

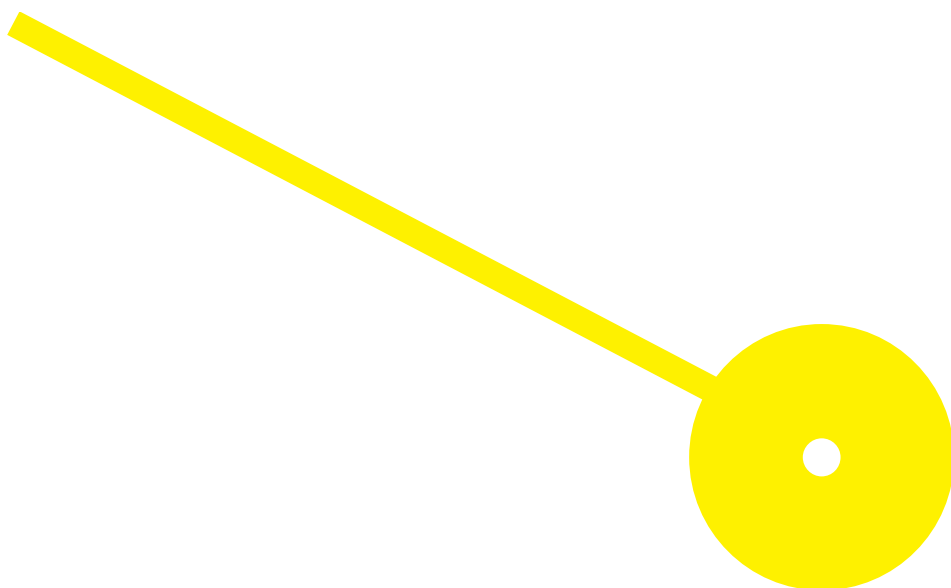
M

MESTRADO
TERAPIA DA FALA

Avaliação complementar da voz através de medidas acústicas de longo termo em vozes disfónicas

Joana Costa Frutuoso Faria

12/2020





Avaliação complementar da voz através de medidas acústicas de longo termo em vozes disfónicas

Autor

Joana Costa Frutuoso Faria

Orientador

Pedro André Guerreiro Martins Araújo (Prof. Doutor), Centro de Investigação e Reabilitação (CIR),
Escola Superior de Saúde (ESS), Instituto Politécnico do Porto (IPP)

Dissertação apresentada para cumprimento dos requisitos necessários à obtenção do grau de **Mestre em Terapia da Fala** pela Escola Superior de Saúde do Instituto Politécnico do Porto.

Agradecimentos

Ao Prof. Doutor André Araújo pela disponibilidade na orientação, pela partilha de conhecimento e pelo incentivo constante na concretização de cada etapa deste trabalho. Foi um privilégio ter sido orientada por um profissional que admiro.

À Prof. Doutora Sandra Alves pelo apoio no tratamento estatístico dos dados desta Dissertação.

À Unidade Local de Saúde de Matosinhos – Hospital Pedro Hispano, em especial ao Serviço de Otorrinolaringologia, pela colaboração e pela possibilidade de recolha de dados.

Aos membros do painel de peritos pela disponibilidade em colaborar neste projeto.

A todos os que aceitaram participar e emprestar a sua voz a este trabalho.

Aos professores e colegas que encontrei na ESS-IPP: todos contribuíram de alguma forma para o meu crescimento.

Às minhas colegas de profissão e de estudo. Em especial: à Liliana Ferreira, por toda a amizade e companheirismo demonstrado ao longo deste percurso e por tantas conversas motivadoras; à Rita Portela que partilhou comigo tantas fases deste projeto e “aturou” muitas das minhas dúvidas existenciais; e às Maduras, pela partilha de estados de alma.

A todos os meus amigos que sempre me apoiaram no decorrer desta etapa, sobretudo com a sua boa-disposição.

Aos meus pais, irmãos, sogros e cunhadas por todo o apoio, incentivo e compreensão. São, com toda a certeza, o meu ponto de apoio.

