roberto Alagna e do cantor de teatro musical Tony Vincent,

durante a emissão da nota Ab3 (415 Hz). O cantor lírico novamente mostrou menor abertura de boca que o popular. Alagna apresentou equilíbrio de energia entre os harmônicos mais baixos, e intenso formante do cantor amplificando H5 e H6; enquanto o espectro de Vincent foi marcado por grande amplitude de H2, como o de Calloway, e H1 com pouca intensidade (Titze e Worley, 2009).

O mesmo método do estudo de Titze e Worley (2009) foi aplicada por Titze et al. (2011) para modelar o trato vocal de vozes femininas, a fim de comparar *belters* do teatro musical americano e sopranos operísticas. Foram analisados os formatos de boca das emissões de três cantoras da *Broadway*, nas notas C#4 (554 Hz) e F4 (698 Hz), e de três sopranos operísticas na mesma região de frequências, nas notas D4 (587 Hz) e E4 (659 Hz), e uma oitava acima, C#5 (1109 Hz) e D#5 (1245 Hz). Foram ainda analisadas, em duas sopranos operísticas, emissões consideradas extremamente agudas, nas notas F5 (1397 Hz) e Bb5 (1867 Hz). As maiores aberturas de boca foram identificadas nas *belters*, seguido das notas extremamente agudas das sopranos (F5 e Bb5), em sequência nas notas agudas das sopranos (C#5 e D#5), e por último nas notas médias das sopranos (D4 e E4).

Nos inertogramas formulados a partir da modelagem dos tratos vocais das sopranos, foi identificado que, nas notas médias, H1 estava na região de inertância positiva abaixo de F1, o que corresponde à forte amplitude de H1 medida nos espectros radiados das cantoras. Em contrapartida, os espectros das cantoras *belters* foram caracterizados por um forte H2, em concordância com os tratos vocais modelados para suas emissões, cujos inertogramas apresentaram H2 logo abaixo de F1, em sua região positiva de inertância. Dessa forma, o F1 das cantoras de teatro musical estava uma oitava acima do das sopranos operísticas, aproximando-se do limite possível previsto pelos inertogramas. Os autores comentam que essas notas costumam ser, para muitas *belters*, o limite de extensão, uma vez que acima delas não é possível manter a amplificação de H2 pelo F1, sonoridade característica do *belting*. As sopranos, ao ascender de sua região média para as notas agudas, também começam a abrir mais a boca para subir a frequência de F1, porém para manter a grande amplitude de H1. Como H1 está uma oitava abaixo de H2, as sopranos têm ainda uma oitava além das *belters* até chegar no limite da manutenção da sintonia de F1. De fato, nos inertogramas modelados para as emissões no extremo agudo das

sopranos, todos os harmônicos estavam acima de F1, o que indica não haver nenhuma participação desse formante na amplificação de harmônicos nessa região de f0. Os resultados são relacionados às preferências de vogais no *belting*, em que são evitadas vogais com F1 baixo, de modo que /i/ e /u/ tendem a ser modificados em direção a /ɛ/ e /ɔ/, respectivamente. Outra possibilidade levantada pelos autores é a utilização de uma posição mais alta de laringe e uma faringe mais estreita, o que aumentaria a inertância geral do trato vocal, além de manter F1 alto. Os autores também sugerem que tais limites estão relacionados à utilização do registro modal nas notas agudas pelas *belters*, que normalmente precisam fazer uma troca brusca para um registro mais leve quando cantam acima do E4 (659 Hz) (Titze et al., 2011).

O objetivo da pesquisa de Henrich et al. (2011) foi investigar se cantores utilizam sintonia dos formantes F1 e F2 a H1 ou outros harmônicos, em diferentes regiões de f0. Participaram da pesquisa 22 cantores, entre nacionalmente conhecidos e amadores: seis sopranos, quatro contraltos, quatro barítonos e oito tenores. Foram gravadas as emissões de diferentes vogais pelos cantores, ao mesmo tempo em que um ruído de amplo espectro sintetizado, composto de ondas senoides igualmente espaçadas, foi introduzido no trato vocal dos cantores. Para tanto, um microfone e uma fonte acústica foram posicionados lado a lado, levemente em contato com os lábios dos sujeitos. O ruído de amplo espectro interage com o trato vocal do sujeito, tendo suas frequências amplificadas ou atenuadas de acordo com as ressonâncias do trato vocal, de modo que seu espectro apresenta picos exatamente nas frequências de ressonância. O sinal captado pelo microfone contém tanto a emissão da vogal pelo cantor, quanto o ruído de amplo espectro modificado pelo trato vocal. Os sujeitos fizeram emissões sustentadas das palavras em inglês *hard*, *hoard*, *who'd* e *heard*, correspondentes respectivamente às vogais /ɑ/, /ɔ/, /u/ e /ɜ/. Os cantores foram instruídos a executar a tarefa em intensidade confortável, e ela foi repetida em todas as notas da tessitura confortável dos sujeitos. Após as emissões com a voz cantada, os sujeitos falaram as palavras na sua fala habitual.

As sopranos exibiram F1 e F2 mais altos que as contraltos na fala, para quase todas as vogais. Não foi identificada diferença estatística nas frequências dos formantes de tenores e barítonos, nem na fala nem no canto. Todos os sujeitos apresentaram variações de F1 e F2 na voz cantada em função de f0. Nas notas abaixo

de B4 (494 Hz), houve maior variação proporcional nas frequências de F1 e F2. Nas notas mais agudas, essa variação caiu pela metade, provavelmente por haver menos possibilidades de sintonia dos formantes, por estarem os harmônios progressivamente mais afastados. Tal achado foi considerado uma indicação do uso deliberado da sintonia de formantes (Henrich et al., 2011).

Todas as sopranos utilizaram sintonia de F1 a H1 em alguma parte da escala em todas as vogais, com início em f0 mais graves na vogal de F1 mais baixo /u/. As contraltos também exibiram o uso da mesma sintonia, a partir das notas em que f0 atingiu a frequência do F1 medido nas suas emissões faladas. Foi identificada também a amplificação de H2 em algumas notas de algumas cantoras, pela sintonia de F2 nas sopranos e de F1 nas contraltos. Tenores e barítonos, por cantarem notas mais graves, que tem os harmônicos mais próximos, possuem mais possibilidades de sintonia. Houve grande variabilidade inter-sujeitos, mas a maioria apresentou, em alguma região de f0, a sintonia de F1 a H4, H3 ou H2, e vários exibiram a amplificação deliberada de H5, H4 ou H3 pela aproximação de F2. Apenas um dos tenores apresentou o uso esporádico, nas notas mais agudas, da estratégia de aumentar a amplitude de H1 pela sintonia de F1, característico das sopranos (Henrich et al., 2011).

Os autores também analisaram se houve preferência por posicionar os formantes logo acima dos harmônicos, como sugerem as teorias que preveem instabilidades com as variações de impedância do trato vocal. Os achados não confirmam a hipótese. Nas mulheres, F1 estava na maioria das emissões logo abaixo de H1, e F2 mostrou igual probabilidade de estar acima ou abaixo. Entretanto, considerada a margem de erro do método empregado, os autores não descartam a possibilidade de equívocos nas estimativas, uma vez que um pequeno desvio nas medidas poderia modificar a distribuição dos formantes ao redor do harmônico mais próximo (Henrich et al., 2011).

O objetivo do estudo de Echternach et al. (2014) foi analisar a influência das vogais nas configurações do trato vocal em diferentes registros, ao longo e acima da zona de passagem de tenores eruditos de diferentes subclassificações vocais. Os autores sugeriram haver um registro especial acima da passagem, uma vez que no canto lírico o registro modal não pode ser estendido ao agudo da mesma forma que é produzido no grave, e o uso do registro de falsete, por sua qualidade sonora, não

parece provável em tenores profissionais. Foram hipotetizadas duas soluções: o uso de um registro de falsete modificado, com aumento da adução glótica e/ou com modulações do trato vocal para seleção de harmônicos específicos; ou o uso de um registro modal modificado, com ajustes de trato vocal para melhorar interações fonte-filtro. Os autores designaram essas notas agudas como voz de palco acima da passagem (*stage voice above the passaggio*), por ser um nome neutro que não sugere o uso de um ou outro registro.

O estudo (Echternach et al., 2014) contou com quatro tenores profissionais, todos com fortes carreiras internacionais. Foram selecionados cantores de diferentes repertórios, da voz mais leve à mais dramática. Foram capturadas imagens de ressonância magnética e áudio dos sujeitos, durante emissões de escalas ascendentes nas vogais /a/, /e/, /i/, /o/, /u/ e /æ/, entre as notas C3 (262 Hz) e A3 (440 Hz), iniciadas no registro modal. A tarefa foi repetida duas vezes: na primeira foi pedido que os cantores trocassem para o registro de falsete entre as notas E3 (330 Hz) e F3 (349 Hz), e na segunda que cantassem como fariam no palco. Em cada *frame* das imagens, foram efetuadas as medidas: abertura de lábios e mandíbula, dorso de língua, protrusão de mandíbula, largura da faringe, posição da úvula e da laringe, e ângulo da laringe.

Os sujeitos exibiram claras modificações entre registros, mas houve muitas diferenças interindividuais. As vogais com F1 alto /a/ e /æ/ apresentaram maior abertura de lábios e mandíbula do que as vogais com F1 baixo /i/ e /u/. A abertura dos lábios e da mandíbula se mostraram dependentes, com alta correlação. Todos os sujeitos aumentaram a largura da faringe na vogal /a/ na mudança para a voz de palco. Nas vogais /e/ e /i/, os tenores leves apresentaram aumento da largura faríngea, enquanto os mais dramáticos apresentaram diminuição. Os autores sugerem que a diminuição da largura da faringe foi utilizada para aumentar F1. Em geral, a largura da faringe foi menor no falsete do que na voz de palco. Foi constatada consistência na maior abertura de lábios e mandíbula e na posição mais baixa da laringe para a “voz de palco” em relação ao falsete, independentemente da vogal. Outras articulações, assim como a largura da faringe, dependeram da vogal emitida e do tipo de repertório do cantor, com as vozes mais dramáticas exibindo movimentos articulatórios mais

amplos do que as mais leves. Os autores relacionaram as grandes diferenças entre os indivíduos ao fato de cantarem diferentes repertórios (Echternach et al., 2014).

O estudo de Echternach et al. (2016) teve como objetivo verificar a ocorrência de sintonia de formantes em tenores de ópera. Participaram dessa pesquisa os mesmo quatro cantores, atuantes em diferentes repertórios, que foram analisados no estudo anterior (Echternach et al., 2014), acerca das configurações do trato vocal na zona de passagem. Os sujeitos cantaram escalas ascendentes entre as notas C3 (262 Hz) e A3 (440 Hz), em intensidade confortável, nas vogais /a/, /e/, /i/, /o/ e /u/. Os cantores foram instruídos a manter, até a nota mais aguda, a qualidade vocal que utilizam no palco, ou seja, não trocar para o falsete. As mesmas tarefas foram repetidas em duas sessões: uma em que foram captados sinal de áudio e imagens do plano médio-sagital em ressonância magnética dinâmica; e outra em que foram gravados sinais de áudio, EGG, fluxo e pressão, os dois últimos por meio do uso de uma máscara de Rothenberg. Ambos os métodos foram utilizados para calcular as frequências de F1, F2 e F3. Da gravação da ressonância magnética, foi selecionado um *frame* para cada nota de cada vogal, no qual foi manualmente traçado o contorno do trato vocal, para estimar a função de área do trato vocal, gráfico que mostra a área da seção transversal do trato vocal, da glote até os lábios. A partir desta, foi calculada a função de transferência do trato vocal, na qual foram medidos os valores dos formantes. Da gravação de áudio da segunda sessão dos procedimentos, foi executado o procedimento de filtragem inversa, com auxílio do sinal do EGG, no *software* DeCap (Svante Granqvist, Suécia). A filtragem inversa consiste na estimativa manual das frequências dos formantes. Os dois métodos mostraram alta correlação entre os resultados.

Houve grandes diferenças inter-individuais nos resultados, possivelmente pelo fato de os sujeitos cantarem diferentes repertórios, além das diferenças anatômicas. Apenas na vogal /i/ foi identificada uma tendência clara da frequência de F1 subir paralelamente à f0; a vogal /a/ não exibiu esse comportamento, mas nas outras vogais foi identificada certa correlação entre a subida de f0 e F1, o que implica na mudança da qualidade das vogais. Apenas na vogal /i/ ocorreram aproximações de F1 a f0, e um sujeito exibiu o mesmo comportamento na vogal /u/. Nas vogais restantes, F1 permaneceu sempre bem mais alto que H1. Ao contrário do que prevê a teoria de

interação não-linear fonte-filtro, não foram observadas instabilidades na fonação quando ocorreram cruzamentos entre harmônicos e formantes, nem nos vários casos em que um formante estava logo abaixo do parcial mais próximo. Os autores compararam seus resultados às hipóteses levantadas por Neumann et al. (2005) e Schutte et al. (2005), acerca da sintonia de F2 a H3 como estratégia principal para tenores operísticos acima da zona de passagem: em apenas dois dos sujeitos ocorreu alguma aproximação de F2 a H3, na vogal /o/ na nota A3 (440 Hz), mas não foi identificada a subida de F2 junto à subida de H3 (Echternach et al., 2016).

O estudo (Echternach et al., 2016) não identificou a aplicação sistemática de sintonia de formantes entre os cantores analisados, e não confirmou a hipótese de que tenores dramáticos possam precisar recorrer à sintonia de formantes mais do que os tenores mais leves. Os autores chamam atenção ao fato de os métodos utilizados poderem interferir na emissão dos cantores: a máquina de ressonância magnética faz muito barulho, o que obriga os cantores a usarem fones de ouvido para atenuar o ruído e fornecer aos sujeitos *feedback* auditivo das suas vozes; a máscara de Rothenberg restringe o grau de liberdade do movimento da mandíbula; e a filtragem inversa, por ser um procedimento manual, é sujeita a gerar desvios das frequências dos formantes nas estimativas (os procedimentos de detecção de erros nas medidas são detalhados na publicação).

Titze e Palaparthi (2020) investigaram como diferentes valores de declínio espectral afetam o *loudness*, medido em “sones”, que foi comparado com os valores obtidos para o NPS. Sones é uma medida da percepção do *loudness*, e é sabido que não há proporção direta entre essa medida e os valores obtidos em decibéis para NPS, em função de o aparelho auditivo ter maior sensibilidade para determinadas frequências. Os valores em sonos são obtidos por meio da conversão das medidas em decibéis de NPS para medidas em “phons”, que são finalmente convertidas para sonos. Os valores em phons são dados pelas curvas isofônicas (*equal-loudness curves*), que levam em consideração a percepção da intensidade de sons simples em cada frequência da audição humana. Os valores em phons são convertidos para sonos, pois estes representam mais fielmente a percepção humana para sons complexos. Por definição, 1 sone equivale ao *loudness* de uma onda senóide de 1000 Hz com 40 dB de NPS.

O estudo (Titze e Palaparthi, 2020) foi baseado em simulações computacionais, nas quais foram testados os efeitos dos declínios espectrais de -3, -6, -9 e -12 dB/oitava sobre o *loudness*, em uma extensão de f_0 entre 125 e 1000 Hz, o que cobre a maior parte da produção vocal humana. Foram simulados apenas espectros glóticos com declínio espectral gradual, ou seja, não foram consideradas funções de transferência do trato vocal, a fim de isolar apenas a influência do ajuste glótico sobre o *loudness*.

Os resultados revelaram que na f_0 mais grave, 125 Hz, há muito contribuição dos harmônicos H5 a H10 para a percepção de *loudness*, por estarem em uma região de frequências às quais o ouvido humano é bastante sensível, em torno de 1000 Hz. Foi constatada pouca contribuição dos harmônicos acima de 2000 Hz. Tanto o *loudness* quanto o NPS diminuíram com os declínios espectrais mais acentuados, porém a perda foi maior em sones, de 12 para 3, enquanto o NPS caiu de 56 para 51 dB. Na frequência mais aguda, 1000 Hz, os harmônicos H2, H3 e H4 foram os que mais contribuíram para o *loudness*, por estarem na região de frequências de maior sensibilidade para o ouvido humano, entre 2000 e 3000 Hz. As quedas de *loudness* e NPS em função do aumento do declínio espectral foram semelhantes às da f_0 mais grave. Verificou-se que as medidas de *loudness* são muito sensíveis a pequenas mudanças de NPS, com um aumento menor do que 5 dB no NPS causando uma quadruplicação de *loudness* em sones (Titze e Palaparthi, 2020).

Os autores ainda utilizaram perfis de extensão vocal (PEV) de cinco mulheres, elaborados em um estudo anterior, a partir dos quais foram computados os declínios espectrais para fonações fracas e fortes em registro modal e de falsete. No registro modal, foi constatada muito pouca variação nos declínios das emissões fortes, que ficaram por volta de -3 dB/oitava, enquanto houve uma variação ampla e irregular na fonação fraca. Os autores associam o resultado à dificuldade de se manter fonações de muito baixa intensidade no registro modal, em que pode haver variações no contato das pregas vocais. Os resultados para o registro de falsete foram semelhantes entre as emissões fortes e fracas, o que indica que nesse registro o *loudness* não é tão controlado pelo declínio espectral da fonte (Titze e Palaparthi, 2020).

Os autores concluíram que variar o declínio espectral é uma forma muito efetiva de regular o *loudness* no registro modal. Eles alertam, porém, que a utilização apenas

da pressão pulmonar para alterar o declínio tem limites: em baixas pressões, próximo ao limiar de fonação, pode ser difícil manter consistência no intervalo do fechamento glótico; e em altas pressões, é difícil obter um declínio menos acentuado que do -12 dB/oitava sem recorrer ao aumento da adução glótica com uma fonação tensa. Os autores sugerem ser eficiente aqui utilizar as interações fonte-filtro, com os benefícios da alta inércia do trato vocal supraglótico (Titze e Palaparthi, 2020).

O estudo de Titze (2021) buscou verificar a influência da sintonia de um formante a um único harmônico na percepção de *loudness*, e comparar esses achados com os da publicação anterior, sobre os efeitos das variações do declínio espectral da fonte sobre o mesmo parâmetro (Titze e Palaparthi, 2020). Aqui também foram utilizadas simulações computacionais, nas quais foram alteradas as amplitudes de harmônicos individuais em 10, 20, 30 e 40 dB, em f_0 de 125 a 1000 Hz, com um espectro de ao menos doze harmônicos para cada f_0 . Foram então medidas as mudanças em *loudness* em sonos. O declínio espectral da fonte foi mantido, para todas as computações, em -6 dB/oitava.

A sintonia de H1 mostrou que o NPS é determinado principalmente pela energia da fundamental. Na computação com f_0 a 125 Hz, o NPS total subiu proporcionalmente aos incrementos de 10 dB na amplitude de H1, de 50 para 90 dB. O aumento do *loudness* não acompanhou o aumento de NPS linearmente, indo de 7 para 25 sonos, menos de dois dobramentos. Os autores salientam que favorecer a amplitude de H1 em uma frequência de fonação tão baixa não apresenta benefícios, e nem seria muito possível, uma vez que as ressonâncias do trato vocal estão bem acima de 200 Hz para todas as vogais. Na nota f_0 mais aguda, 1000 Hz, os resultados são diferentes: o *loudness* aumentou de 10 para 40 sonos, pois H1 está agora em uma região de maior sensibilidade auditiva. A estratégia de sintonizar F1 para amplificar H1 é frequentemente encontrada em sopranos operísticas. Porém, o aumento do NPS é ainda muito maior que o de *loudness* (Titze, 2021).

A sintonia de H2 é de interesse na região de f_0 por volta de 500 Hz, por ser utilizada por homens cantores de ópera e por mulheres *belters* no teatro musical. O aumento de *loudness* foi de 12 para 35 sonos, um pouco menor do que a sintonia da H1 com f_0 em 1000 Hz. A sintonia de H3 não é tão eficaz quanto a de H2, quando f_0 está em 500 Hz. Nessa frequência H3 está em 1500 Hz, um pouco acima da região

de grande sensibilidade auditiva em 1000 Hz. Uma f_0 um pouco mais baixa, por volta de 400 Hz, seria mais favorável ao aumento de amplitude em H3. Os autores relatam que o famoso tenor operístico Luciano Pavarotti frequentemente exibia um forte H3 quando cantava na região entre 400 e 500 Hz, porém também mantinha quase tão grande amplitude no H2 (Titze, 2021).

Em comparação aos ganhos de NPS com as variações de declínio espectral, aumentar a amplitude de H1 – e em certa medida, de H2 e H3 – gera mudanças ainda mais expressivas. Porém, o mesmo efeito não se verifica no aumento de *loudness* em sonos. Tal achado sugere que o NPS não é um bom índice para a percepção de *loudness*. Os autores concluem que mudanças no declínio espectral da fonte são mais eficientes do que a sintonia de harmônicos específicos para se obter uma ampla gama de dinâmicas musicais na voz cantada (Titze, 2021).