À minha cunhada Miriam pela revisão cuidada do texto.

À minha irmã Inês, um agradecimento especial pelas várias ajudas técnicas.

Às minhas filhas, Maria e Carolina, que todos os dias me inspiram a ser mais e melhor.

E por último, ao meu marido, que desde o primeiro momento me apoiou incondicionalmente. Obrigada, Pedro, por fazeres caminho comigo e acreditares nos meus sonhos.

Resumo

A avaliação acústica da voz inclui medidas acústicas de curto ou longo termo, sendo as primeiras as mais utilizadas na prática clínica. Contudo, apresentam limitações, sobretudo na análise da fala encadeada. O objetivo desta investigação foi estudar a contribuição de tarefas de fala encadeada, medidas acústicas de longo termo e medidas combinadas para complementar a avaliação da qualidade vocal, tendo como referência medidas de curto termo e áudio-percetivas. Realizou-se um estudo analítico, observacional, quantitativo transversal. Doze indivíduos, com e sem disfonia, foram sujeitos a avaliação por ORL, gravação da voz (vogal sustentada e leitura em diferentes intensidades) e avaliação áudio-percetiva (por três peritos, usando a escala CAPE-V adaptada), ficando distribuídos por três graus: sem alterações, grau ligeiro e grau moderado de disfonia. Foram calculadas medidas de curto e longo termo e da análise estatística (correlação de *Pearson*) concluiu-se que tarefas de leitura em diferentes intensidades são importantes na avaliação acústica, verificando-se correlação inversa entre a dinâmica de intensidades, $(H1-H2)_{LTAS}$, $CPPS_{fala}$ e $CPPS_{vogal}$ e o grau de disfonia. Verificou-se, também, a relevância dos índices DSI e AVQI na caracterização do grau de disfonia. Foram feitas considerações sobre as práticas de avaliação vocal pelos terapeutas da fala para uma melhor caracterização vocal.

Palavras-chave: disfonia, avaliação acústica, parâmetros acústicos de longo termo, DSI, AVQI

Abstract

The acoustic voice assessment includes short and long-term acoustic measures, the first ones being the most used in clinical practice. However, they have limitations, especially in speech analysis. The aim of this investigation was to study the contribution of long-term and combined acoustic measures to complement the assessment of vocal quality in speech tasks, using short-term acoustic and audio-perceptual measures as reference. An analytical, observational, quantitative cross-sectional study was carried out. Twelve individuals, with and without dysphonia, were submitted to laryngeal assessment by ENT, voice recording (sustained vowel and reading at different intensities) and audio-perceptual evaluation (by three experts, using the adapted CAPE-V scale), being distributed among three degrees: normal voice, slight and moderate degree of dysphonia. Short and long term measures were calculated and using statistical analysis (Pearson's correlation) results revealed that reading tasks at different intensities are important in acoustic evaluation, with an inverse correlation between parameters as intensity dynamics, $(H1-H2)_{LTAS}$, $CPPS_{fala}$ and $CPPS_{vocal}$ and the degree of dysphonia. The relevance of DSI and AVQI indexes was also verified in the characterization of the degree of dysphonia. Considerations were made about assessment practices of speech behaviours by speech-language therapists for better vocal characterization.

Keywords: dysphonia, acoustic voice evaluation, long-term acoustic parameters, DSI, AVQI

Índice

1. Introdução.....	1
2. Enquadramento Teórico	2
2.1. Definição de voz e de disfonia	2
2.2. Avaliação áudio-percetiva da voz	3
2.3. Avaliação acústica da voz	5
2.3.1. Medidas acústicas de curto termo	6
2.3.2. Medidas acústicas de longo termo.....	10
2.3.3. Índices vs. medidas isoladas.....	17
2.3.4. Tarefas vocais.....	19
2.4. Objetivos.....	21
3. Métodos.....	22
3.1. Tipo e desenho do estudo	22
3.2. População e amostra	23
3.3. Instrumentos	23
3.4. Procedimentos.....	26
3.4.1. Análise áudio-percetiva.....	27
3.4.2. Análise acústica.....	27
3.4.3. Análise de dados	29
4. Resultados.....	29
4.1. Caracterização da amostra	30
4.2. Relações entre variáveis.....	40
5. Discussão.....	43
6. Conclusão.....	49
Referências Bibliográficas	52
Apêndices	60
Apêndice 1 – Ficha de Caracterização do Participante	61
Apêndice 2 – Protocolo de Gravação da Voz – Tarefas.....	62
Apêndice 3 – Registo da avaliação áudio-percetiva (página Web)	63
Apêndice 4 – Autorização da Comissão de Ética da ULSM para realização do estudo.....	64
Apêndice 5 – Termo de Consentimento Informado	65
Anexo – Tarefas de leitura.....	66

Índice de abreviaturas, acrónimos e sinais

AVQI	<i>Acoustic Voice Quality Index</i>
CAPE-V	<i>Consensus Auditory Perceptual Evaluation of Voice</i>
CPP	<i>Cepstral Peak Prominence</i>
CPP _{fala}	<i>Cepstral Peak Prominence</i> em tarefa de fala encadeada
CPP _{vogal}	<i>Cepstral Peak Prominence</i> em tarefa de vogal sustentada
CPPS	<i>Cepstral Peak Prominence Smoothed</i>
CPPS _{fala}	<i>Cepstral Peak Prominence Smoothed</i> em tarefa de fala encadeada
CPPS _{vogal}	<i>Cepstral Peak Prominence Smoothed</i> em tarefa de vogal sustentada
CSID	<i>Cepstral Spectral Index of Dysphonia</i>
DSI	<i>Dysphonia Severity Index</i>
FFT	<i>Fast Fourier Transform</i>
f_0	Frequência fundamental
HNR	<i>Harmonic-to-Noise Ratio</i>
Int _{subj}	Intensidade vocal subjetiva
LTAS	<i>Long Term Average Spectrum</i> (espectro médio de longo termo)
ORL	Otorrinolaringologia
TMF	Tempo Máximo de Fonação

Índice de tabelas

Tabela 1 – Características de escalas de avaliação perceptiva da voz	5
Tabela 2 – Valores de frequência fundamental na população adulta com e sem patologia vocal em diferentes tarefas vocais	7
Tabela 3– Valores de referência para medidas do LTAS	14
Tabela 4 – Valores de referência das medidas CPP e CPPS (Praat) em indivíduos sem disfonia	16
Tabela 5– Caracterização demográfica da amostra	30
Tabela 6– Distribuição da amostra considerando diagnóstico ORL	30
Tabela 7– Coeficientes de correlação <i>intra</i> avaliador	31
Tabela 8 – Coeficiente de correlação intraclassa (ICC)	31
Tabela 9 – Resultados da avaliação áudio-perceptiva (média peritos)	31
Tabela 10 – Distribuição da amostra considerando o grau de disfonia	32
Tabela 11– Resultados das medidas acústicas na tarefa de Leitura em Intensidade subjetiva média	33
Tabela 12– Resultados das medidas acústicas na tarefa de Leitura em Intensidade subjetiva fraca	34
Tabela 13– Resultados das medidas acústicas na tarefa de Leitura em Intensidade subjetiva forte	35
Tabela 14– Resultados da dinâmica de intensidades	36
Tabela 15– Resultados das medidas acústicas na tarefa de vogal [a] sustentada	37
Tabela 16– Resultados dos índices DSI e AVQI	39
Tabela 17 – Coeficiente de correlação de Pearson – medidas acústicas de longo termo e avaliação áudio-perceptiva	41
Tabela 18– Coeficiente de correlação de Pearson – medidas acústicas de longo e curto termo	42
Tabela 19– Coeficiente de correlação de Pearson – Grau de disfonia e índices AVQI e DSI	43