Herbst e Story (2022) conduziram um estudo teórico, com o objetivo de investigar como a intensidade sonora da voz radiada é influenciada por f_0 e por características da fonte glótica, por todo o espaço vocálico (gráfico que mostra as relações possíveis de F1 e F2 de um indivíduo) de adultos. Foi utilizado um modelo computacional simples, que considera igualmente todas as combinações de F1 e F2 dentro do espaço vocálico determinado. Foram modelados um espaço vocálico masculino e um feminino, baseados em um estudo anterior, em que foram parametrizadas funções de área do trato vocal, a partir de imagens de ressonância magnética. Foram simuladas três diferentes configurações glóticas, chamadas de “forte”, “intermediária” e “fraca”, respectivamente com -6, -12 e -18 dB/oitava de declínio espectral, em todos os semitons entre as notas C1 (65 Hz) e G5 (1568 Hz), f_0 extremas da extensão vocal humana. Foram comparadas as variações de diferença de NPS entre o espectro radiado e o da fonte. Os resultados apontaram que a transmissão de energia sonora pelo trato vocal torna-se progressivamente mais dependente de sintonias de F1 e F2 com algum harmônico, à medida que f_0 aumenta. Em relação à fonte, o ajuste “fraco”, com declínio espectral mais acentuado, sempre tem um output sonoro menos forte, e a sintonia de formantes se torna mais importante para aumentar a intensidade. Na fonte “forte”, por outro lado, com maior amplitude em harmônicos mais agudos, não prevalece a necessidade de sintonia de formantes. Por fim, o efeito da radiação labial, que adiciona em média 6 dB/oitava no espectro

radiado, se mostrou mais relevante em f_0 mais agudas, de modo que tende a haver um aumento de NPS à medida que f_0 sobe.

O estudo (Herbst e Story, 2022) concluiu haver fundamentalmente duas estratégias para maximizar a intensidade da voz no canto. Uma se baseia em otimizar a fonte sonora, “aplanando” o seu espectro, com o aumento da amplitude dos harmônicos agudos. Isso é alcançado por meio de ajustes na adução glótica ou nos registros; de ajustes da parte inferior do trato vocal supraglótico, para se beneficiar de interações não-lineares; ou ainda pelo aumento da P_{sub} . Essa é a estratégia de preferência para as notas mais graves, até o C3 (262 Hz). Os autores especulam que essa estratégia, em notas mais agudas, seja limitada pela anatomia laríngea, de forma que o cantor com uma mucosa mais espessa e maleável poderia empregá-la mais facilmente sem aduzir em excesso as pregas vocais. A segunda estratégia é a sintonia de formantes deliberada, que se torna necessária em f_0 mais agudas, acima do C3 (262 Hz), e mais relevante a partir do C4 (523 Hz). As duas estratégias podem ainda ser utilizadas em conjunto, aumentando as possibilidades da expressão artística. Dada a melhor aplicabilidade de cada técnica para diferentes regiões de f_0 , os autores concluem que diferentes classificações vocais podem se beneficiar mais de uma ou de outra.

4. Método

4.1. Preceitos éticos

O projeto desta pesquisa foi submetido e aprovado pelo Comitê de Ética em Pesquisa da Pontifícia Universidade Católica de São Paulo, sob nº 65785922.3.0000.5482 (Anexo 1). Os sujeitos voluntários do estudo foram informados detalhadamente sobre a pesquisa e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (Anexo 2).

4.2. Casuística

A amostra é intencional, uma vez que os sujeitos foram selecionados por serem representativos do objeto de pesquisa e por se adequarem aos objetivos específicos (Fontanella et al., 2008). Os sujeitos foram indicados por meio de contatos pessoais e profissionais do pesquisador. Participaram do estudo cinco tenores eruditos profissionais, atuantes em óperas.

Os critérios de inclusão foram: ser do sexo masculino; ser cantor profissional de ópera há pelo menos um ano; ser tenor; ter no seu repertório de performance, ou já ter estudado, a ária selecionada para o trabalho pelo pesquisador; e ausência de queixa vocal. No dia marcado para a coleta, os cantores voluntários foram contatados via aplicativo de mensagens para conferir se estavam em bom estado de saúde para o procedimento. Quando não estavam, a coleta foi remarcada para outro dia.

4.3. Seleção da ária e das vogais

No presente estudo, foi selecionada uma ária de ópera para ser executada pelos cantores. A ópera é um gênero musical que toma forma em um drama musicado, em que os cantores representam os papéis e cantam, acompanhados por uma orquestra. Além das partes cantadas, a ópera pode ter diálogo falado. As porções cantadas podem ser divididas em recitativo e ária. No recitativo se desenvolvem os fatos da história, logo, a melodia tende a ser mais rápida e a seguir a entoação e o ritmo natural das palavras na fala, e a instrumentação tende a ser mais esparsa. A ária, por sua vez, é mais melodiosa e fluida, e representa os pensamentos e emoções dos personagens, suas reações aos acontecimentos da trama (Bennett, 1998).

Foi escolhida a ária de ópera “*Un’aura amorosa*” (partitura no Anexo 3), do personagem Ferrando, no primeiro ato da obra *Così fan tutte*. Com música composta por Wolfgang Amadeus Mozart sobre libreto escrito por Lorenzo Da Ponte, o trabalho teve sua estreia em 1790. A obra apareceu em 12º lugar nas listas das óperas com maior número de produções e de encenações no mundo entre 2018 e 2022, em pesquisa efetuada dia 6 de fevereiro de 2023 na base de dados do *site* Operabase. O portal documenta as atividades relacionadas às produções de ópera desde 1996, com registros de mais de 3 mil teatros do mundo inteiro (Operabase, 2023).

Para a execução de algumas das tarefas, foram selecionadas as vogais /a/, /e/, /i/, /o/ e /u/, por elas compreenderem valores altos, médios e baixos de frequências de F1 e F2 (conferir a Figura 6). As vogais /i/ e /u/ tem os valores mais baixos de F1 (em torno de 300 Hz), enquanto a vogal /a/ tem o F1 mais alto (entre 700 e 850 Hz). Em relação a frequência de F2, /u/ tem a mais baixa (por volta de 850 Hz) e /i/ a mais alta (2500 Hz), enquanto /a/ apresenta um valor intermediário (1200 Hz) (Denes e Pinson, 2016)¹⁴. As vogais /e/ e /o/ apresentam valores intermediários de F1. As vogais /i/ e /u/ são de especial dificuldade para o canto: por terem F1 baixo, elas propiciam mais oportunidades para que ocorram interações não-lineares entre fonte e filtro, inclusive interações de Nível 2 (descontinuidades na fonação quando H1 ultrapassa a frequência de F1), em notas relativamente não tão agudas (Titze, 2008a). As notas mais agudas da ária selecionada para o estudo, por exemplo, são A3 (440 Hz), razoavelmente acima das frequências usuais de F1 das vogais /i/ e /u/, e a ária exige a emissão da vogal /i/ nessa nota.

As vogais selecionadas ainda compreendem posições contrastantes dos articuladores, fato diretamente relacionado aos valores diferentes de F1 e F2. A vogal /a/ apresenta o corpo da língua projetado para a parede posterior da faringe. A vogal /i/ é caracterizada pelo corpo da língua projetado para o palato duro, lábios estirados e tendência à laringe mais alta. E a vogal /u/ apresenta corpo da língua projetado para o véu palatino e ponta levemente retraída, lábios arredondados e tendência à laringe

¹⁴Os valores de Denes e Pinson (2016) são médias de falantes do sexo masculino e feminino de língua inglesa. Os formantes das vogais variam entre idiomas, em função de suas particularidades articulatorias, mas os valores máximos e mínimos possíveis para cada formante não sofrem grandes alterações, e são dependentes principalmente da anatomia do falante. A ordem de vogais do F1 mais baixo ao mais alto não se altera, assim como a do F2.

mais baixa. O posicionamento dos articuladores na vogal /e/ é intermediário aos das vogais /i/ e /a/, e os da vogal /o/ aos das vogais /u/ e /a/ (Sundberg, 2018).

4.4. Instrumentos

O procedimento consistiu na gravação do sinal de áudio dos cantores e foi realizado em estúdios acusticamente tratados. A captação do sujeito T1 ocorreu no laboratório de rádio da Faculdade de Filosofia, Comunicação, Letras e Artes (FAFICLA), na PUC-SP, enquanto os demais sujeitos foram gravados em um estúdio particular em Porto Alegre. Em ambos os estúdios foi utilizado o mesmo microfone condensador omnidirecional, com curva de resposta de 18 Hz a 20000 Hz plana, modelo EMM-6 (Dayton Audio, EUA), conectado a um computador com sistema operacional Windows 10 (Microsoft Corporation, Estados Unidos), por meio de uma mesa de áudio MGP32X (Yamaha Corporation, Japão) no estúdio da PUC-SP, e pela interface 2408 (MOTU, Estados Unidos) no estúdio privado. As gravações foram captadas pelo *software Sound Forge Pro* versão 11 (Magix Software GmbH, Alemanha) no estúdio da PUC-SP, e pelo REAPER versão 6.78 (Cockos Incorporated, Estados Unidos) no estúdio privado. Os arquivos de áudio foram salvos no formato .wav com 48000 Hz de taxa de amostragem. Um teclado musical foi disponibilizado para referência de tonalidade e para o aquecimento dos cantores, além de uma estante de partituras colocada em frente aos cantores para consulta ao material impresso.

4.5. Procedimentos

Foi realizado um estudo piloto para verificar os procedimentos e determinar a melhor ordem e o número de repetições das tarefas, além de aspectos como o posicionamento do cantor e a distância microfone-lábios. A ária exigida no estudo, por ser critério de inclusão, foi informada aos cantores no primeiro contato. Confirmada a participação dos sujeitos, lhes foi enviado um questionário de identificação (Anexo 4), para ser respondido *online* via Google Forms. As questões foram adaptadas do Questionário de Identificação de Aspectos de Saúde e Bem-estar Vocal, formulado por Andrada e Silva e Duprat (2014). O instrumento é aplicado pelos autores nos atendimentos no Ambulatório de Artes Vocais da Santa Casa de São Paulo, do qual são coordenadores.

No dia da coleta, antes do início do procedimento, os sujeitos leram e assinaram o TCLE. O pesquisador descreveu brevemente o procedimento do início até o fim e sanou quaisquer dúvidas dos sujeitos. Os cantores tiveram então até 10 minutos para fazer livremente seu aquecimento vocal corriqueiro. Para tal, puderam utilizar o teclado musical. Não houve qualquer interferência do pesquisador no aquecimento.

Após o aquecimento, os cantores foram posicionados em pé em frente ao microfone, de forma que esse ficasse na altura de seus lábios a 20cm de distância, medidos pelo pesquisador com uma trena. Sempre que considerado necessário durante o procedimento, essa medida foi novamente conferida.

Para ajuste de ganho na mesa de som, os cantores foram instruídos a emitir a vogal /a/ em três intensidades diferentes (média, fraca e forte), e a cantar alguns trechos de maior intensidade da peça musical. Dessa forma, foi evitado o *clipping* do sinal e a distorção do áudio das gravações.

Para posterior calibração do sinal, foi reproduzido em um alto falante o áudio de uma onda senoidal de 1000 Hz, previamente sintetizada no *software* Praat versão 6.2.23 (Paul Boersma e David Weenink, Holanda). Simultaneamente, foi posicionado ao lado do microfone um medidor de NPS, cuja leitura da intensidade da onda senoide foi anunciada pelo pesquisador, juntamente ao nome do sujeito, à data e à distância do microfone (Švec e Granqvist, 2018).

A gravação foi iniciada com um trecho de fala espontânea de aproximadamente 1 minuto, em que os sujeitos falaram livremente sobre si, e uma contagem regressiva de 20 a 1. Na sequência, foram executadas quatro tarefas:

(1) Execução das sentenças-veículo: “Diga pápa baixinho”, “Diga pêpe baixinho”, “Diga pípi baixinho”, “Diga pôpo baixinho” e “Diga púpu baixinho”. As sentenças-veículo permitem que se analise as vogais faladas dentro de um mesmo contexto fonético, de forma a reduzir possíveis influências de coarticulação dos fonemas que ocorrem antes e depois da vogal.

(2) Emissão sustentada, em tom confortável habitual, com qualidade de voz cantada, das vogais /a/, /e/, /i/, /o/ e /u/. Foi pedido aos sujeitos que cantassem as vogais da forma que o fariam normalmente em uma música.

(3) Execução da ária. Foi selecionado um trecho de aproximadamente 2 minutos de duração. Foi disponibilizada para consulta aos cantores a partitura da peça. O trecho executado corresponde, na partitura no Anexo 3, do início da ária até o final da segunda página, onde na letra em italiano se lê pela última vez “*bisogno non ha*”.

(4) Execução de duas escalas diatônicas ascendentes de 8 tons em dó maior, entre as notas C2 (131 Hz) e C3 (262 Hz), e as notas C3 (262 Hz) e C4 (523 Hz), nas vogais /a/, /e/, /i/, /o/ e /u/. Segundo Miller (2008), a nota C4 (523 Hz) é uma das notas mais agudas a ser esperada de um tenor de ópera. Ela está próxima ao limite da sua extensão fonatória (Miller, 1993). Considerou-se então essa nota uma referência razoável a ser utilizada como limite agudo da tarefa. Os cantores foram instruídos a manter na execução das escalas a qualidade vocal utilizada na peça musical. Como o cantor no estudo piloto inicialmente sacrificou a qualidade vocal para “manter a pureza e a inteligibilidade das vogais”, por ter entendido ser esse o intuito da tarefa, optou-se por instruir expressamente os sujeitos a priorizarem a qualidade vocal que utilizariam no palco, modificando a articulação das vogais, caso considerassem necessário ao longo da escala, da mesma forma que fariam se tivessem que cantar em uma peça musical essas vogais, nessas notas.

As tarefas foram repetidas quando a execução foi considerada insatisfatória pelo pesquisador ou pelos sujeitos, entre uma e três vezes para cada emissão.

4.6. Análise dos dados

As gravações foram inicialmente processadas e analisadas no *software* Praat. Análises posteriores foram realizadas no *software* Sopran versão 1.0.28 (Tolvan Data, Suécia). Foram extraídos:

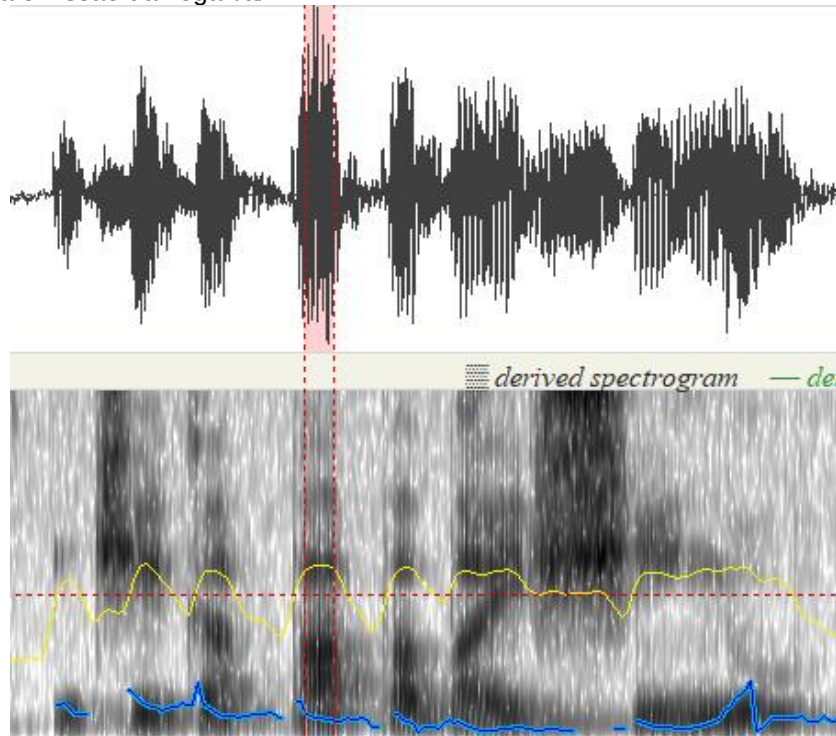
(1) valores das frequências de F1 e F2 das vogais faladas nas sentença-veículo (sílabas tônicas da palavra alvo) e das vogais cantadas nas escalas musicais;

- (2) amplitudes de H1, H2, H3, H4 e do pico na região do formante do cantor das vogais cantadas nas escalas;
- (3) f0 das vogais faladas nas sentenças-veículo (sílabas tônicas da palavra alvo) e das vogais cantadas nas escalas;
- (4) NPS das vogais faladas nas sentenças-veículo (sílabas tônicas da palavra alvo) e das vogais cantadas nas escalas, da fala espontânea e da execução da peça musical;
- (5) EMLT e correspondente declínio espectral da fala espontânea e da peça musical.

Para as análises das vogais cantadas nas escalas, foram selecionadas notas alternadas, ou seja, C2 (131 Hz), E2 (165 Hz), G2 (196 Hz), B2 (247 Hz), D3 (294 Hz), F3 (349 Hz), A3 (440 Hz) e C4 (523 Hz). Optou-se pela seleção de notas alternadas a fim de reduzir o número total de emissões de vogais a serem analisadas, que seria de 400, para 225. As notas consideradas compreendem ampla tessitura, do grave até o agudo dos cantores, e estão em intervalos regulares e próximas, o que ainda permite a análise de variações dos parâmetros ao longo da escala.

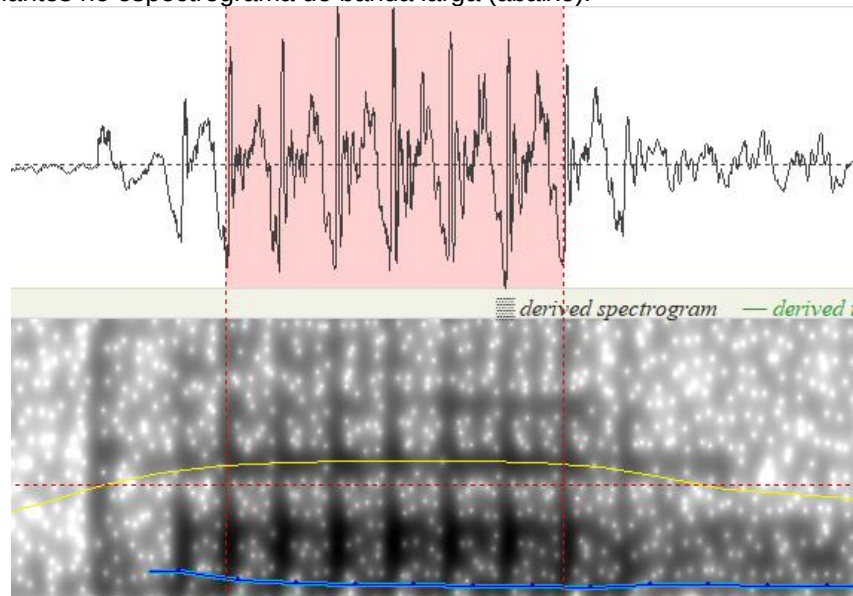
Para as extrações dos valores das frequências de F1 e F2 das vogais faladas nas sentenças-veículo e das notas selecionadas nas escalas musicais, os arquivos de áudio foram primeiramente editados e convertidos para a taxa de amostragem de 11000 Hz no Praat. Para cada emissão de voz falada, foi selecionado um trecho de articulação e fonação estáveis do meio da vogal da sílaba tônica da palavra alvo, excluindo a articulação das consoantes antes e após a vogal falada. Para tanto, foi gerado um espectrograma de banda larga (por meio dos segundos comandos no Praat: menu *Spectrogram* > *Spectrogram settings*, parâmetro *Window length*=0,005 s), que mostra as faixas de formantes. Os critérios para a área selecionada foram a presença de formantes bem definidos e com frequências centrais estáveis, e forma de onda regular, exemplificados nas Figuras 7 e 8. O início e o final da seleção foram alinhados a um cruzamento de zero na forma de onda, como ilustrado na Figura 9. A duração média das seleções das vogais faladas foi de 0,073 s.

Figura 7 – Captura de tela do Praat, mostra a frase “Diga pápa baixinho”: no painel superior a forma de onda, no inferior o espectrograma de banda larga. Área vermelha corresponde a porção selecionada, em que ocorre a emissão da vogal /a/.



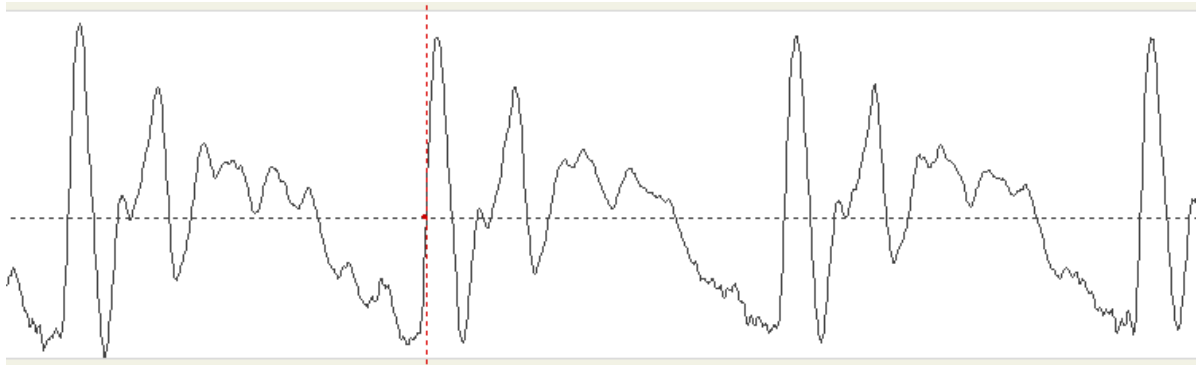
Fonte: Autor.

Figura 8 – Captura de tela do Praat, zoom na vogal /a/ da imagem acima, mostra a seleção do trecho a ser analisado, em vermelho, baseado na regularidade dos pulsos na forma de onda (acima) e das faixas dos formantes no espectrograma de banda larga (abaixo).



Fonte: Autor.

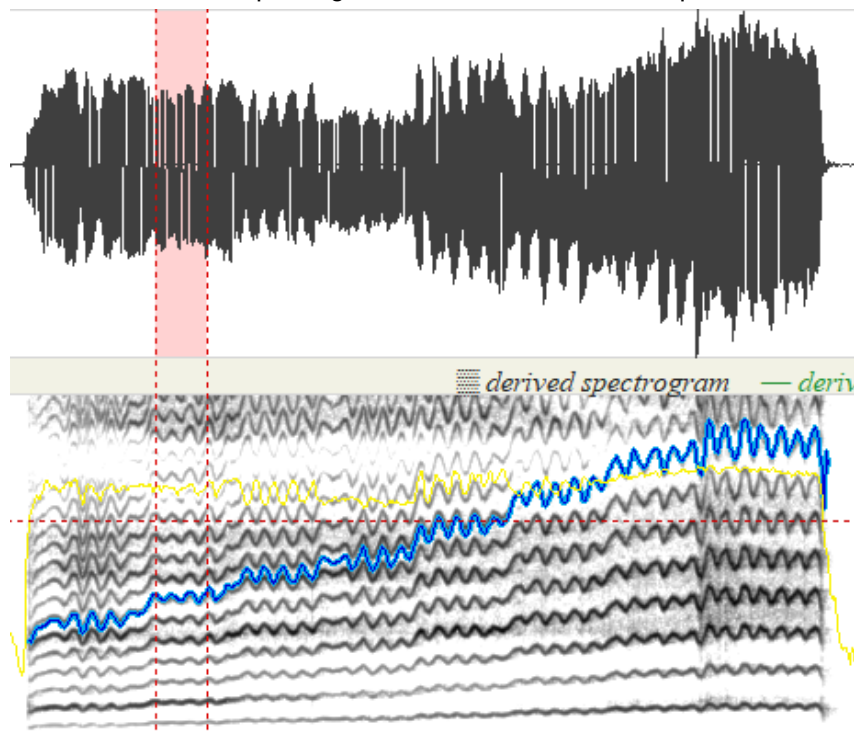
Figura 9 – *Zoom* nos pulsos da forma de onda, possibilita alinhar o ponto de seleção a um cruzamento de zero, demarcado aqui pela interseção entre a linha de seleção (linha pontilhada vertical vermelha) e o eixo zero da amplitude do sinal de áudio (linha horizontal pontilhada preta).



Fonte: Autor.

As porções analisadas das notas das escalas foram determinadas pela regularidade da fonação e de f_0 , identificada no espectrograma de banda estreita no Praat, que mostra cada harmônico individualmente (comandos no Praat: menu *Spectrogram > Spectrogram settings*, parâmetro *Window length*=0,05 s). Ainda foi utilizada como referência a estabilidade do vibrato (Figura 10). Foram desconsiderados nas seleções o início e o final de cada nota, a fim de evitar possíveis variações decorrentes da sua articulação com as notas anteriores e posteriores. A duração média dos trechos de voz cantada foi de 0,383 s.

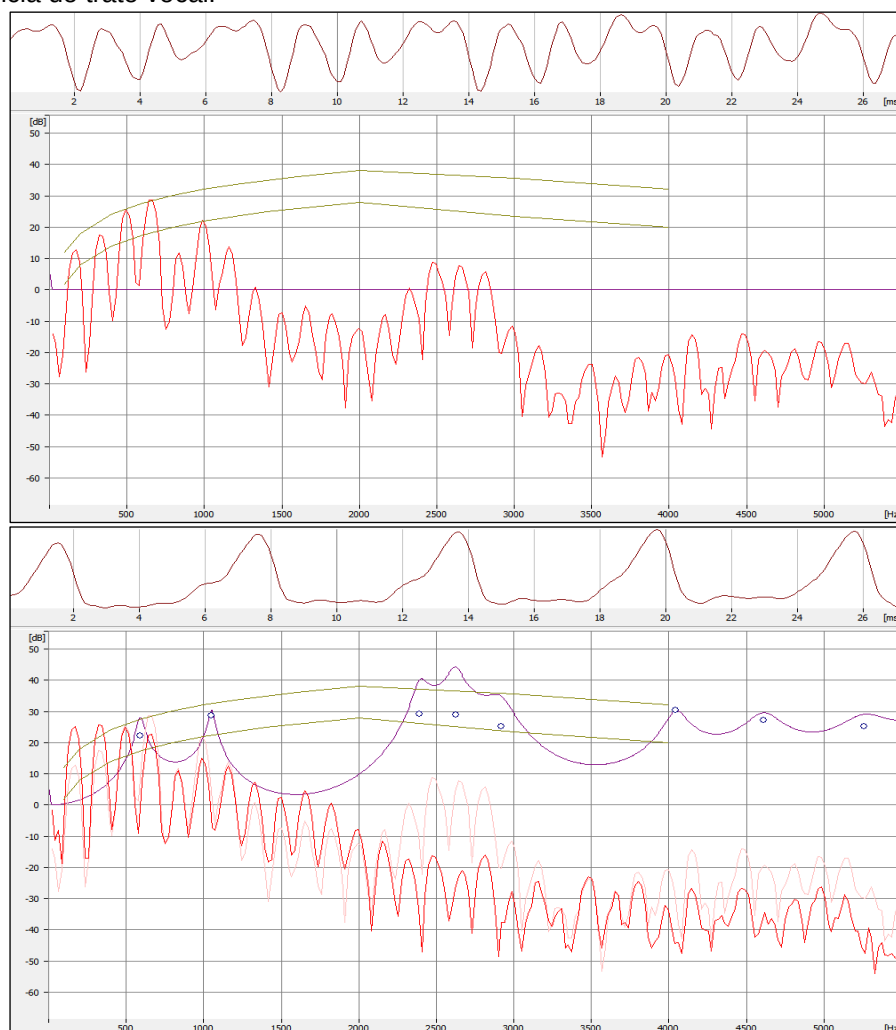
Figura 10 – Captura de tela do Praat, em vermelho a seleção do trecho da nota D3, considerando a estabilidade de f_0 e do vibrato no espectrograma de banda estreita, no quadro de baixo.



Fonte: Autor.

Os valores das frequências de F1 e F2 foram estimados no *software* Sopran, por meio do módulo DeCap, no qual é executado o procedimento de filtragem inversa. As medidas foram feitas em um ponto no meio das emissões. O *display* do *software* foi configurado para mostrar a forma de onda da fonte glótica, o espectro radiado e o espectro filtrado. O método consiste em manualmente estimar as frequências e largura de banda dos formantes, posicionando-os com o *mouse* nos pontos desejados. O *software* aplica então um filtro com frequências contrárias àquelas do trato vocal da emissão analisada, cancelando os efeitos do filtro do trato vocal sobre o som da fonte, revelando dessa forma o espectro e a forma de onda da fonte glótica (Salomão, 2008). A Figura 11 ilustra a tela do *software* antes e depois do processo de filtragem.

Figura 11 – Tela do módulo DeCap do *software* Sopran, no qual é executado o procedimento de filtragem inversa. A imagem acima apresenta o áudio antes do procedimento, assim, a forma de onda e o espectro correspondem ao emitido pelo cantor. Na imagem abaixo, após a filtragem, observa-se a forma de onda da fonte glótica, com uma fase fechada bem definida, e seu respectivo espectro, com declínio espectral regular. Atrás do espectro da fonte glótica, é visível ainda em menor intensidade o espectro radiado. Os círculos pequenos correspondem aos formantes, e a curva sobre eles à função de transferência do trato vocal.



Fonte: Autor.

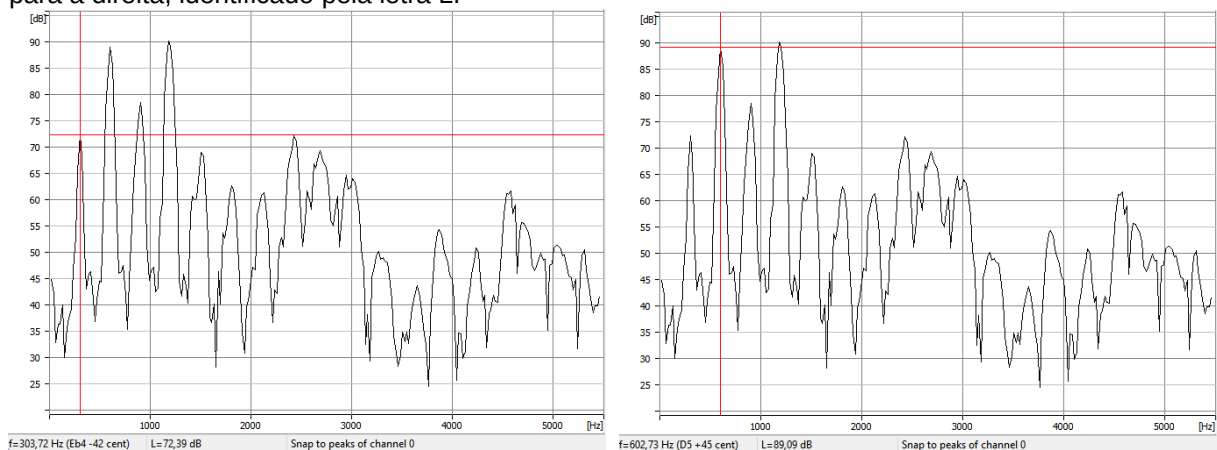
A estimativa dos formantes por meio de filtragem inversa manual foi selecionada pela pouca confiabilidade de extração automática de formantes em f_0 agudas (Salomão e Sundberg, 2009), pois os harmônicos são mais distantes, ou seja, há pouca informação para a extração por *software*. Esse método exige do pesquisador atenção e conhecimentos acerca dos formantes da vogal cantada, das características espectrais e da forma de onda da fonte, que devem ser coerentes com a qualidade vocal da emissão. Por outro lado, o método permite ajustes finos para se chegar ao resultado esperado.

Por se tratar de um método manual, são necessários cuidados. A correta estimativa dos formantes tem como parâmetros: forma de onda regular, com fase fechada definida e o mínimo possível de ondulações; e espectro da fonte com declínio espectral o mais regular possível (ver Figura 11). No procedimento é normalmente utilizado também o sinal de EGG (Björkner, 2008; Salomão e Sundberg, 2009; Borch e Sundberg, 2011): a derivada do EGG mostra com maior precisão o momento do fechamento glótico, o que oferece maior certeza do resultado obtido na forma de onda.

A filtragem inversa é muito empregada na análise do comportamento fonatório isolado do filtro, por possibilitar a verificação precisa do fluxo de ar transglótico. Este trabalho não se propõe a avaliar parâmetros glóticos, nem a detalhar com precisão o filtro. Dessa forma, considera-se que a não utilização do EGG nas gravações não traz grande prejuízo aos resultados apresentados. Limita-se aqui a se estimar as frequências de F1 e F2, para identificar a sintonia de formantes. Herbst e Story (2022) apontam que F1 e F2 são as ressonâncias mais importantes na produção das vogais e se mostram centrais na sintonia de formantes no canto. Para maior confiabilidade nos resultados, além das referências supracitadas (fase fechada definida e espectro glótico com declínio regular), foi utilizado o *software* sintetizador de voz Madde versão 3.0.0.2 (Tolvan Data, Suécia). Os valores obtidos para os primeiros seis formantes na filtragem inversa foram inseridos no *software*, e a voz sintetizada foi perceptualmente avaliada pelo pesquisador. Quando a qualidade da vogal ou o timbre do som produzido pelo *software* foram muito diferentes da emissão analisada, ajustes foram feitos nos formantes no Sopran até que o resultado fosse considerado auditivamente satisfatório.

A Figura 12 ilustra as leituras das amplitudes dos dois primeiros harmônicos de uma mesma emissão. Uma vez que o espectro da fonte glótica apresenta o declínio gradual de energia com a aumento da frequência dos parciais, são os harmônicos mais graves, por terem maior amplitude, os mais propícios para a sintonia com F1 e F2, com o objetivo de aumentar a intensidade da voz radiada (Herbst e Story, 2022). Por isso, foram medidas as amplitudes de H1, H2, H3 e H4, além do pico na região do formante do cantor, das notas das escalas musicais.

Figura 12 – Tela do espectro de banda estreita no *software* Sopran. À esquerda a leitura da amplitude de H1, e à esquerda a de H2. O cruzamento das linhas vermelhas marca o pico de cada harmônico. O valor da intensidade pode ser identificado na barra inferior da tela, no segundo campo da esquerda para a direita, identificado pela letra L.



Fonte: Autor.

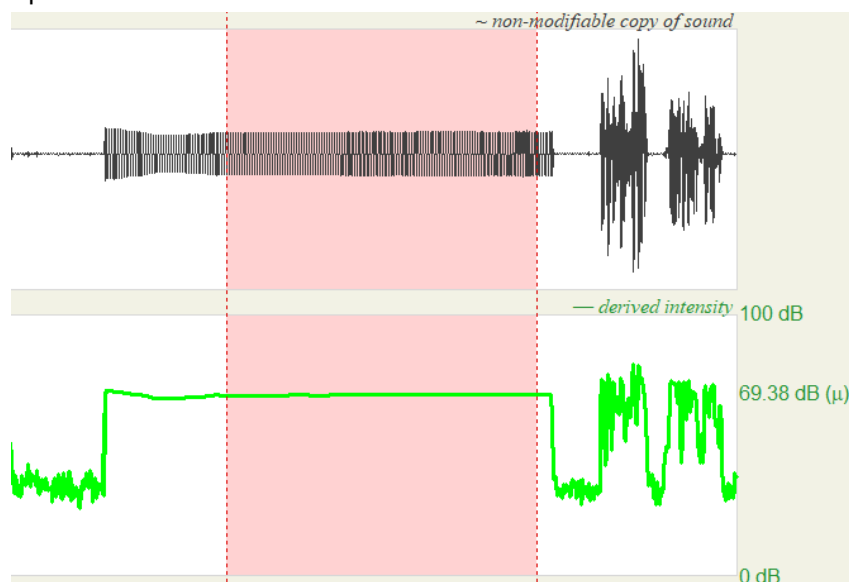
As extrações foram feitas em espectros médios da seleção da porção estável das emissões (conforme a Figura 10). A seleção arbitrária de um ponto no meio da emissão estaria sujeita a possíveis micro variações no ajuste vocal e às oscilações de f_0 decorrentes do vibrato (Schutte et al., 2005). No *software* Sopran, os valores foram medidos no espectro de banda estreita (*Analysis > Spectrum*), configurações de FFT com janela de Hamming e largura de banda de 50 Hz. O item *Selection* foi marcado nas opções *LTAS*, para gerar o espectro médio da parte selecionada do áudio.

As extrações de f_0 das vogais faladas nas sentenças-veículo e das vogais cantadas nas escalas foram feitas no *software* Praat (menu *Pitch > Get pitch*), nos mesmos trechos que foram utilizados para as estimativas de F1 e F2 e para as medições das amplitudes dos harmônicos e do formante do cantor.

As medidas do NPS foram feitas no Praat (menu *Intensity > Get intensity*), com referência no sinal de calibração de 1000 Hz. Uma vez que a intensidade do sinal é

conhecida, da leitura do medidor de NPS durante sua reprodução, pode-se calcular a diferença entre a intensidade identificada pelo medidor e aquela identificada pelo Praat, a partir da seleção de uma porção estável da forma de onda do sinal de calibração. A Figura 13 exemplifica esse procedimento. A leitura do *software* é de 69,38 dB, enquanto a do medidor de NPS foi de 87 dB, portanto 17,62 dB acima. Esse valor pode ser somado ao da leitura de intensidade de qualquer porção do arquivo de áudio para a obtenção da intensidade real desse trecho, considerada sempre a distância do microfone. O NPS das vogais foi medido nas mesmas seleções utilizadas nos procedimentos acima descritos. O NPS da fala espontânea e da execução da ária foi medido nos trechos integrais das emissões.

Figura 13 – Captura de tela do Praat, mostra o sinal de calibração: no painel superior, a forma de onda; no inferior, a linha de intensidade. Para a calibração, foi selecionada a área em vermelho, com forma de onda com amplitude estável e linha de intensidade constante.



Fonte: Autor.

Para a extração dos EMLT da fala espontânea e da execução da peça musical, os trechos foram salvos em arquivos separados no Praat, e então normalizados e salvos no formato .wav com taxa de amostragem convertida para 22000 Hz no *software Audacity* versão 2.4.2 (*The Audacity Team*, EUA). A normalização modifica a amplitude do sinal de áudio sem modificar as amplitudes relativas entre os harmônicos, tendo como referência um valor máximo padronizado para todos os arquivos. Dessa forma, as diferentes gravações ficam com intensidades semelhantes sem alterar o perfil dos espectros, o que permite que eles sejam comparados de forma equivalente.

Os arquivos normalizados foram abertos no Praat, onde foram gerados os EMLT. Com o arquivo de áudio selecionado na janela *Praat objects*, foi utilizada a opção *Analyse spectrum > To Ltas (pitch corrected)...*, que, por meio de uma série de análises automatizadas, elimina possíveis erros no espectro decorrentes da influência de f_0 . O EMLT resultante é representativo das ressonâncias do trato vocal e do envelope do espectro da fonte glótica (Boersma e Kovacic, 2006). Os parâmetros utilizados para gerar os espectros foram: *Minimum pitch*=75 Hz e *Maximum pitch*=600 Hz (região de frequências onde o *software* analisa f_0), e *Maximum Frequency*=5000 Hz (frequência máxima do espectro a ser analisada); e para desenhar os gráficos: *Pitch range*=0 a 5000 Hz, e *Power range*=-10 a 60 dB. Na janela *Praat objects*, com o EMLT selecionado, o declínio espectral foi obtido pelo comando *Query > Get slope...*, parâmetros *Low band*=0-1000 Hz e *High band*=1000-5000 Hz.

Todos os valores obtidos nas análises foram inseridos em planilhas no *software* Excel versão 2312 (*Microsoft Corporation*, Estados Unidos).

Os valores de F1 e F2 foram plotados em gráficos no Excel em função de f_0 , para analisar sua evolução ao longo das escalas. Nos gráficos foram ressaltadas as relações dos formantes com os primeiros harmônicos de cada nota. Foram elaborados ainda, a partir de F1 e F2, para cada cantor, espaços vocálicos das vogais faladas, da nota mais grave e da mais aguda da escala, C2 (131 Hz) e C4 (523 Hz). O espaço vocálico é um gráfico que relaciona as frequências de F1 com as de F2 das vogais de um falante. As coordenadas das vogais, conectadas por linhas, delimitam uma área correspondente às possibilidades de ajustes de F1 e F2 do sujeito. Quanto maior for a área, maior é a variabilidade de F1 e F2 entre as vogais, ou seja, maior o grau de extensão articulatória que o falante é capaz de realizar (Titze et al., 2017).

As amplitudes de H1, H2, H3, H4 e do pico do formante do cantor foram também plotadas no Excel em gráficos em função de f_0 , o que permite a análise da variação dos harmônicos dominantes no espectro, ou seja, os de maior amplitude, nota a nota.

A ocorrência de sintonias de formantes foi levantada com base no cálculo da diferença entre os valores de frequências dos formantes e dos harmônicos (frequência do formante menos frequência do harmônico). Foram consideradas as sintonias de formantes mais frequentemente descritas em tenores na literatura – o primeiro

formante com um dos dois primeiros harmônicos, e o segundo formante com o terceiro ou o quarto harmônicos (Neumann et al., 2005; Schutte et al., 2005; Echternach et al., 2016). As estratégias foram quantificadas por vogais, por sujeitos e por três diferentes zonas de f_0 : abaixo da zona de passagem, na zona de passagem, e acima da zona de passagem, definida como as notas entre B2 (247 Hz) e F3 (349 Hz) (Miller, 2019). Foi considerada a ocorrência de sintonia quando um formante estava no máximo 50 Hz abaixo ou acima do harmônico mais próximo (Henrich et al., 2011).

Os valores de NPS foram comparados entre fala espontânea e execução da ária, e entre as vogais faladas e cantadas.

Foi analisado o perfil geral dos EMLT, identificando bandas de frequência com maior e menor amplitude. As comparações foram feitas com base na sobreposição dos gráficos da fala espontânea e da peça musical de cada sujeito. Os EMLT e os declínios espectrais foram comparados entre fala espontânea e execução musical de cada cantor, e entre os cantores.

Considerando o grande número de resultados, optou-se por deixar os áudios: contagem regressivas, vogais cantadas sustentadas em tom habitual, e medidas de f_0 , extraídas das vogais faladas das sentenças-veículo e das vogais cantadas nas escalas, para serem analisadas em um outro trabalho.