Índice de figuras

Figura 1 - Espectros.....	11
Figura 2- LTAS	12
Figura 3- Representação do Cepstrum para voz sem perturbação	14
Figura 4 - Procedimentos do estudo	26
Figura 5- Gráfico da Avaliação áudio-percetiva.....	32
Figura 6 - Gráfico LTAS do indivíduo VN2f	36
Figura 7 - Gráfico LTAS do indivíduo OF9f	37
Figura 8 - Resultado do DSI no Praat - indivíduo VN3m	38
Figura 9 - Índice AVQI indivíduo VN3m	38
Figura 10 - Índice AVQI indivíduo FU12m.....	39
Figura 11- Resultados dos índices DSI e AVQI	40

1. Introdução

A avaliação da voz exige uma abordagem multidimensional de forma a descrever e a explicar um fenómeno também ele multidimensional. Contempla, por parte do otorrinolaringologista, a avaliação instrumental por laringoscopia, por parte do terapeuta da fala a avaliação perceptiva, a análise acústica e a avaliação de medidas aerodinâmicas, e, por parte do próprio indivíduo, a avaliação auto-percetiva (Dejonckere et al., 2001).

A avaliação acústica da voz inclui várias medidas acústicas, que podem ser de curto ou longo termo, sendo as de curto termo as mais utilizadas na prática clínica. Por um lado, estas medidas são relativamente fáceis de obter através de vogais sustentadas, mantendo estáveis a altura tonal e a intensidade de modo a garantir a periodicidade do sinal acústico. Por outro lado, existem limitações, sobretudo nos casos de vozes com disфония (vozes com alteração da qualidade vocal), nomeadamente a fiabilidade dos parâmetros de perturbação que é baseada na periodicidade do sinal. Nas vozes com disфония, o sinal vocal apresenta um grau de aperiodicidade muito variável que muitos programas não conseguem calcular eficazmente (Freitas, 2012). Assim, nas vozes classificadas com grau de alteração global moderado a extremo, poderá ser difícil, ou até mesmo impossível, a extração dos valores das medidas de curto termo (Núñez-Batalla, Díaz-Fresno, Álvarez-Fernández, Muñoz Cordero, & Llorente Pendás, 2017).

Torna-se necessário complementar a avaliação acústica com tarefas de fala mais representativas da voz falada. A análise acústica da fala encadeada é possível através de medidas acústicas de longo termo, de que são exemplo as medidas obtidas através do Espectro Médio de Longa Duração (*Long Term Average Spectrum* - LTAS) e do *Cepstrum*. Contudo, estas medidas representam um comportamento extremamente variável (p.e. a fala) e contemplam os contributos de todo o mecanismo de produção vocal, incluindo fonte laríngea e todos os processos de ressonância. Por este motivo, podem ser de difícil extração e interpretação, razão pela qual, apesar de comuns em investigação em voz, não são usadas de forma corrente na prática clínica. Serão estas medidas capazes de se relacionar com o grau de alteração da qualidade vocal encontrada em tarefas de fala encadeada?

A presente dissertação pretende dar contributos para a prática dos terapeutas da fala na avaliação da qualidade vocal, sobretudo na avaliação acústica da fala encadeada e encontra-se organizada em cinco partes. A primeira parte, capítulo segundo, corresponde ao enquadramento teórico no qual, com base na literatura consultada, são revistos os conceitos de voz e disфония, são descritos e explorados vários métodos e medidas de avaliação áudio-percetiva e avaliação acústica da voz e, por fim, são apresentados os objetivos do estudo. No capítulo três são descritos os métodos utilizados na investigação, nomeadamente os procedimentos para a obtenção das amostras vocais e para as diferentes análises dos dados (análise áudio-percetiva e acústica da voz e tratamento estatístico dos dados). O capítulo quatro foca os resultados obtidos, sendo feita a descrição demográfica da amostra, a

sua caracterização vocal e apresentados os resultados das correlações obtidas. A discussão dos resultados é feita no capítulo cinco. Por fim, o capítulo seis apresenta as conclusões gerais do estudo, assim como sugestões para trabalhos futuros na mesma linha de investigação.