5. Resultados

A Tabela 1 contém os dados de identificação dos cinco sujeitos participantes da pesquisa. Todos são tenores profissionais, atuantes há entre 10 e 34 anos em palcos de ópera, e tinham entre 31 e 54 anos de idade na data do procedimento de gravação.

Tabela 1 – Identificação dos participantes.

Sujeito	Idade (anos)	Tempo de profissão (anos)	Classificação vocal
T1	42	20	Tenor lírico ligeiro
T2	31	10	Tenor lírico
T3	42	26	Tenor lírico spinto
T4	54	34	Tenor lírico ligeiro
T5	53	30	Tenor lírico ligeiro
Média	44,4	24	

Fonte: Autor.

Do questionário de identificação e aspectos de saúde e bem-estar vocal (Anexo 4), foram destacadas as respostas dadas ao último item, no qual foi solicitado aos sujeitos que comentassem livremente acerca de modificações e ajustes que consideravam necessários à produção das notas agudas. A Tabela 2 apresenta as respostas, nas quais foram identificados cinco aspectos levantados pelos cantores: (1) qualidade vocal, citada por todos os participantes; (2) ajustes laríngeos, mencionados por quatro sujeitos; (3) adaptações corporais e (4) ajustes supraglóticos, observados por três sujeitos; e (5) outras percepções subjetivas, descritas por dois sujeitos.

Tabela 2 – Respostas dos tenores (T) em relação ao questionamento dos ajustes necessários para a produção das notas agudas.

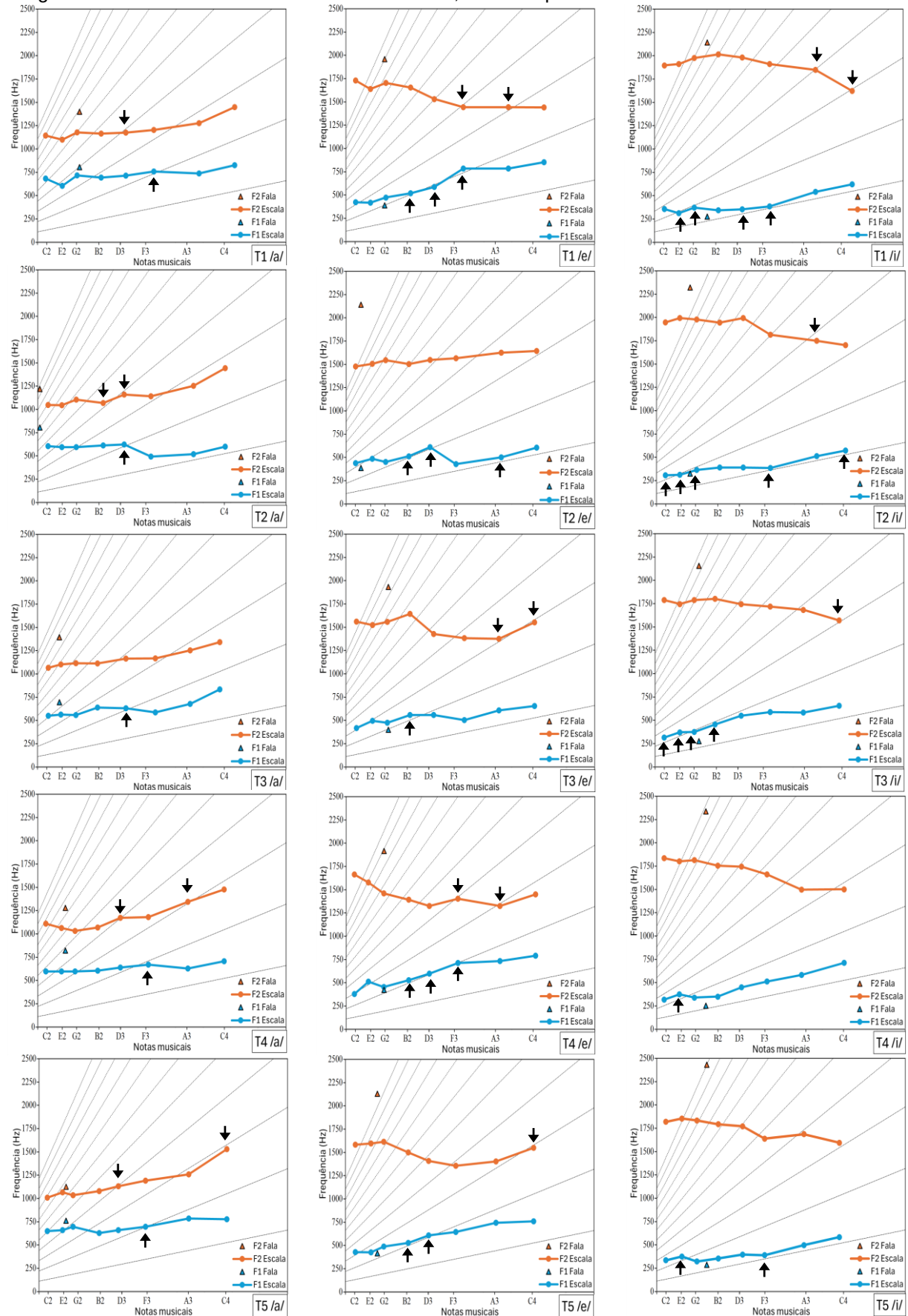
Aspectos	T1	T2	T3	T4	T5
Qualidade vocal	Qualidade do som	Modificar as vogais	Manter a qualidade sonora	Observar condução e cor da voz	“Ligar a voz de peito, para evitar agudos minguados”
Ajustes laringeos	Equilíbrio muscular	Manter a adução das pregas vocais	“Fechamento pequeno das cordas vocais, menor ponto de fonação”		Usar os TA, subir sem “desligar o motor”
Adaptações corporais		Maior empenho físico para manter o equilíbrio entre costelas abertas e a contração abdominal		Apoio diafragmático; Bem relaxado	Força no “core”; Equilíbrio entre força e relaxamento
Ajustes supraglóticos		Abrir a boca; Língua mais neutra; Elevar lábio superior, mostrando os caninos; Lábios mais arredondados	“Abertura da laringe”; Mais bocejado		Língua baixa; Boca relaxada e bem aberta
Outras percepções subjetivas			Sensação de maior pressão subglótica; Pensar em algo profundo; Pensar maior e menor ao mesmo tempo; Ato mental, o corpo se adapta		Pensar na voz para baixo, em vez de espremida no rosto, para manter o espaço para os agudos

Fonte: Autor.

O Quadro 1 apresenta gráficos do primeiro e do segundo formantes (F1 e F2) das vogais /a/, /e/ e /i/, faladas (sentenças-veículo) e cantadas (escalas), dos cinco sujeitos; no Quadro 2 estão os mesmos resultados para as vogais /o/ e /u/. Os valores das frequências dos formantes estão plotados em função das notas musicais, o que permite que se observe sua evolução com o aumento de f_0 .

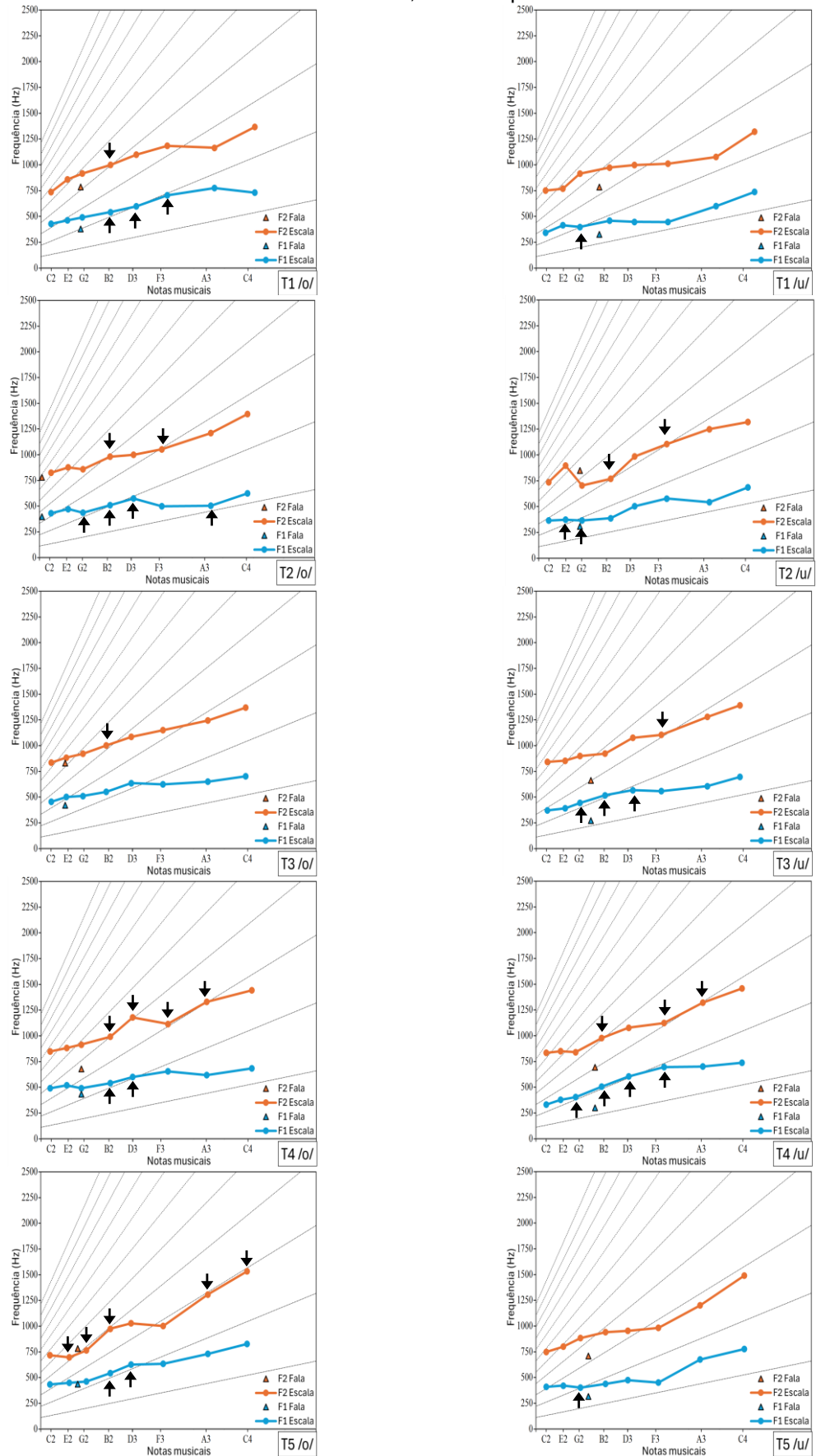
No fundo do gráfico existem linhas diagonais cinzas, que correspondem aos valores dos harmônicos (H1 até H11) para cada f_0 , de forma que é possível visualizar as relações entre os harmônicos e os formantes. Nas notas em que foi detectada a ocorrência de sintonia de um formante a um harmônico (F_n-H_n), há uma seta abaixo, no caso de F1, ou acima, no caso de F2, da coordenada do formante no gráfico.

Quadro 1 – F1 (azul) e F2 (laranja) das vogais /a/, /e/ e /i/ faladas (triângulo) e cantadas (linha e ponto). Diagonais cinzas indicam os harmônicos H1 a H11, de baixo para cima. Setas indicam sintonia F_n-H_n .



Fonte: Autor.

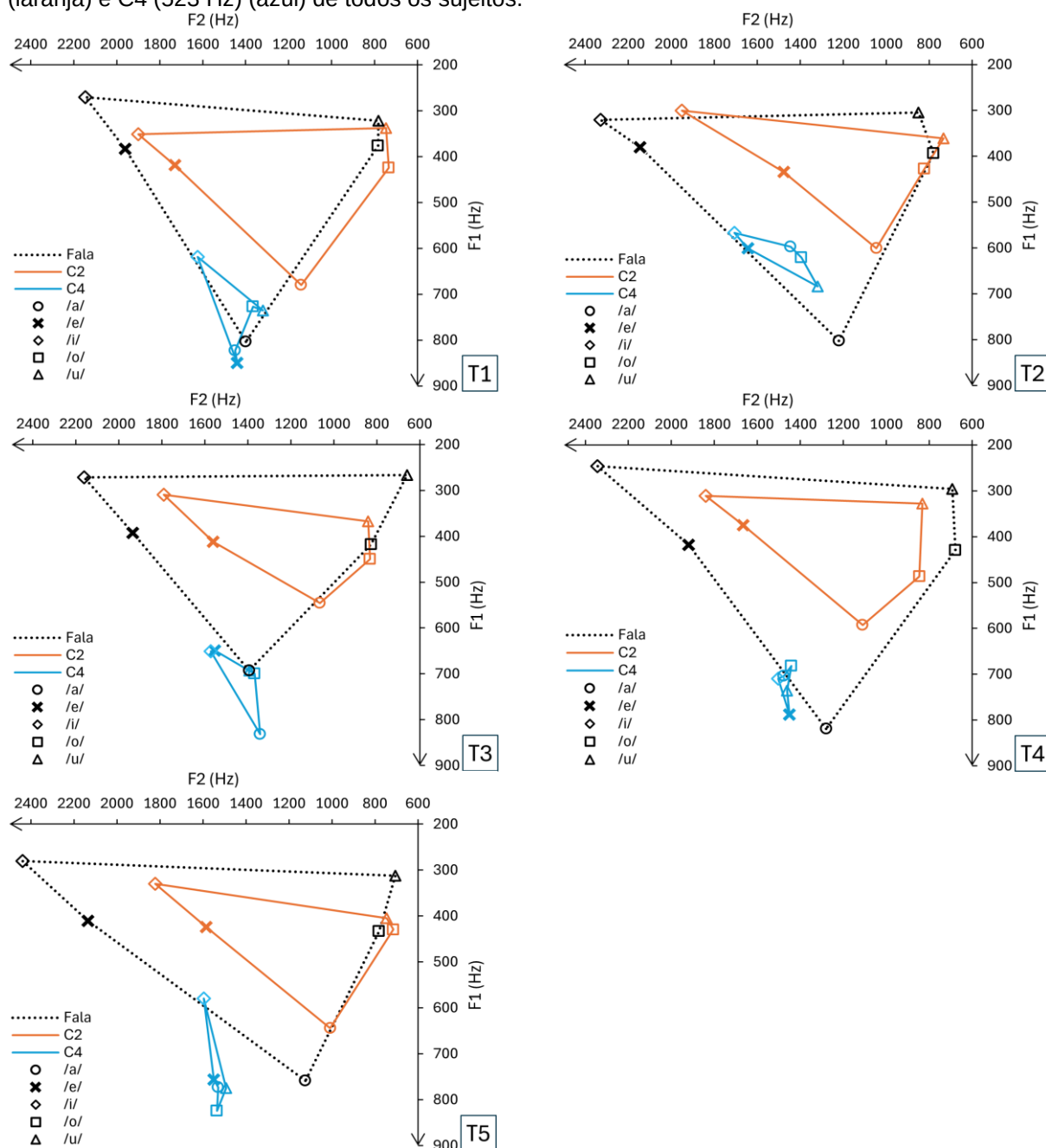
Quadro 2 – F1 (azul) e F2 (laranja) das vogais /o/ e /u/ faladas (triângulo) e cantadas (linha e ponto). Diagonais cinzas indicam os harmônicos H1 a H11, de baixo para cima. Setas indicam sintonia F_n - H_n .



Fonte: Autor.

No Quadro 3 estão sobrepostos os espaços vocálicos das vogais faladas e das notas mais grave, C2 (131 Hz), e mais aguda, C4 (523 Hz), extraídas das escalas. Todos os cantores apresentaram diminuição progressiva da área do espaço vocálico das vogais faladas para a nota C2, e desta para a nota C4.

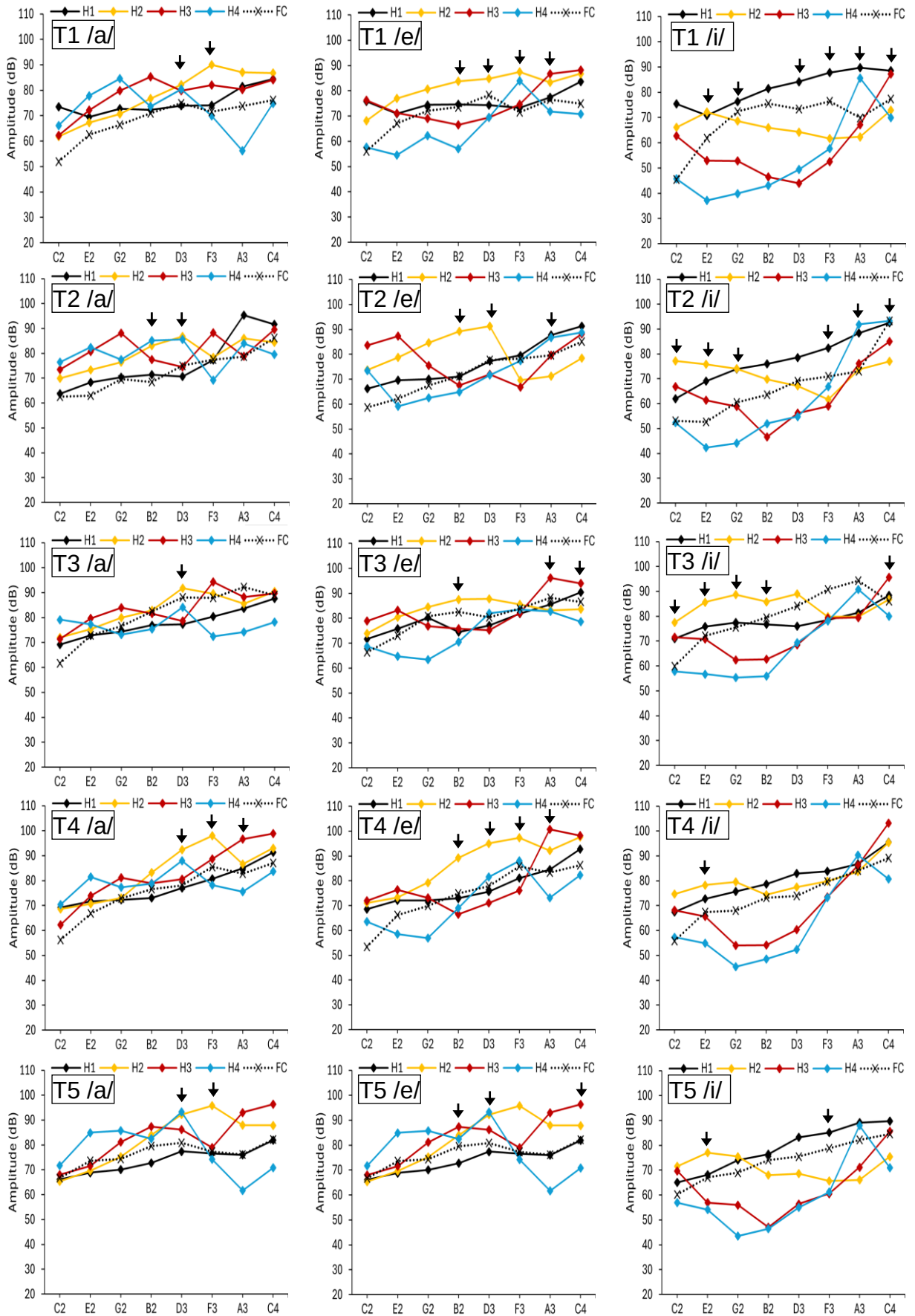
Quadro 3 – Espaços vocálicos das vogais faladas (linha pontilhada preta) e das notas C2 (131 Hz) (laranja) e C4 (523 Hz) (azul) de todos os sujeitos.



Fonte: Autor.

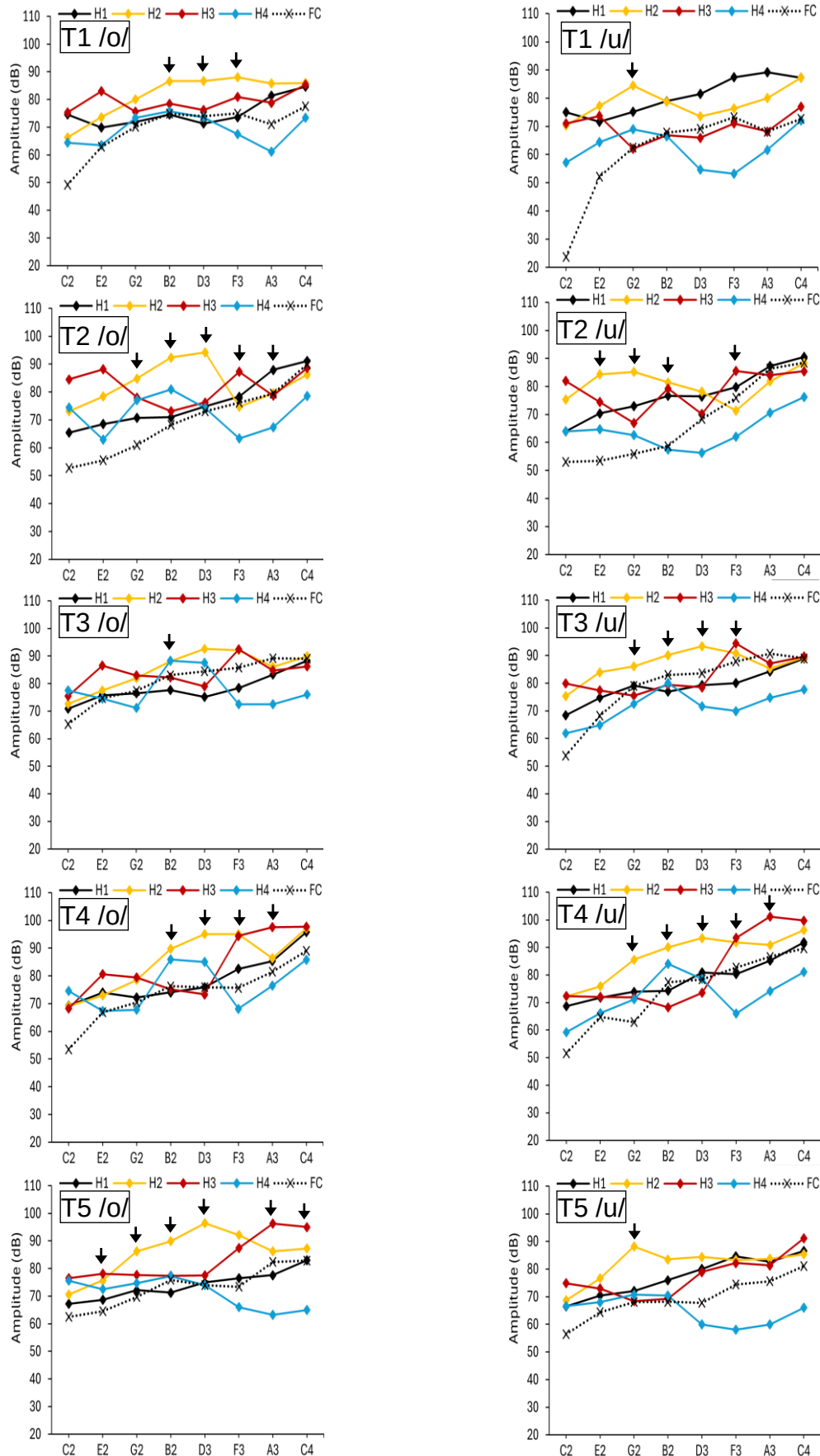
As variações de amplitude de H1, H2, H3 e H4, e do pico na região do formante do cantor (FC), estão relacionadas às notas da escala nos gráficos dos quadros 4 e 5. As notas em que há sintonia de formantes estão identificadas com um seta.

Quadro 4 – Amplitude de H1 (preto), H2 (amarelo), H3 (vermelho), H4 (azul) e do pico na região do FC (linha pontilhada), em função das notas musicais, das vogais /a/, /e/ e /i/. Setas indicam sintonia F_n-H_n .



Fonte: Autor.

Quadro 5 – Amplitude de H1 (preto), H2 (amarelo), H3 (vermelho), H4 (azul) e do pico na região do FC (linha pontilhada), em função das notas musicais, das vogais /o/ e /u/. Setas indicam sintonia F_n - H_n .



Fonte: Autor.

As Tabelas 3, 4 e 5 exibem o número de ocorrências das sintonias de formantes mais frequentemente verificadas em tenores (do primeiro formante com um dos dois primeiros harmônicos, e do segundo formante com o terceiro ou o quarto harmônicos). Nas tabelas estão relacionadas as estratégias de sintonia respectivamente com: vogais, sujeitos e regiões de f_0 , divididas em notas abaixo da zona passagem (C2, E2 e G2), notas da zona de passagem (B2, D3 e F3) e notas acima da zona de passagem (A3 e C4).

Tabela 3 – Número de ocorrências de sintonia formantes-harmônicos nas cinco vogais.

Tipo de sintonia	/a/	/e/	/i/	/o/	/u/	Total
F1-H2	5	11	11	10	11	48
F1-H1	0	1	5	1	0	7
F2-H4	5	2	2	8	1	18
F2-H3	2	5	2	5	5	19
Total	12	19	20	24	17	92

Fonte: Autor.

Tabela 4 – Número de ocorrências de sintonia formantes-harmônicos em cada um dos tenores (T).

Tipo de sintonia	T1	T2	T3	T4	T5	Total
F1-H2	10	11	9	11	7	48
F1-H1	2	4	0	0	1	7
F2-H4	4	4	1	5	4	18
F2-H3	2	3	4	6	4	19
Total	18	22	14	22	16	92

Fonte: Autor.

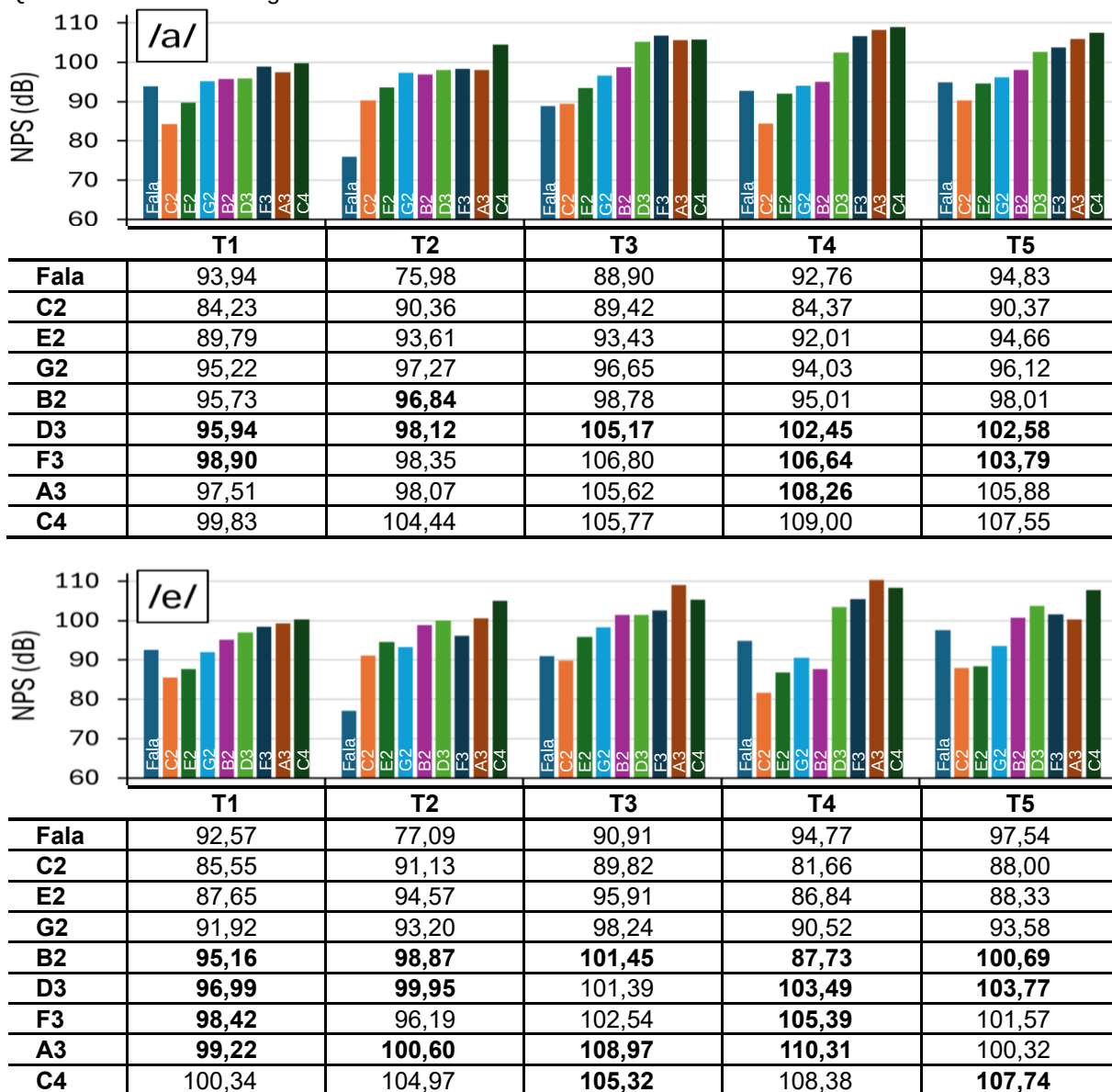
Tabela 5 – Número de ocorrências de sintonia formantes-harmônicos por região de f_0 , em função da zona de passagem.

Tipo de sintonia	Abaixo da passagem	Zona de passagem	Acima da passagem	Total
F1-H2	17	31	0	48
F1-H1	0	4	3	7
F2-H4	2	14	2	18
F2-H3	0	6	13	19
Total	19	55	18	92

Fonte: Autor.

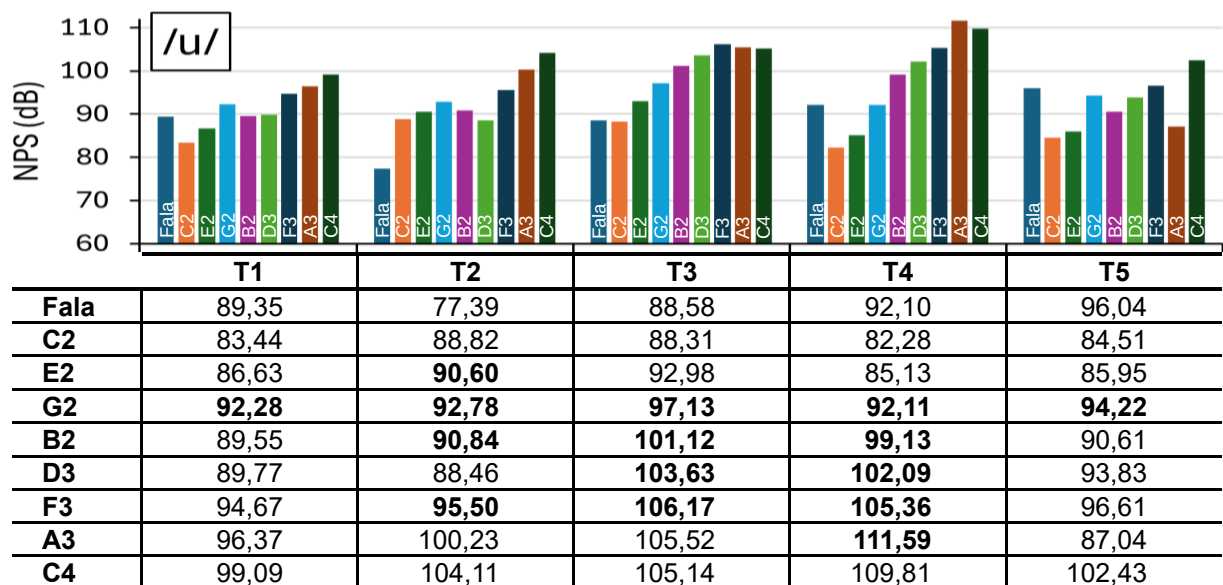
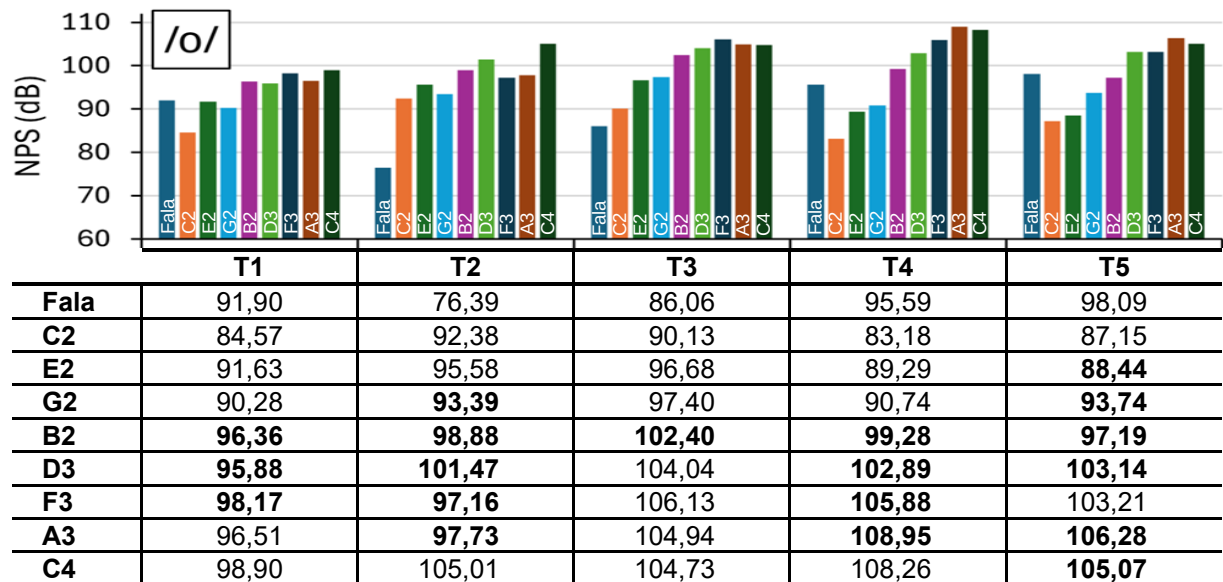
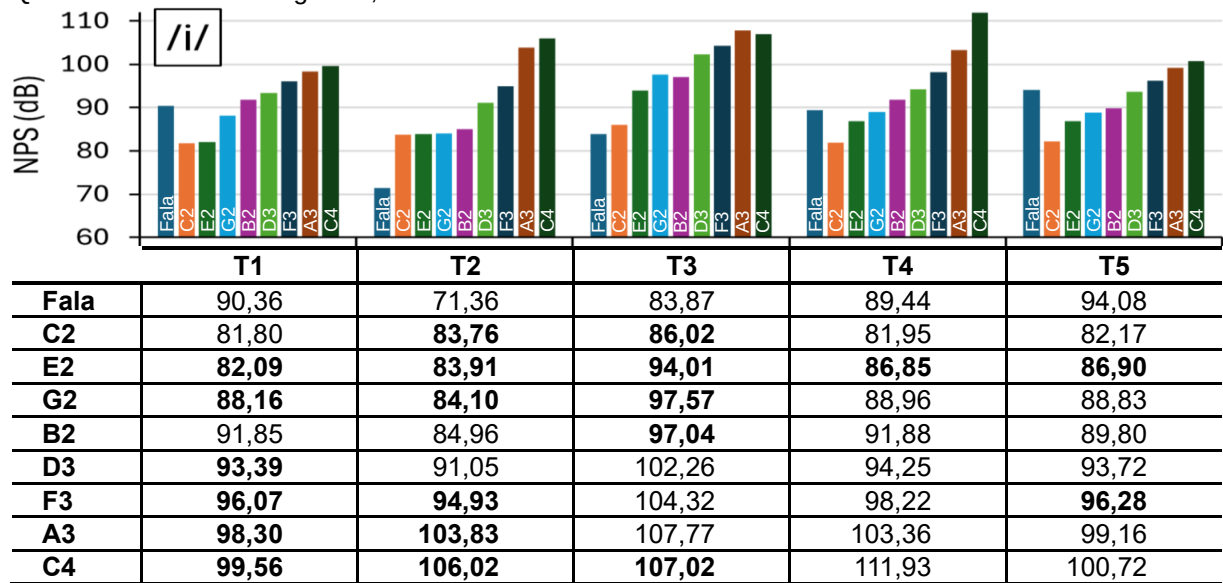
Os valores de nível de pressão sonora (NPS) das vogais faladas extraídas das sentenças-veículo, bem como os das vogais cantadas nas escalas, estão exibidos nas tabelas e nos gráficos do Quadro 6 (para as vogais /a/ e /e/) e do Quadro 7 (vogais /i/, /o/ e /u/). As notas nas quais foi observada a ocorrência de sintonia de formantes estão identificadas nas tabelas em **negrito**. Há variabilidade da intensidade das vogais faladas em relação às cantadas, mas todos os sujeitos mostraram a tendência geral de aumentar progressivamente o NPS no decorrer das escalas, da nota mais grave à mais aguda.

Quadro 6 – NPS das vogais /a/ e /e/ faladas e cantadas.



Fonte: Autor.

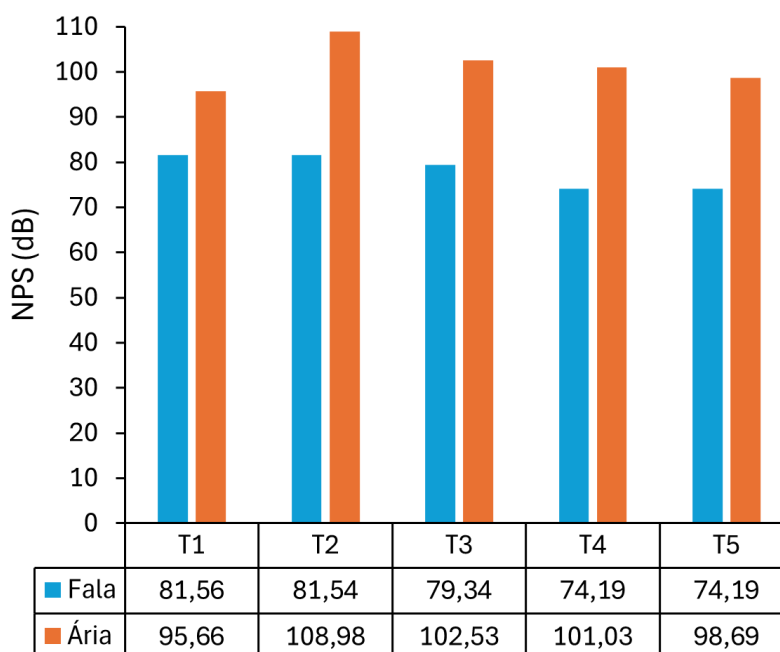
Quadro 7 – NPS das vogais /i/, /o/ e /u/ faladas e cantadas.



Fonte: Autor.

O gráfico e a tabela na Figura 14 apresentam os valores de NPS medidos para a fala espontânea e para a performance da ária de todos os cantores. Os cinco exibiram números mais elevados na execução na ária do que no trecho falado.

Figura 14 – NPS da fala espontânea (azul) e da execução da ária (laranja) de todos os sujeitos.



Fonte: Autor.

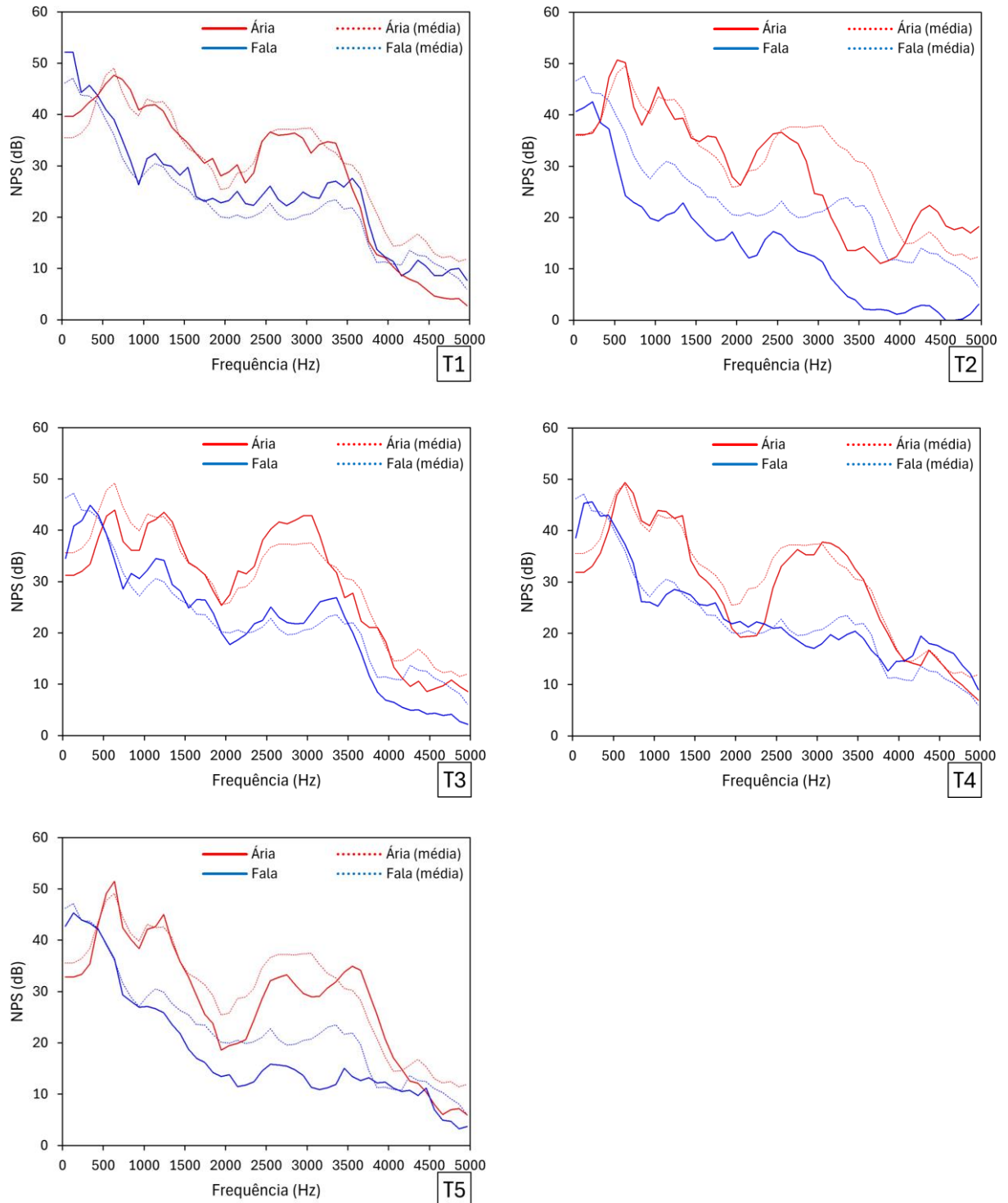
O Quadro 8 apresenta a sobreposição dos espectros médios de longo termo (EMLT) das falas espontâneas e das execuções da ária de cada sujeito, comparados às respectivas médias. Nos EMLT, os grandes picos de energia por volta de 2500 e 3000 Hz, presentes nas curvas das execuções das árias, correspondem à região do formante do cantor. Em algumas das curvas de fala espontânea, um discreto aumento de amplitude entre 3500 e 4500 Hz corresponde às frequências do formante do falante. Na Tabela 6 são exibidos os valores de declínio espectral dos correspondentes EMLT. Todos os sujeitos apresentaram declínios menos acentuados no canto.