2. Enquadramento Teórico

Neste capítulo serão abordados conceitos relacionados com a voz, disfonia e avaliação (áudio-perceptiva e acústica) da qualidade vocal que suportam este estudo. No final serão definidos os objetivos da investigação.

2.1. Definição de voz e de disfonia

A voz é um som perceptível ao ouvido que resulta de uma complexa relação entre a pressão e velocidade do fluxo de ar expirado, os diferentes padrões de mobilidade das pregas vocais, as características de reflexão e configuração do trato vocal (Guimarães, 2007). A produção vocal é multidimensional, envolvendo mecanismos fisiológicos, biomecânicos e aerodinâmicos (Patel et al., 2018); e resulta de uma atividade muscular complexa, dependendo da integridade dos tecidos do aparelho fonador (Behlau, Azevedo, & Pontes, 2001).

A voz possibilita a comunicação (através da fala) sendo um importante meio de participação da pessoa nos diversos contextos da vida, dado que permite expressar diferentes ideias, emoções e intenções, e assume relevância como “biomarcador de identidade pessoal” (SPTF, 2020), sendo uma das mais fortes expressões de personalidade (Behlau, Azevedo, et al., 2001). A saúde vocal e a saúde mental estão relacionadas, tendo sido já demonstrado o efeito positivo do treino vocal na melhoria tanto de uma quanto de outra (Nusseck, Immerz, Spahn, Echternach, & Richter, 2019).

Para Aronson (1985 in Pinho, 1998) as características da voz normal são: altura tonal da voz¹ apropriada à idade e sexo, intensidade vocal subjetiva² apropriada ao ambiente, qualidade agradável e sem ruídos, e flexibilidade. Por outro lado, quando existem alterações nas características e parâmetros vocais, percebidos pelo falante ou pelo ouvinte, que impeçam ou dificultem a emissão vocal, considera-se perturbação vocal.

Existem várias definições de **disfonia** ou perturbação da voz: o desvio num ou mais parâmetros vocais (SPTF, 2020); o efeito da alteração de uma ou mais características acústicas da voz (Le Huche, &

¹ Altura tonal da voz – originalmente, em inglês, *pitch*, corresponde à sensação auditiva da frequência fundamental vocal, numa escala de grave a agudo (SPTF, 2020)

² intensidade vocal subjetiva (Int_{subj}) – originalmente, em inglês, *loudness*, corresponde à percepção auditiva de amplitude, altura de uma onda sonora, numa escala de fraco a forte (SPTF, 2020).

Allali, 2003); a impossibilidade de produção natural da voz devido a uma qualquer perturbação na emissão vocal, manifestada por alterações vocais e/ou cinestésicas (Behlau, Madazio, Feijó, & Pontes, 2001); a alteração da qualidade vocal, da intensidade, da frequência fundamental ou da flexibilidade vocal para a idade, género ou cultura do indivíduo (Aronson & Bless, 1990; SPTF, 2020); “dificuldade na emissão da voz com suas características naturais” (Tarneaud, 1941 in Pinho, 1998); perturbação de comunicação que tem significância pessoal, social e económica (Aronson, 1980 in Fawcus, 2001; SPTF, 2020) e que pode limitar a participação do indivíduo (SPTF, 2020).

As alterações da qualidade vocal podem estar associadas a causas orgânicas (alterações estruturais, neurológicas, endócrinas e patologias/lesões da laringe), funcionais e psicogénicas que afetem o sistema de produção vocal (Greene & Mathieson, 1989 in Fawcus, 2001).