Tabela 6 – Declínio espectral dos EMLT da fala espontânea e da ária.

Declínio espectral do EMLT (dB)						
	T1	T2	T3	T4	T5	Média
Ária	-10,27	-10,82	-1,64	-8,54	-10,31	-8,316
Fala	-21,26	-22,75	-14,39	-19,94	-23,14	-20,296
Diferença	10,99	11,93	12,75	11,4	12,83	11,98

Fonte: Autor.

Quadro 8 – Sobreposição dos EMLT da fala espontânea (azul) e da execução da ária (vermelho). As curvas pontilhadas representam as médias dos 5 sujeitos.



Fonte: Autor.

6. Discussão

No presente estudo sobre notas agudas, foram analisadas as vozes de cinco tenores eruditos. O projeto contemplava inicialmente a comparação de cantores eruditos com cantores populares, no caso pagode, em relação aos aspectos acústicos da emissão aguda. A ideia era abordar o canto de um gênero musical popular brasileiro. A dificuldade em encontrar cantores de pagode dispostos a participar da pesquisa foi enorme, o que limitou a amostra para cantores eruditos. Foram convidados 19 tenores, oito não responderam e quatro mostraram interesse, mas não participaram por motivos de saúde ou por não terem data disponível para agendar os procedimentos. Foram gravados para o estudo sete tenores, dos quais os dois primeiros serviram como estudo piloto para definição da coleta de dados.

A apreciação dos resultados deste trabalho deve levar em consideração determinadas limitações intrínsecas ao método. A análise de vogais em escalas, desprovidas de um contexto semântico, não é representativa da prática artística do cantor, que normalmente canta palavras associadas a uma melodia, que apresenta uma expressividade musical. As vogais faladas nas sentenças-veículo igualmente não são representativas da fala natural. Não se pode também descartar a possibilidade de que os cantores, sabendo que seriam gravados e de alguma forma avaliados, tenham ficado mais atentos à sua produção vocal. Sendo assim, pode-se hipotetizar que algumas vogais tenham sido produzidas em um ajuste mais confortável, e possivelmente diferente do que ocorre no palco nas notas mais agudas. Entende-se, porém, que o nível de detalhamento das análises aqui propostas demanda tal rigor metodológico, com uma alta padronização das tarefas e das condições do procedimento para todos os cantores.

A amostra da pesquisa foi composta por cinco tenores profissionais, atuantes em óperas, com diferentes tempos de experiência e diferentes tipologias vocais. Os sujeitos T1 e T3 tinham, à época da coleta, 42 anos de idade e em torno de 23 anos de profissão; o sujeito T2 era o mais jovem, com 30 anos de idade e 10 anos de profissão; e os sujeitos T4 e T5 tinham 54 e 53 anos, com 32 anos de experiência em média. Quanto às classificações vocais, três dos sujeitos eram tenores lírico ligeiro (T1, T4 e T5), um era tenor lírico (T2) e um era tenor lírico *spinto* (T3). Dessa forma, o estudo abarcou cantores em diferentes estágios do desenvolvimento vocal e da

carreira, e vozes relativamente mais leves e mais pesadas. As possíveis implicações dessas características serão relacionadas aos resultados durante a discussão.

As respostas dadas pelos sujeitos da pesquisa acerca das modificações necessárias à produção das notas agudas (Tabela 2) foram agrupadas em cinco temas. A temática mais citada, considerada pelos cinco tenores, foi a qualidade vocal, com colocações como a necessidade de manter a qualidade sonora, modificar as vogais, “evitar agudos minguados”. A segunda questão mais citada foi a dos ajustes laríngeos, sobre os quais foram mencionados o uso dos músculos tireoaritenóideos (TA) para “não desligar o motor”, manter a adução das pregas vocais e o equilíbrio muscular. Em terceiro lugar foram apontadas as adaptações corporais, relacionadas ao apoio, força no “core” e manter o equilíbrio entre a abertura das costelas e a contração abdominal. O estudo de Sousa et al. (2010), apesar de não ter abordado especificamente notas agudas, retornou resultados semelhantes, em respostas de professores de canto sobre os conceitos que trabalham em aula e as metáforas que utilizam. O apoio, a qualidade da emissão e o equilíbrio dos registros foram igualmente os tópicos mais levantados. A necessidade de adaptar o apoio também foi apontada por professores de canto (Fernandes, Andrada e Silva, 2020b) e sopranos (Fernandes, Andrada e Silva, 2020b, 2020a), quando perguntados sobre a produção das notas superagudas da soprano erudita, acima do C5 (1046 Hz).

A importância dos ajustes supraglóticos para as notas agudas também foi levantada por 3 sujeitos da pesquisa, como a abertura de boca, e língua baixa ou “mais neutra” (Tabela 2). Esses mesmos ajustes foram citados pelas sopranos e pelos professores de canto entrevistados no trabalho de Fernandes e Andrada e Silva (2020b). A produção vocal dos superagudos das sopranos apresenta desafios fisiológicos diferentes daqueles que os tenores enfrentam nas suas notas mais altas, por ocorrerem em diferentes registros e faixas de f0. Ambos precisam, entretanto, lidar com a proximidade entre F1 e suas frequências de fonação, utilizando, porém, diferentes estratégias, uma vez que objetivam qualidades vocais distintas.

A relevância do corpo na emissão dos agudos foi outro ponto em comum levantado tanto pelos sujeitos da pesquisa quanto pelas sopranos superagudas (Fernandes, Andrada e Silva, 2020b, 2020a), que citaram respectivamente o aumento do empenho físico e o equilíbrio entre força e relaxamento, e a liberdade de movimento

e a participação do corpo todo na produção. A atenção que cantores e cantoras de ópera relatam dar ao corpo, a fim de garantir a qualidade técnica das suas produções vocais, tem implicações sobre suas performances no palco. O envolvimento emocional com a performance, por sua vez, também tem influências sobre o corpo e a técnica dos cantores. O artista de ópera precisa, ao mesmo tempo, cantar, utilizando técnicas de altíssima demanda, e atuar, ser um personagem, com sua postura corporal, suas atitudes e suas emoções. Para Coutinho et al. (2019), é provável que alguns aspectos do canto limitem a gama e a intensidade de emoções que o cantor de ópera sente. Segundo os autores, cantores relatam ser necessário encontrar um meio termo entre o envolvimento emocional com o papel e o controle fisiológico demandado pela técnica adequada. As emoções têm íntima relação com ajustes glóticos, articulatórios e respiratórios. Importantes variações em parâmetros glóticos, relacionados às dimensões de hipo e hiperfunção na fonação, foram identificadas nos glotogramas de fluxo de cantores que expressaram diferentes emoções no canto (Sundberg et al., 2021).

Mello et al. (2015) apresentaram uma discussão semiológica a respeito da relação entre autoimagem e voz de cantores líricos. Apesar de a publicação não ter foco específico em uma ou outra questão técnica, dados os desafios para a produção de notas agudas elencados no presente trabalho e os aspectos proprioceptivos citados nos parágrafos acima, cabe incluir aqui algumas das reflexões levantadas no texto. Para os autores, ocorre nos cantores uma alteração de autopercepção, em virtude das ferramentas interpretativas empregadas para caracterizar a personalidade de um personagem. As modificações de imagem corporal têm a capacidade de produzir alterações na qualidade vocal, bem como o contrário também pode ocorrer. Além das dificuldades intrínsecas à fisiologia e aos aspectos acústicos das emissões em notas agudas, fica evidente aqui o desafio imposto também pelas camadas de emoções e expressividade que integram a práxis artística do cantor de ópera, uma vez que elas têm o poder de desestabilizar o corpo do artista. Se torna imprescindível, além de um alto desenvolvimento técnico vocal, o refinamento da consciência corporal do cantor.

As medidas dos formantes obtidas das escalas e das vogais extraídas das sentenças-veículo (Quadros 1 e 2) mostram diferentes variações entre voz falada e

voz cantada, considerando a região da fala, a depender da vogal. Na vogal /a/, as frequências de F1 e F2 da voz cantada tenderam a ficar abaixo ou muito próximas das da voz falada. Nas vogais anteriores /e/ e /i/, todos os sujeitos exibiram na voz cantada valores de F1 muito próximos ou logo acima daqueles da fala, enquanto os valores de F2 ficaram bem abaixo. Nas vogais posteriores /o/ e /u/, tanto F1 quanto F2 da voz cantada tenderam a ficar acima ou próximos das medidas da voz falada.

Os movimentos de F1 e F2 ao longo das escalas também variaram em função das vogais. A frequência de F1 da vogal /a/ mostrou diferentes tendências entre os sujeitos: se manteve razoavelmente estável para os sujeitos T1, T4 e T5; no sujeito T2 apresentou uma leve queda, e no sujeito T3 um movimento ascendente. Todos os sujeitos mostraram algum grau de elevação da frequência de F2 da vogal /a/ em direção às notas agudas. O F1 da vogal /e/ subiu para os sujeitos T1, T3, T4 e T5; enquanto que, para o sujeito T2, o formante iniciou a escala com uma subida, teve uma forte queda entre as notas D3 e F3, e voltou a subir. A queda de F1 no limite superior da zona de passagem em tenores foi descrito por Neumann et al. (2005) nas vogais /a/ e /o/, segundo os autores para que F1 fique mais grave que H2 e saia da sua proximidade, de forma a diminuir sua amplitude. O F2 da vogal /e/ apresentou um movimento geral descendente para todos os sujeitos, à exceção do T2, cujo F2 subiu ligeiramente, de forma que se manteve quase estável. Na vogal /i/, todos os sujeitos apresentaram subida da frequência de F1 e queda da de F2. Nas vogais posteriores /o/ e /u/, tanto F1 quanto F2 subiram em todos os sujeitos. A tendência de tenores para a elevação de F1 em direção a f0 agudas foi verificada por Titze et al. (1994). Echternach et al. (2016) identificaram em tenores clara tendência na subida de F1 na vogal /i/, e correlação entre as subidas de F1 e f0 nas vogais /e/, /o/ e /u/. Apenas na vogal /a/ não foi observada tal relação. Henrich et al. (2011) constataram variações de F1 e F2 em função de f0 nas vogais /a/, /ɔ/, /u/ e /ɜ/, não apenas em tenores, como também em sopranos, contraltos e barítonos. A subida de F1 está diretamente relacionada ao aumento da abertura da boca (Titze e Worley, 2009; Titze et al., 2011). A maior abertura de lábios e mandíbula na região aguda de tenores, acima da zona de passagem, foi constatada por meio da observação de imagens de ressonância magnética (Echternach et al., 2014).

Os espaços vocálicos (Quadro 3) ilustram a tendência de todos os sujeitos a modificarem as vogais da voz falada para a voz cantada, e de f0 graves para f0 agudas. Os espaços vocálicos de todos os cantores são amplos na voz falada; diminuem em direção ao centro do gráfico na nota mais grave da escala, e reduzem ainda mais na nota aguda, de forma que as vogais tendem a convergir para uma pequena área. Quanto maior for a área de um espaço vocálico, maior também serão as possibilidades de variação das frequências de F1 e F2 entre as vogais, ou seja, maior será o grau de liberdade articulatória que o cantor terá (Titze et al., 2017). Dessa forma, observa-se que, à medida que f0 aumenta, o cantor tem progressivamente menos possibilidades de variações das vogais e do timbre.

Do espaço vocálico da fala para o da nota grave da voz cantada, C2 (131 Hz), há especial deslocamento das vogais /a/, /e/ e /i/, caracterizado pela diminuição da frequência de F2 para as três vogais, e diminuição de F1 na vogal /a/ e aumento de F1 na vogal /i/. Mudanças nas vogais /o/ e /u/ variam entre os sujeitos. Da nota grave para a nota aguda, C4 (523 Hz), o espaço vocálico de todos os sujeitos sofre uma importante diminuição de área e um grande deslocamento para uma região centralizada em um F1 médio alto, por volta de 700 Hz, e F2 intermediário, por volta de 1600 Hz. Tais achados estão de acordo com a pesquisa de Hollien et al. (2000). Entretanto, não há padrão no formato dos espaços vocálicos agudos entre os sujeitos. Por exemplo, o sujeito T1 apresentou variação aproximada de 200 Hz para F1 e 300 Hz para F2, enquanto o sujeito T4 apresentou uma área extremamente reduzida, com aproximadamente 100 Hz de variação para F1, e menos de 100 Hz para F2. Millhouse e Kenny (2008) elaboraram espaços vocálicos semelhantes para seis cantores eruditos: três tenores, um barítono e dois baixos. Foram apresentados, para cada sujeito, cinco espaços vocálicos em notas entre E1 (85 Hz) e C#3 (277 Hz). Apesar de não serem apresentados dados para a região de notas agudas, todos os sujeitos mostraram a tendência de centralização e diminuição do espaço vocálico em função da subida da frequência fundamental.

A partir das estimativas das frequências dos formantes das vogais cantadas dos sujeitos, foram investigadas as ocorrências de sintonias dos formantes F1 e F2 com os harmônicos H1, H2, H3 e H4 (Tabelas 3, 4 e 5). As diferentes possibilidades de sintonia estão relacionadas nas tabelas às vogais, aos sujeitos e a três regiões de

f0 delimitadas pela zona de passagem – B2 (247 Hz) a F3 (349 Hz) –, ou seja, abaixo, dentro e acima dela. Nas 200 emissões de vogais analisadas no estudo, foram verificados 92 casos de sintonia de formantes: 48 do primeiro formante com o segundo harmônico (F1-H2), 19 do segundo formante com o terceiro harmônico (F2-H3), 18 do segundo formante com o quarto harmônico (F2-H4) e 7 do primeiro formante com o primeiro harmônico (F1-H1). A sintonia F2-H4 ocorreu 9 vezes em conjunto com F1-H2, e a sintonia F2-H3 ocorreu uma vez simultaneamente à F1-H2. Assim, os 92 casos de sintonia de formantes foram verificados em 82 emissões do total de 200.

Todos os sujeitos da pesquisa empregaram a sintonia de um formante a um harmônico em algum ponto das escalas, em todas as vogais. 55 das ocorrências foram na zona de passagem, das quais 31 foram a sintonia F1-H2 e 14 a sintonia F2-H4. Essas duas estratégias são consideradas por Neumann et al. (2005) e por Schutte et al. (2005) como o padrão para tenores na zona de passagem. Acima da zona de passagem, a sintonia mais identificada, com 13 ocorrências, foi F2-H3, igualmente considerada uma das principais estratégias das notas agudas do tenor (Neumann et al., 2005; Schutte et al., 2005). Evitar a sintonia de F1-H2 acima da segunda passagem e passar a utilizar F2-H3 para as notas agudas é também a recomendação pedagógica de Miller (2000) e Bozeman (2013). Entretanto, a prevalência desses ajustes não foi verificada por Henrich et al. (2011) e Echternach et al. (2016). Ambos os estudos identificaram diferentes possibilidades de ressonâncias para as notas agudas, não necessariamente envolvendo sintonias precisas de formantes a harmônicos. No presente trabalho, tais estratégias se mostraram frequentes e, portanto, relevantes, porém não houve indícios de serem ajustes obrigatórios ou preferenciais, correspondendo a 48% (F1-H2) e 26% (F2-H3) das emissões em cada zona de frequências.

Nenhum cantor aplicou consistentemente uma ou outra sintonia nas diferentes vogais. Houve consistência, porém, entre todos os cantores, em não utilizar a sintonia F1-H2 acima da zona de passagem, de forma que todos mantiveram F1 abaixo de H2 a partir da nota F3 ou acima dela. Todos os sujeitos exibiram a maioria das ocorrências de suas sintonias de formantes na zona de passagem. Tal achado pode indicar a maior relevância dessas estratégias para manejar as mudanças de preponderância entre os músculos tireoaritenóideos (TA) e os músculos cricotireóideos (CT) na

laringe, que ocorrem justamente na região de f_0 da zona de passagem (Roubeau et al., 2009; Kochis-Jennings et al., 2014). Dessa forma evita-se que ocorra a mudança para o registro de falsete, ou que se obtenha uma qualidade vocal de grito ou chamado – *yell* (Bozeman, 2013) ou *calling* (Miller, 2019) –, que seria mais apropriada para o *belting* (Schutte e Miller, 1993).

Em concordância com estudos anteriores (Titze et al., 1994; Schutte et al., 2005; Henrich et al., 2011; Echternach et al., 2016), a sintonia de F1-H1 ocorreu raramente, empregada apenas pelos sujeitos T1, T2 e T5. Essa sintonia é considerada como a sonoridade típica da voz aguda feminina no canto erudito (Henrich et al., 2011; Titze et al., 2011), e quando ocorre na voz do tenor, comumente está associada ao formante do cantor (Schutte et al., 2005).

A teoria de interação não-linear fonte-filtro prevê que possam ocorrer descontinuidades na fonação, caso um formante esteja logo abaixo de um harmônico, colocando-o em uma região de reatância negativa. É esperado que ocorra o maior risco de instabilidades quando f_0 estiver logo acima de F1 (Titze, 2008a). Houve nos sujeitos da pesquisa apenas sete emissões com sintonia de F1 a H1, e em todas elas a frequência do formante foi mantida logo acima da do harmônico. Entretanto, em 35 das outras 85 medidas, nas quais foram identificadas sintonias de formantes com os harmônicos mais agudos (Tabelas 3, 4 e 5), o formante estava ligeiramente abaixo da frequência do harmônico mais próximo. Outros pesquisadores também não observaram preferência por se situar os formantes acima dos harmônicos (Henrich et al., 2011; Henrich Bernardoni et al., 2014; Echternach et al., 2016), nem interferências perceptíveis na fonação com o formante logo abaixo de um harmônico (Sundberg et al., 2016). Ao invés de apenas evitarem que os formantes fiquem logo abaixo de um harmônico, Sundberg et al. (2016) hipotetizaram que os cantores recorreriam a outras manobras para evitar tais instabilidades.

A sintonia exata de um formante não parece ser estritamente necessária para que haja um harmônico dominante em destaque no espectro. Nos gráficos que exibem as amplitudes dos quatro primeiros harmônicos e do formante do cantor (Quadros 4 e 5), há notas em que há um harmônico de muito maior amplitude, em que não foi identificada a sintonia com um formante, e há casos em que foi observada sintonia, e não há um harmônico de especial destaque. Schutte et al. (2005) descreveram a

possibilidade de se aproximar F1 por baixo e F2 por cima de um harmônico, ainda mantendo uma distância deste. Como a aproximação de dois formantes causa um aumento de amplitude na função de transferência na região entre eles (Sundberg, 2018), esse harmônico ganharia intensidade sem a necessidade de sintonia. A nota A3 na vogal /o/ do sujeito T1 pode ser potencialmente um exemplo dessa estratégia: o H2 é o harmônico mais forte (Quadro 5), e abaixo e acima dele estão F1 e F2 (Quadro 2).

A análise dos harmônicos de maior intensidade na nota mais aguda, C5 (523 Hz), permite perceber que a estratégia mais utilizada foi a distribuição de energia entre vários harmônicos, o que pode ser identificado em 15 do total de 25 emissões dessa nota. Estes resultados estão de acordo com achados de Titze et al. (1994) e Sundberg et al. (2013). O único sujeito que mostrou mais ocorrências de um harmônico dominante (sempre o H3), foi o sujeito T5. Os sujeitos T2 e T3 mostraram consistência na utilização do formante do cantor tão forte quanto os parciais mais graves, fato coerente com o fato de os dois terem as vozes menos leves da amostra (Schutte et al., 2005).

Nas vogais cantadas nas escalas (Quadros 6 e 7), todos os cantores apresentaram uma clara tendência de aumento do nível de pressão sonora (NPS) do grave ao agudo, fato previsto por Titze (2000). O sujeito T1 mostrou a menor variabilidade de intensidade, com aumento do NPS em 19 dB do grave ao agudo, enquanto o do sujeito T5 variou 30 dB. Os outros sujeitos exibiram um aumento médio de 23 dB. Na opinião do pesquisador, o sujeito T1 apresentou a qualidade vocal mais leve entre todos os sujeitos, percepção condizente com o repertório do cantor, que conta com óperas de Mozart e Rossini, que exigem leveza e grande agilidade vocal. Essa achado está de acordo com Miller et al. (2001), que em um estudo que comparou um tenor lírico e um tenor dramático durante a emissão de uma nota aguda, com diminuição de intensidade do forte para o fraco, constataram uma menor variação entre os extremos de intensidade (NPS máximo e mínimo) no tenor lírico. O tenor *spinto* da presente pesquisa (T3), entretanto, não apresentou a maior variação, mantendo-se na média com os sujeitos T4 e T5.

Em todos os sujeitos, as vogais fechadas /i/ e /u/ tenderam a apresentar valores de NPS menores que as outras vogais, especialmente no início das escalas. Pabon e

Ternström (2020) argumentaram ser clara a tendência do NPS a cair nas vogais fechadas, uma vez que a área de radiação da boca é menor que nas vogais abertas. Os sujeitos T1, T2 e T4 exibiram valores semelhantes para todas as vogais nas notas agudas, enquanto o sujeito T5 manteve menores intensidades para as vogais /i/ e /u/ até o fim da escala. O sujeito T3 mostrou regularidade entre todas as vogais do grave ao agudo, fato que pode ser associado também a sua classificação vocal, tenor lírico *spinto*, que exige maior potência vocal inclusive nas notas médias e graves.

Não foi identificada relação entre a ocorrência de sintonia de formantes e o NPS das vogais cantadas, havendo ocasiões em que as notas com sintonia estavam mais fracas, mais fortes, ou com igual intensidade à das notas anteriores ou posteriores, nas quais não ocorreu sintonia. Pabon e Ternström (2020) conduziram uma pesquisa na qual elaboraram perfis de extensão vocal de cantores amadores e não cantores, a fim de relacionar as variações de NPS com mudanças em outros diversos parâmetros. Os pesquisadores concluíram haver pouco suporte para a expectativa de que uma ressonância aumentando a amplitude de um único harmônico teria efeito significativo no SPL total de emissão. Em um estudo de simulações computacionais, Titze (2021) demonstrou que a estratégia mais eficiente para o aumento de NPS é deixar a curva do declínio espectral mais reta. Para tanto, são necessários ajustes da fonte, relacionados a uma adução eficiente e a uma maior velocidade de fechamento, com aumento da taxa máxima do declínio de fluxo. Herbst e Story (2022), também na análise de simulações computacionais, concluíram haver eficiência no aumento de NPS com a sintonia de formantes a harmônicos em notas mais agudas. Relatam, porém, que um forte sinal acústico vindo da fonte dispensaria a necessidade de se utilizar tal estratégia.

Considerando o NPS dos excertos da execução da ária e da fala espontânea (Figura 14), os cinco tenores apresentaram maiores valores na voz cantada do que na voz falada. Dada a natureza do canto erudito, em que é necessário o cantor projetar a sua voz por um teatro, através do som de uma orquestra, sem o auxílio de amplificação eletrônica, é esperado que a voz cantada apresente maior intensidade. O sujeito T1 apresentou o menor NPS da ária e o maior da voz falada, ou seja, a menor diferença entre eles. Esse achado pode ser relacionado aos resultados das escalas do cantor, que apresentou a menor variabilidade de NPS da nota mais grave

à mais aguda entre os 5 sujeitos, bem como à sua qualidade vocal, caracterizada pela leveza e pela agilidade. O sujeito T5, tenor *spinto*, apresentou valores médios em ambas as tarefas. É sabido entretanto, que o NPS não é um bom indicador de percepção de *loudness* (Titze e Palaparthi, 2020; Titze, 2021), de forma que ainda seria possível a voz do sujeito T3 ser percebida pelos ouvintes como a mais forte e dramática entre os cinco participantes. Echternach et al. (2024) apontaram o vibrato com maiores amplitude e frequência, que não foram medidas no presente estudo, como características definitivas das vozes dramáticas.

Todos os cantores apresentaram declínio espectral dos espectros médios de longo termo (EMLT) (Tabela 6) menos acentuado na execução da ária, em média 11,98 dB maior que o da fala espontânea. Os valores do sujeito T3 foram razoavelmente maiores que os dos outros sujeitos, que apresentaram números semelhantes entre si. É notável que o declínio espectral do espectro radiado também não mostrou relação com o NPS. Quanto menos acentuado for o declínio espectral, ou seja, quanto mais próximo de zero for o seu valor, maior é o conteúdo de energia na região aguda do espectro. A concentração de energia nas frequências altas do espectro radiado não depende unicamente da configuração do filtro, sendo essencial um sinal acústico robusto, rico em harmônicos agudos, na fonte glótica (Sundberg, 2018; Herbst et al., 2023). O principal correlato fisiológico de um espectro da fonte glótica rico e com declínio menos acentuado, é a taxa máxima de declínio de fluxo (Pabon e Ternström, 2020), ou seja, a velocidade de fechamento das pregas vocais a cada ciclo glótico. O momento de fechamento da glote é o principal responsável pela excitação do trato vocal (Salomão e Sundberg, 2009). O declínio menos acentuado do espectro da fonte e, conseqüentemente, do espectro radiado, está associado à maior percepção de *loudness* (Titze e Palaparthi, 2020; Titze, 2021; Herbst e Story, 2022). Novamente, o resultado diferente do sujeito T3 pode ser associado a sua categoria vocal mais dramática, caracterizada pelo uso de maior pressão subglótica e de maior adução das pregas vocais (Echternach et al., 2024).

Os cinco tenores da pesquisa apresentaram os EMLT (Quadro 8) semelhantes, tanto entre as execuções da ária quanto entre as amostras de fala espontânea. Na ária, todos apresentaram o pico do formante do cantor (FC) por volta de 3000 Hz, à exceção do sujeito T2, cujo FC ficou um pouco mais baixo, logo acima de 2500 Hz. O

pico do sujeito T2 também se caracterizou por ser menos largo que o dos outros cantores. Apesar disso, nas vogais das escalas (Quadros 4 e 5), o sujeito T2 foi o segundo que apresentou maior amplitude no FC, especialmente nas notas agudas, perdendo apenas para o tenor lírico *spinto* T3. O tenor T2 já atua profissionalmente há 10 anos, tendo cantado papéis secundários em importantes casas de ópera de São Paulo, mas é o mais jovem de todos os sujeitos (30 anos na data da coleta dos dados), e não se pode descartar a possibilidade de a sua voz ainda não ter atingido seu máximo desenvolvimento. O sujeito T3 apresentou no EMLT da voz cantada um FC especialmente forte, de amplitude próxima à da região dos dois primeiros formantes, típico de vozes mais dramáticas (Schutte et al., 2005), mais uma vez sugerindo um alto conteúdo de energia aguda advindo do rápido fechamento da glote.

A presença do FC na voz cantada é esperada em tenores eruditos (Cleveland et al., 2001; Sundberg, 2001; Schutte et al., 2005; Millhouse, 2012). Nenhum sujeito apresentou o FC na voz falada. T1 e T3 apresentaram um pico de energia por volta de 3500 Hz, porém mais discreto que o FC, o que corresponde ao formante do falante (Cleveland et al., 2001).

As notas agudas se mostram desafiadoras fisiológica e acusticamente para os tenores de ópera. As estratégias de sintonia de formante mais comumente citadas na literatura científica (Neumann et al., 2005; Schutte et al., 2005; Henrich et al., 2011; Echternach et al., 2016), e recomendada por pedagogo vocais influentes (Miller, 2008; Bozeman, 2013), como F1-H2 abaixo e F2-H3 acima da segunda passagem, foram encontradas no presente trabalho. Tais estratégias, porém, não se mostraram tão frequentes ou obrigatórias, uma vez que outras soluções também foram identificadas, como um forte formante do cantor ou a distribuição de energia acústica entre mais de um dos parciais mais graves. Echternach et al. (2016) encontraram semelhantes resultados.

Numa crítica aos trabalhos de Schutte et al. (2005), no qual foram analisadas gravações comerciais de tenores célebres do século XX, e Neumann et al. (2005), em que foram analisados sujeitos de pesquisa, Echternach et al. (2016) chamaram atenção ao fato de não terem sido medidas as frequências dos formantes, elas foram apenas inferidas a partir dos parciais de maior amplitude no espectro. Os autores comentam: “Desse modo, se a sintonia de formantes é definida com base na

ocorrência de parciais proeminentes, o termo sintonia parece um tanto problemático nesses casos”¹⁵ (Echternach et al., 2016, p. 304). Como pode ser visto nos resultados dos sujeitos do presente trabalho, é possível que um harmônico tenha destaque no espectro sem que seja necessária uma aproximação de um formante, tal que se possa dizer que tenha sido feita uma sintonia exata (ou muito próxima) de frequências.

Howell (2016), na sua teoria Cor Absoluta do Som Espectral (*Absolute Spectral Tone Color*), compara as diferentes frequências do espectro a cores, que o cantor pode escolher para “colorir” a sua voz. Segundo Plomp (1966), sons simples (ondas senoides) também têm um timbre característico típico, que depende somente da sua frequência. Esses sons simples, quando situados dentro de um som complexo, mantêm sua cor absoluta. Assim, a qualidade sonora inerente dessas frequências não se perde, ela pode ser expressada ou suprimida, dependendo da sua amplitude e das amplitudes das frequências próximas a ela (Howell, 2016). Dessa forma, cada frequência de maior amplitude no espectro contribui um pouco para a qualidade sonora final que será percebida pelo ouvinte.

Conclui-se esta dissertação com a seguinte reflexão: tanto a grande amplitude de H3, quanto o formante do cantor ou a distribuição equilibrada de energia entre os parciais mais graves, mais do que apenas estratégias para garantir projeção vocal ou estabilidade fonatória, podem todos ser “cores” características do canto do tenor operístico. Elas estão entre as possibilidades de timbres e sonoridades que o artista tem à sua disposição para expressar musicalmente o conteúdo emocional dos personagens que ele vive nos palcos de ópera.

¹⁵ Tradução do pesquisador, do original em inglês: “Hence, if formant tuning is defined on the basis of the occurrence of prominent partials, the term tuning appears quite problematic in these cases” (Echternach et al., 2016, p. 304).

7. Conclusão

O presente estudo sobre os aspectos acústicos das notas agudas em tenores eruditos sugere serem necessárias modificações do trato vocal para a produção dessas notas. As vogais se tornam, do grave ao agudo, progressivamente menos diferenciáveis, dada a convergência dos valores das frequências dos dois primeiros formantes. As estratégias acústicas mais comumente associadas às notas agudas de tenores eruditos – a sintonia do primeiro formante ao segundo harmônico na zona de passagem, e do segundo formante ao terceiro harmônico acima dela – foram frequentemente observadas nos sujeitos da pesquisa. O uso de sintonias foi principalmente observado na zona de passagem ou logo acima dela, enquanto que, na nota mais aguda emitida pelos cantores, foi mais frequente a distribuição de energia entre os harmônicos mais graves. Entretanto, não foi identificada a ocorrência de um padrão de ajustes específicos dos formantes em relação aos harmônicos, que fosse comum a toda a extensão aguda, a todas as vogais e a todos os sujeitos. Os tenores eruditos estudados nesta pesquisa exibiram o emprego de estratégias variadas para manejar a zona de passagem e as emissões nas frequências fundamentais mais altas, o que indica a necessidade de diferentes ajustes das ressonâncias do trato vocal ao longo da sua tessitura aguda.

.

8. Referências Bibliográficas

- Andrada e Silva MA de, Duprat A de C. Avaliação do paciente cantor. Em: Marchesan IQ, Silva HJ da, Tomé MC, organizadores. Tratado das especialidades em fonoaudiologias. Editora Guanabara Koogan Ltda; 2014.
- Bennett R. Elementos Básicos da Música. Rio de Janeiro: Jorge Zahar; 1998.
- Björkner E. Musical Theater and Opera Singing—Why So Different? A Study of Subglottal Pressure, Voice Source, and Formant Frequency Characteristics. *Journal of Voice*. setembro de 2008;22(5):533–40.
- Boersma P, Kovacic G. Spectral characteristics of three styles of Croatian folk singing. *The Journal of the Acoustical Society of America*. 1º de março de 2006;119(3):1805–16.
- Borch DZ, Sundberg J. Some Phonatory and Resonatory Characteristics of the Rock, Pop, Soul, and Swedish Dance Band Styles of Singing. *Journal of Voice*. setembro de 2011;25(5):532–7.
- Bozeman K. Practical vocal acoustics: pedagogic applications for teachers and singers. Hillsdale, NY: Pendragon Press; 2013. (Vox musicae: the voice, vocal pedagogy, and song).
- Cleveland TF, Sundberg J, Stone RE. Long-Term-Average Spectrum Characteristics of Country Singers During Speaking and Singing. *Journal of Voice*. março de 2001;15(1):54–60.
- Coutinho E, Scherer KR, Dikken N. Singing and Emotion. Em: Welch GF, Howard DM, Nix J, organizadores. *The Oxford Handbook of Singing* [Internet]. Oxford University Press; 2019 [citado 16 de maio de 2024]. p. 296–314. Disponível em: <https://academic.oup.com/edited-volume/35194/chapter/299587424>
- Denes PB, Pinson EN. *The Speech Chain: The Physics and Biology of Spoken Language*. Pickle Partners Publishing; 2016.
- Dr Dan's Voice Essentials. How Vocal Tract Acoustics Affect Singing Registers | Ken Bozeman with #DrDan [Internet]. 2018 [citado 24 de outubro de 2023]. Disponível em: <https://www.youtube.com/watch?v=9qBTbDXRyik>
- Echternach M, Birkholz P, Sundberg J, Traser L, Korvink JG, Richter B. Resonatory Properties in Professional Tenors Singing Above the Passaggio. *Acta Acustica united with Acustica*. 1º de março de 2016;102(2):298–306.
- Echternach M, Burk F, Kirsch J, Traser L, Birkholz P, Burdumy M, et al. Articulatory and acoustic differences between lyric and dramatic singing in Western classical music. *The Journal of the Acoustical Society of America*. 1º de abril de 2024;155(4):2659–69.

- Echternach M, Traser L, Richter B. Vocal Tract Configurations in Tenors' Passaggio in Different Vowel Conditions—A Real-Time Magnetic Resonance Imaging Study. *Journal of Voice*. março de 2014;28(2):262.e1-262.e8.
- Fernandes DE, Andrada e Silva MA de. Superagudos: análise perceptivo-auditiva da voz e autorrelato em sopranos profissionais. *CoDAS*. 2020a;32(4):e20190068.
- Fernandes DE, Andrada e Silva MA de. What do Sopranos and Singing Teachers Think About Upper High Range Notes? *Journal of Voice*. 16 de outubro de 2020b;Publicado online.
- Fontanella BJB, Ricas J, Turato ER. Amostragem por saturação em pesquisas qualitativas em saúde: contribuições teóricas. *Cad Saúde Pública*. janeiro de 2008;24(1):17–27.
- Garcia M. *Mémoire sur la voix humaine: présenté à l'Académie des sciences en 1840*. 2º ed. Paris: Imprimerie d'E. Duverger; 1847.
- Henrich Bernardoni N, Smith J, Wolfe J. Vocal tract resonances in singing: Variation with laryngeal mechanism for male operatic singers in chest and falsetto registers. *The Journal of the Acoustical Society of America*. janeiro de 2014;135(1):491–501.
- Henrich N. Mirroring the voice from Garcia to the present day: some insights into singing voice registers. *Logopedics Phoniatrics Vocology*. janeiro de 2006;31(1):3–14.
- Henrich N, Smith J, Wolfe J. Vocal tract resonances in singing: Strategies used by sopranos, altos, tenors, and baritones. *The Journal of the Acoustical Society of America*. fevereiro de 2011;129(2):1024–35.
- Herbst CT, Story BH. Computer simulation of vocal tract resonance tuning strategies with respect to fundamental frequency and voice source spectral slope in singing. *The Journal of the Acoustical Society of America*. 1º de dezembro de 2022;152(6):3548–61.
- Herbst CT, Story BH, Meyer D. Acoustical Theory of Vowel Modification Strategies in Belting. *Journal of Voice*. abril de 2023;S0892199723000048.
- Hollien H, Mendes-Schwartz AP, Nielsen K. Perceptual confusions of high-pitched sung vowels. *Journal of Voice*. junho de 2000;14(2):287–98.
- Howell I. *Parsing the spectral envelope: towards a general theory of vocal tone color* [Tese de doutorado]. [Boston, MA]: The New England Conservatory of Music; 2016.
- Kochis-Jennings KA, Finnegan EM, Hoffman HT, Jaiswal S, Hull D. Cricothyroid Muscle and Thyroarytenoid Muscle Dominance in Vocal Register Control: Preliminary Results. *Journal of Voice*. setembro de 2014;28(5):652.e21-652.e29.

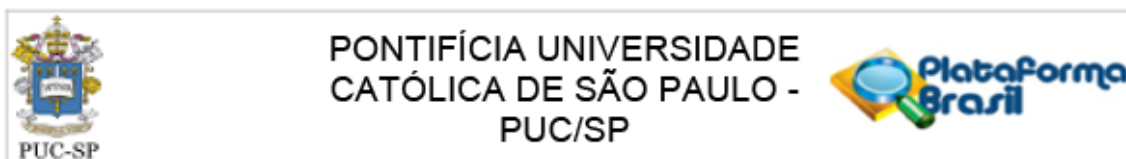
- Lamesch S, Expert R, Castellengo M, Henrich N, Chuberre B. Investigating voix mixte: A scientific challenge towards a renewed vocal pedagogy. Maimets-Volt K, Parncutt M, Marin M, Ross J, organizadores. Proceedings of the third Conference on Interdisciplinary Musicology (CIM07). 2007;
- Leino T. Long-Term Average Spectrum in Screening of Voice Quality in Speech: Untrained Male University Students. *Journal of Voice*. novembro de 2009;23(6):671–6.
- Master S, De Biase N, Chiari BM, Laukkanen AM. Acoustic and Perceptual Analyses of Brazilian Male Actors' and Nonactors' Voices: Long-term Average Spectrum and the "Actor's Formant". *Journal of Voice*. março de 2008;22(2):146–54.
- Med B. Teoria da Música. 4ª. Brasília: MusiMed; 1996.
- Mello ÊL, Ballesterio LRB, Andrada e Silva MA de. Postura corporal, voz e autoimagem em cantores líricos. *Per musi*. junho de 2015;(31):74–85.
- Michaelis On-line. Whoop. Em: Michaelis On-line [Internet]. 2023 [citado 24 de outubro de 2023]. Disponível em: <https://michaelis.uol.com.br/moderno-ingles/busca/ingles-portugues-moderno/whoop/>
- Miller DG. Registers in Singing: Empirical and Systematic Studies in the Theory of the Singing Voice [Tese de doutorado]. [Groningen]: Rijksuniversiteit Groningen; 2000.
- Miller DG. Resonance in singing: voice building through acoustic feedback. Princeton, NJ: Inside View Press; 2008.
- Miller DG, Schutte HK, Doing J. Soft Phonation in the Male Singing Voice: A Preliminary Study. *Journal of Voice*. 2001;15(4):483–91.
- Miller R. Training tenor voices. New York : Toronto : New York: Schirmer Books ; Maxwell Macmillan Canada ; Maxwell Macmillan International; 1993.
- Miller R. A Estrutura do Canto: sistema e arte na técnica vocal. 1º ed. São Paulo: É Realizações; 2019.
- Millhouse TJ. Observations of the Higher Formant Structure in the Male Operatic Vowel. Em: Proceedings of the 14th Australasian International Conference on Speech Science and Technology. Sydney: ASSTA; 2012.
- Millhouse TJ, Kenny DT. Vowel placement during operatic singing: "come si parla" or "aggiustamento"? Em: Proceedings of Interspeech 2008 [Internet]. Brisbane: ISCA; 2008 [citado 16 de maio de 2024]. p. 245–8. Disponível em: https://www.isca-archive.org/interspeech_2008/millhouse08_interspeech.html
- Neumann K, Schunda P, Hoth S, Euler HA. The Interplay between Glottis and Vocal Tract during the Male Passaggio. *Folia Phoniatri Logop*. 2005;57(5–6):308–27.