Como já foi referido, a avaliação da voz é multidimensional e multidisciplinar. O diagnóstico de disфония deve envolver vários procedimentos específicos que incluam o diagnóstico clínico e a avaliação da qualidade vocal. A avaliação da qualidade vocal é a primeira fase da intervenção e reabilitação em voz (Behlau, Madazio, et al., 2001). Implica, primeiro, a realização de uma entrevista estruturada. Depois, a realização de um exame orofaríngeo, a avaliação da interação respiração-fonação, a avaliação áudio-percetiva da voz, a avaliação instrumental (laríngea, acústica e aerodinâmica). Também deverá ser contemplada a auto-avaliação do impacto psicossocial da disфония, de forma informal ou formal através de questionários construídos para o efeito (SPTF, 2020).

Em suma, há dificuldade em avaliar a qualidade (ou a perturbação) da voz falada com um método único e de maneira ampla e precisa, devido ao facto de a voz falada contemplar o sistema mais complexo da comunicação humana (Behrman, 2005; Guimarães, 2007).

Considerando o objeto do presente estudo, esta investigação centra-se na avaliação áudio-percetiva e na avaliação acústica da voz falada.

2.2. Avaliação áudio-percetiva da voz

A qualidade vocal é um fenómeno perceptual (Barsties & De Bodt, 2015) e corresponde à sensação evocada no ouvinte pelo sinal acústico da voz (Brinca, Batista, Tavares, Pinto, & Araújo, 2015).

A avaliação áudio-percetiva da voz baseia-se na apreciação de amostras de voz segundo vários parâmetros de qualidade vocal (Guimarães, 2007). É fundamental para a prática clínica, continua a ser o *gold standard* da avaliação da qualidade vocal (Barsties & De Bodt, 2015) e tem sido usada para a validação de medidas acústicas (Gerratt, Kreiman, Antonanzas-Barroso, & Berke, 1993; Kreiman & Gerratt, 2002). Dado o seu carácter perceptivo, é uma avaliação subjetiva. Diversos estudos demonstram que existe maior consenso na avaliação áudio-percetiva de vozes normais e extremamente alteradas comparativamente a vozes com alterações ligeiras e moderadas (Cannito, Woodson, Murry, & Bender,

2004; Dedivitis et al., 2004; Gerratt et al., 1993; Kreiman, Gerratt, Kempster, Erman, & Berke, 1993; Rabinov, Kreiman, Gerratt, & Bielamowicz, 1995) e que a fiabilidade das avaliações aumenta com o aumento do grau de disфония (Kreiman & Gerratt, 2011; Rabinov et al., 1995). Os avaliadores com maior experiência em perturbações da voz apresentam maior consenso na avaliação áudio-percetiva do que avaliadores sem esta experiência (Rabinov et al., 1995; Sofranko & Prosek, 2012; Zraick, Wendel, & Smith-Olinde, 2005). A qualidade da avaliação áudio-percetiva depende de vários fatores: treino e tempo de experiência do avaliador, conhecimento do diagnóstico médico, tipo e duração do estímulo vocal, grau da patologia, instruções da prova de avaliação da qualidade vocal e o fornecimento de estímulos âncora, tipo de escala e a dimensão/número de amostras a avaliar (Almeida, 2016).

Frequentemente, é pela avaliação áudio-percetiva que o indivíduo percebe a alteração vocal, e é também pelos correlatos desta avaliação que o profissional baseia grande parte da sua orientação. (Carding, Wilson, MacKenzie, & Deary, 2009)

São várias as escalas que permitem avaliar a voz de forma áudio-percetiva. As escalas de avaliação podem centrar-se em avaliar a fonte (nível laríngeo) e o filtro (nível supralaríngeo). Quando se considera a alteração de vibração da mucosa das pregas vocais e se avalia o som produzido na fonte glótica, analisa-se a fonte (Pinho & Pontes, 2008). Para este tipo de avaliação existem várias escalas, tais como:

- GRBAS (G – *grade* – avaliação global da disфония; R – *roughness* – aspereza/rugosidade; B – *breathiness* – sopro; A – *astheny* – astenia; S – *strain* – tensão) (Hirano, 1981 in Guimarães, 2007);
- RASAT/RASATI (rouquidão, aspereza, sopro, astenia, tensão, instabilidade) (Pinho & Pontes, 2008);