- Operabase. Verified Opera Statistics | Performances | Composers | City | Production | Operabase. 2023 [citado 6 de fevereiro de 2023]. Disponível em: <https://www.operabase.com/statistics/en>
- Pabon P, Ternström S. Feature Maps of the Acoustic Spectrum of the Voice. *Journal of Voice*. janeiro de 2020;34(1):161.e1-161.e26.
- Plomp R. Experiment on tone perception. Rijksuniversiteit Utrecht; 1966.
- Roubeau B, Henrich N, Castellengo M. Laryngeal Vibratory Mechanisms: The Notion of Vocal Register Revisited. *Journal of Voice*. julho de 2009;23(4):425–38.
- Salomão GL. Registros vocais no canto: aspectos perceptivos, acústicos, aerodinâmicos e fisiológicos da voz modal e da voz de falsete [Tese de doutorado]. [São Paulo]: Pontifícia Universidade Católica de São Paulo; 2008.
- Salomão GL, Sundberg J. What do male singers mean by modal and falsetto register? An investigation of the glottal voice source. *Logopedics Phoniatrics Vocology*. janeiro de 2009;34(2):73–83.
- Schutte HK, Miller DG. Belting and pop, nonclassical approaches to the female middle voice: Some preliminary considerations. *Journal of Voice*. junho de 1993;7(2):142–50.
- Schutte HK, Miller DG, Duijnste M. Resonance Strategies Revealed in Recorded Tenor High Notes. *Folia Phoniatr Logop*. 2005;57(5–6):292–307.
- Sousa JMD, Andrada e Silva MAD, Ferreira LP. O uso de metáforas como recurso didático no ensino do canto: diferentes abordagens. *Rev soc bras fonoaudiol*. 2010;15(3):317–28.
- Story BH, Laukkanen AM, Titze IR. Acoustic impedance of an artificially lengthened and constricted vocal tract. *Journal of Voice*. dezembro de 2000;14(4):455–69.
- Sundberg J. Level and Center Frequency of the Singer's Formant. *Journal of Voice*. junho de 2001;15(2):176–86.
- Sundberg J. *Ciência da voz: fatos sobre a voz na fala e no canto*. 1. ed. 1. reimpr. rev. São Paulo: Editora da Universidade de São Paulo; 2018.
- Sundberg J, Lã FMB, Gill BP. Formant Tuning Strategies in Professional Male Opera Singers. *Journal of Voice*. maio de 2013;27(3):278–88.
- Sundberg J, Lã FMB, Gill BP. Voice Source Variation Between Vowels in Male Opera Singers. *Journal of Voice*. setembro de 2016;30(5):509–17.
- Sundberg J, Salomão GL, Scherer KR. Analyzing Emotion Expression in Singing via Flow Glottograms, Long-Term-Average Spectra, and Expert Listener Evaluation. *Journal of Voice*. janeiro de 2021;35(1):52–60.

- Švec JG, Granqvist S. Tutorial and Guidelines on Measurement of Sound Pressure Level in Voice and Speech. *J Speech Lang Hear Res.* 15 de março de 2018;61(3):441–61.
- Titze IR. Principles of voice production. 2nd printing. Iowa City, IA: National Center for Voice and Speech; 2000.
- Titze IR. Nonlinear source–filter coupling in phonation: Theory. *The Journal of the Acoustical Society of America.* maio de 2008a;123(5):2733–49.
- Titze IR. The Human Instrument. *Sci Am.* janeiro de 2008b;298(1):94–101.
- Titze IR. The Effect of Single Harmonic Tuning on Vocal Loudness. *Journal of Voice.* novembro de 2021;35(6):832–7.
- Titze IR, Baken RJ, Bozeman KW, Granqvist S, Henrich N, Herbst CT, et al. Toward a consensus on symbolic notation of harmonics, resonances, and formants in vocalization. *The Journal of the Acoustical Society of America.* maio de 2015;137(5):3005–7.
- Titze IR, Mapes S, Story B. Acoustics of the tenor high voice. *The Journal of the Acoustical Society of America.* fevereiro de 1994;95(2):1133–42.
- Titze IR, Maxfield LM, Walker MC. A Formant Range Profile for Singers. *Journal of Voice.* maio de 2017;31(3):382.e9–382.e13.
- Titze IR, Palaparthi A. Vocal Loudness Variation With Spectral Slope. *J Speech Lang Hear Res.* 22 de janeiro de 2020;63(1):74–82.
- Titze IR, Worley AS. Modeling source-filter interaction in belting and high-pitched operatic male singing. *The Journal of the Acoustical Society of America.* setembro de 2009;126(3):1530–40.
- Titze IR, Worley AS, Story BH. Source-Vocal Tract Interaction in Female Operatic Singing and Theater Belting. *Journal of Singing.* junho de 2011;67(5):561–72.
- Wolfe J, Garnier M, Smith J. Vocal tract resonances in speech, singing, and playing musical instruments. *HFSP Journal.* fevereiro de 2009;3(1):6–23.

Anexos

Anexo 1 – Parecer Consubstanciado do CEP



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Notas agudas: aspectos acústicos em cantores de diferentes gêneros musicais

Pesquisador: GUSTAVO DO AMARAL AMBROS

Área Temática:

Versão: 1

CAAE: 65785922.3.0000.5482

Instituição Proponente: Pontifícia Universidade Católica de São Paulo-PUC/SP

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 5.871.438

Apresentação do Projeto:

Trata-se de um projeto de mestrado pertencente ao curso de Pós-graduação em Comunicação Humana e Saúde da Faculdade de Ciências Humanas e da Saúde, da PUC-SP.

notas agudas são um desafio técnico para cantores, exigindo ajustes da fonte glótica e do filtro. No aspecto fisiológico, o cantor deve

equilibrar a atividade das musculaturas da laringe. No acústico, considerando a teoria de interação não-linear fonte-filtro, é necessário ajustar as

frequências das ressonâncias do trato vocal, por meio de modificações no posicionamento dos articuladores do trato vocal, a fim de colocar os

harmônicos da fonte em regiões que apresentem reatância inercial, auxiliando na estabilidade da oscilação das pregas vocais e na eficiência na

geração da energia acústica. A hipótese é que diferentes gêneros musicais, por apresentarem variadas qualidades vocais, podem apresentar nas

notas agudas ajustes distintos no trato vocal, resultando em diferentes características no espectro acústico radiado. Objetivo: comparar a emissão

de notas agudas, no aspecto acústico, em cantores de diferentes gêneros musicais. Método: a amostra será composta de dez cantores, cinco do

gênero musical operístico e cinco do gênero musical pagode. Como procedimentos, os cantores terão suas vozes gravadas durante a execução de

Endereço: Rua Ministro Godoi, 969 - sala 63C, térreo do Prédio Reitor Bandeira de Mello
 Bairro: Perdizes CEP: 05.015-001
 UF: SP Município: SÃO PAULO
 Telefone: (11)3670-8466 Fax: (11)3670-8466 E-mail: cometica@pucsp.br



PONTIFÍCIA UNIVERSIDADE
CATÓLICA DE SÃO PAULO -
PUC/SP



Continuação do Parecer: 5.871.438

quatro tarefas: emissão das vogais /a/, /i/ e /u/ em tom habitual de fala; execução um trecho de uma peça musical pré-determinada e característica do gênero musical; vocalização do mesmo trecho musical nas vogais supracitadas; vocalização de uma escala ascendente de oito tons, da nota C4 (261 Hz) à nota C5 (523 Hz), nas mesmas vogais. O áudio será captado e analisado no software Praat. Na análise dos dados serão extraídos o nível de pressão sonora e o espectro médio de longo termo da execução da peça musical, no qual será observada a distribuição de energia nas diferenças faixas de frequência; nas escalas, será identificado no espectrograma de banda estreita em que notas ocorrem variações nos harmônicos de maior intensidade na região de frequências das duas primeiras ressonâncias do trato vocal. Desses pontos serão extraídos os valores das frequências de ressonância por meio da análise LPC. A mesma extração será realizada na nota mais aguda das vocalizações da peça musical nas vogais. Comparando essas medidas, será observada a variação das frequências de ressonâncias em função da frequência fundamental. Os dados extraídos serão relacionados por meio de comparações: intra-sujeito, considerando variações entre vogais; inter-sujeito e intra-gênero, observando as tendências de cada gênero musical; e inter-gênero, estabelecendo similaridades e diferenças entre os dois.

Objetivo da Pesquisa:

Objetivo Primário: Comparar os aspectos acústicos da emissão de notas agudas em cantores de diferentes gêneros musicais.

Objetivo Secundário: 1. Identificar no espectro acústico a distribuição de energia nas diferentes faixas de frequência em diferentes vogais nos gêneros musicais operístico e pagode. 2. Identificar as frequências de ressonância do trato vocal na emissão aguda em diferentes vogais nos gêneros musicais operístico e pagode. 3. Relacionar variações nas frequências de ressonância do trato vocal em função da frequência fundamental em diferentes vogais nos gêneros musicais operístico e pagode.

Avaliação dos Riscos e Benefícios:

Por serem utilizados apenas métodos não invasivos de captação de áudio, não há riscos físicos. Há baixo risco psicológico, que pode decorrer apenas de constrangimentos durante o procedimento

Endereço: Rua Ministro Godoi, 969 - sala 63C, térreo do Prédio Reitor Bandeira de Mello
Bairro: Perdizes CEP: 05.015-001
UF: SP Município: SAO PAULO
Telefone: (11)3670-8466 Fax: (11)3670-8466 E-mail: cometica@pucsp.br



PONTIFÍCIA UNIVERSIDADE
CATÓLICA DE SÃO PAULO -
PUC/SP



Continuação do Parecer: 5.871.438

de coleta. Benefícios: Contribuição científica para os campos da voz, da música e do canto. Os achados desta pesquisa, além de contribuírem para expandir os conhecimentos científicos acerca da voz cantada, podem ser de grande utilidade prática na pedagogia vocal, na clínica fonoaudiológica e no processo de estudo de cantores. O conhecimento sobre as características acústicas dos diferentes gêneros de canto pode possibilitar maior precisão e eficiência na intervenção em voz cantada e no processo do desenvolvimento vocal, artístico e estético dos cantores.

Comentários e Considerações sobre a Pesquisa:

O projeto está bem estruturado e atende bem às exigências éticas próprias das pesquisas com seres humanos.

Considerações sobre os Termos de apresentação obrigatória:

Todos os documentos foram apresentados. O TCLE tem uma linguagem clara e precisa. Tudo está de acordo com a resolução CNS 466/2012 e 510/2016.

Recomendações:

Nenhuma.

Conclusões ou Pendências e Lista de Inadequações:

Aprovado, sem pendências.

Considerações Finais a critério do CEP:

Aprovamos integralmente o parecer do relator.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMACOES_BASICAS_DO_PROJETO_2031418.pdf	05/12/2022 11:44:29		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.pdf	05/12/2022 11:43:54	GUSTAVO DO AMARAL AMBROS	Aceito
Folha de Rosto	folha_de_rostoassinada.pdf	01/12/2022 18:24:20	GUSTAVO DO AMARAL AMBROS	Aceito
Outros	parecer_de_merito.jpg	25/11/2022 15:26:22	GUSTAVO DO AMARAL AMBROS	Aceito
Projeto Detalhado / Brochura Investigador	projeto_detalhado_notas_agudas.docx	25/11/2022 15:17:00	GUSTAVO DO AMARAL AMBROS	Aceito

Endereço: Rua Ministro Godoi, 969 - sala 63C, térreo do Prédio Reitor Bandeira de Mello
Bairro: Perdizes CEP: 05.015-001
UF: SP Município: SAO PAULO
Telefone: (11)3670-8466 Fax: (11)3670-8466 E-mail: cometica@pucsp.br



Continuação do Parecer: 5.871.438

Outros	oficio_de_apresentacao_anexo_iii.pdf	25/11/2022 15:14:55	GUSTAVO DO AMARAL AMBROS	Aceito
--------	--------------------------------------	------------------------	-----------------------------	--------

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO PAULO, 01 de Fevereiro de 2023

Assinado por:
Antonio Carlos Alves dos Santos
 (Coordenador(a))

Endereço: Rua Ministro Godoi, 969 - sala 63C, térreo do Prédio Reitor Bandeira de Mello
 Bairro: Perdizes CEP: 05.015-001
 UF: SP Município: SAO PAULO
 Telefone: (11)3670-8466 Fax: (11)3670-8466 E-mail: cometica@pucsp.br

Anexo 2 – Termo de Consentimento Livre e Esclarecido (TCLE)



PONTIFÍCIA UNIVERSIDADE CATÓLICA DE SÃO PAULO

PROGRAMA DE PÓS-GRADUAÇÃO EM
COMUNICAÇÃO HUMANA E SAÚDE

Termo de Consentimento Livre e Esclarecido

Título da pesquisa: Notas agudas: aspectos acústicos em cantores de diferentes gêneros musicais

Pesquisador: Gustavo do Amaral Ambros

Orientador: Profa. Dra. Marta Assumpção de Andrada e Silva

Prezado participante,

Você está sendo convidado a participar, de forma totalmente voluntária, da pesquisa acima especificada. Sua contribuição é essencial para a execução da pesquisa, porém, você não deve participar contra a sua vontade. Antes de decidir participar, é importante que você entenda o porquê da realização do estudo, os procedimentos envolvidos e os possíveis benefícios, riscos e desconfortos.

Este trabalho estudará as notas agudas de cantores de diversos gêneros musicais. Espera-se que os resultados auxiliem na compreensão da voz cantada, otimizando a atuação de fonoaudiólogos e professores de canto no desenvolvimento técnico de cantores.

Sua participação se dará por meio da execução de um trecho musical que você conhece e da vocalização de escalas. Durante as tarefas, sua voz será gravada. O procedimento ocorrerá em um estúdio de gravação musical, em horário previamente agendado de acordo com a disponibilidade do pesquisador e dos sujeitos voluntários.

Sua privacidade será totalmente assegurada. Seus dados pessoais, assim como os áudios gravados, serão tratados com sigilo, sendo acessíveis apenas aos pesquisadores responsáveis. Você não será identificado na pesquisa e a divulgação do trabalho ocorrerá exclusivamente no meio acadêmico.

O benefício de sua participação nessa pesquisa é a sua contribuição científica para os campos da voz, da música e do canto. Sua participação é voluntária e espontânea, não

havendo remuneração financeira. Eventuais gastos com transporte para participação serão ressarcidos.

A participação no estudo não apresenta riscos físicos, e baixo risco em relação aos aspectos psicológicos, decorrentes apenas de possíveis constrangimentos que possam ocorrer durante o procedimento.

A qualquer momento, antes, durante ou após os procedimentos da pesquisa, você poderá solicitar maiores esclarecimentos sobre o estudo, recusar-se a participar ou desistir de participar. Em nenhum desses casos você será prejudicado(a), penalizado(a) ou responsabilizado(a) de qualquer forma.

Em caso de qualquer dúvida a respeito da pesquisa, você poderá entrar em contato com o pesquisador responsável, o mestrando Gustavo do Amaral Ambros, pelo telefone (51) 99298-3793 ou pelo e-mail gustavo.ambros@gmail.com.

Este estudo foi analisado pelo Comitê de Ética em Pesquisa da Pontifícia Universidade Católica de São Paulo (CEP-PUC/SP). Em caso de dúvidas sobre seus direitos como participante do estudo, ou se estiver insatisfeito com a realização deste, você pode entrar em contato com o CEP-PUC/SP – Sede Campus Monte Alegre, situado no Edifício Reitor Bandeira de Mello, sala 63-C, na Rua Ministro Godói, 969 – Perdizes – São Paulo – SP – CEP: 05015-001 – Tel./FAX: (11) 3670-8466 – e-mail: cometica@pucsp.br.

Eu, _____,
telefone () _____, ciente e de acordo com o que foi explicado, aceito participar voluntariamente neste estudo. Esse TCLE será assinado em duas vias, uma ficará com o pesquisador e outra em posse do sujeito voluntário.

São Paulo, _____ de 2023.

Sujeito voluntário

Gustavo do Amaral Ambros (pesquisador)

Profa. Dra. Marta Assumpção de Andrada e Silva (orientador)

Anexo 3 – Partitura da ária “Un’aura amorosa”, do personagem Ferrando no primeiro ato da ópera *Così fan tutte*

Redução da orquestra para piano

Compositor: Wolfgang Amadeus Mozart

Texto em italiano e tradução para alemão

Libretista: Lorenzo Da Ponte

Nº 17. Aria .

Andante cantabile.

FERR.

Un' au-ra a - mo - ro - sa del nostro te - so - ro un
Der O - dem der Liebe er - frischt die See - le, ein

dol - ce ri - sto - ro al cor por - ge - rà, un au - ra a - mo - ro - sa del
Bal - sam so wön - nig, so schmei - chelnd und weich, der O - dem der Lie - be er -

no - stro te - so - ro un dol - ce ri - sto - ro al cor por - ge - rà, un
frischt die See - le, ein Bal - sam so wön - nig, so schmei - chelnd und weich, ein

dol - ce ri - sto - ro al cor por - ge - rà. Al
Bal - sam so wön - nig, so schmei - chelnd und weich. Wer

cresc. *f* *p* *tr*

cor che nu - dri - to da spe - me d'a - mo - re, da
Lie - be ge - nie - sset und treu sie er - fin - det, und

spe - me d'a - more, di un e - sca mi - glio - re - bi - so - gno non ha, di un
treu sie er - findet, be - gehret nichts wei - ter, - ist se - lig und reich, be -

e - sca mi - glio - re bi - so - gno non ha, bi - sogno non ha, bi -
gehret nichts wei - ter, ist se - lig und reich, ist se - lig und reich, ist

sogno non ha. Un' au - ra a - mo - ro - sa del nostro te - so - ro un
se - lig und reich. Der O - dem der Liebe er - frischet die See - le, ein

dol - ce ri - sto - ro al cor por - ge - rà, un' au - ra a - mo - ro - sa del
Bal - sam, so won - nig, so schmeichelnd und weich; der O - dem der Lie - be er -

no - stro te - so - ro un dol - ce ri - sto - ro al cor por - ge - rà, un
frischt die See - le, ein Bal - sam so wonnig, so schmeichelnd und weich, ein

dol - ce ri - sto - ro al - cor por - ge - rà, al -
Bal - sam so won - nig, so schmeichelnd und weich, so -

cor por - ge - rà, al - cor por - ge - rà, un dol - ce ri - sto -
schmeichelnd und weich, so schmeichelnd und weich, ein Bal - sam so won -

ro al cor por - ge - rà. (Ferrando e Guglielmo partono.)
 nig, so schmei - chelnd und weich. (Ferrando und Guglielmo gehen ab.)

p *cresc.*

Fonte: Mozart WA. Così fan tutte - Vollständiger Klavierauszug mit Text. Levi H, editor. Leipzig: Breitkopf und Härtel; 1898.

Anexo 4 – Questionário de identificação da amostra

IDENTIFICAÇÃO PESSOAL

Nome: _____ Data de nascimento: _____
 Classificação vocal (incluir subclassificação, se souber): _____
 Há quanto tempo canta profissionalmente? (anos) _____

HÁBITOS DE SAÚDE

Você fuma? () Não () Sim
 Se sim, há quanto tempo você fuma? _____ Quantos cigarros você fuma por dia? _____
 Se é ex-tabagista, há quanto tempo parou de fumar? _____

Você ingere bebidas alcoólicas? () Não () Sim
 Se sim, quais bebidas alcoólicas você costuma consumir? _____
 Se sim, qual a quantidade/copos? _____ Se sim, qual a frequência semanal? _____

Você utiliza pastilhas e/ou sprays antes ou durante sua atividade vocal profissional? () Não () Sim
 Se sim, qual produto? _____

Faz gargarejo? () Não () Sim Se sim, quantas vezes por semana? _____ Se sim, com o quê? _____

Tem hábito de pigarrear? () Nunca () Raramente () Frequentemente
 Tem hábito de tossir? () Nunca () Raramente () Frequentemente

Utiliza gengibre? () Não () Sim
 Se sim, em qual situação? _____ Se sim, qual a frequência semanal? _____

Utiliza mel? () Não () Sim
 Se sim, em qual situação? _____ Se sim, qual a frequência semanal? _____

Come maçã? () Não () Sim
 Se sim, em qual situação? _____ Se sim, qual a frequência semanal? _____

Ingere água durante a sua apresentação? () Não () Sim
 Qual a sua média diária de ingestão de água diária fora das refeições? (copos) _____

Faz aquecimento vocal? () Não () Sim
 Se sim, faça uma breve descrição dos exercícios que costuma fazer para aquecimento. _____

Você considera a sua alimentação equilibrada (proteínas, grãos, frutas e verduras)? () Não () Sim
 Qual o intervalo médio entre as suas refeições? (horas) _____
 Você tem algum tipo de queixa relacionada à digestão/estômago?
 () Não () Sim Se sim, quais? _____

HÁBITOS DE BEM-ESTAR

Você acorda descansado? Considera seu sono reparador? () Nunca () Raramente () Frequentemente
 Você pratica atividades físicas (no mínimo 2x por semana)? () Não () Sim Quais? _____
 Antes do início da apresentação vocal, você tem algum hábito específico?
 () Não () Sim Se sim, qual? _____

Você costuma ter problemas de voz? () Não () Sim
 Se sim, você: (marque todas as alternativas que se aplicam)
 () se automedica; especifique os medicamentos _____
 () procura um otorrinolaringologista _____
 () procura outra especialidade médica; especifique _____
 () procura um fonoaudiólogo _____
 () procura o professor de canto ou regente _____
 () faz repouso vocal _____
 () outro _____

PERCEPÇÃO DAS NOTAS AGUDAS

Você sente que, para produzir notas agudas, precisa fazer algum tipo de modificação na maneira de cantar? Comente livremente (considere sensações, respiração, movimentos do corpo, do rosto, da boca, da garganta, entre outros aspectos perceptivos e proprioceptivos). _____