Quando também se considera a passagem dos harmónicos produzidos pela fonte glótica através do trato vocal, analisa-se o filtro, englobando aspetos relacionados com a articulação verbal e a ressonância (Pinho & Pontes, 2008). São exemplos de escalas que avaliam o filtro:

- *Consensus Auditory-Perceptual Evaluation of Voice* (CAPE-V) e CAPE-V EP 2ª;
- *Bufallo III - Voice Screening Profile* (Wilson, 1987 in Freitas, 2012);
- *Vocal Profile Analysis Scheme* (VPAS) (Laver et al., 1980 in Beck & Schaeffler, 2015).

Na Tabela 1 indicam-se as principais características das escalas referidas.

TABELA 1 - CARACTERÍSTICAS DE ESCALAS DE AVALIAÇÃO PERCEPTIVA DA VOZ

Escala	GRBAS; RASAT(I)	CAPE-V EP 2ª	Bufallo III	VPAS
Parâmetros	Só considera nível laríngeo Grau de alteração vocal, rouquidão, aspereza, sopro, tensão, astenia, instabilidade	Grau severidade global, rouquidão, sopro, tensão, altura tonal da voz, intensidade, ressonância, consistência, inconsistência; possibilidade de acrescentar parâmetros	Tónus laríngeo, timbre, intensidade vocal subjetiva, altura tonal da voz, ressonância nasal, ressonância oral, suporte respiratório, musculatura de abuso/esforço vocal, o débito da fala, a ansiedade de fala, inteligibilidade da fala e eficiência vocal geral	31 parâmetros, baseados na descrição fonética, divididos em três categorias: - recursos do trato vocal - tensão muscular (laríngeo e trato vocal) - qualidades do som da voz (<i>falseto, creak, creaky, whisper, whispery, breathy, harsh, tremor</i>)
Treino formal	Não	Não	Não	Sim
Provas	Discurso espontâneo e/ou vogal sustentada	Leitura frases, discurso espontâneo, vogais [a, i] sustentadas	Leitura, discurso espontâneo, contagem de números, vogal sustentada	Leitura, discurso espontâneo
Classificação	0 a 3	Escala visual analógica	1 a 5	0 a 4

A CAPE-V é o instrumento de avaliação áudio-perceptiva selecionado para o presente estudo (ver Capítulo 2.3). Foi originalmente desenvolvido pela *American Language-Hearing Association* em 2002 e depois revisto em 2009 (Kempster, Gerratt, Abbott, Barkmeier-Kraemer, & Hillman, 2009). O objetivo principal é estimular a implementação e registo padronizados no sentido de aumentar a consistência da avaliação e a troca de informação entre profissionais. Foi inicialmente traduzido e adaptado para o Português Europeu (PE) no âmbito da elaboração do Protocolo de Avaliação da Qualidade Vocal da Universidade de Aveiro (Santos, 2009) e posteriormente revisto por Almeida (2016) com o intuito de aumentar a sua validade e fiabilidade (sobretudo pela adaptação das frases da prova de leitura).

Neste instrumento utilizam-se as seguintes tarefas fonatórias: produção de vogais sustentadas ([a] e [i]), leitura de seis frases com diferentes contextos fonéticos e produção de discurso espontâneo. São considerados os seguintes parâmetros vocais: grau de severidade global, rouquidão, sopro, tensão, altura tonal, intensidade (subjetiva) e ressonância. Para a avaliação de cada parâmetro utiliza-se uma escala visual analógica com uma linha de 100mm. A escala à esquerda indica nenhum desvio e à direita refere-se ao máximo de alteração/desvio. Abaixo da linha encontram-se indicadores qualitativos: DL- desvio ligeiro, DM – desvio moderado, DS – desvio severo. Avalia-se ainda a consistência de cada parâmetro nas diferentes tarefas. Existe a possibilidade de indicar outros parâmetros que se afigurem relevantes no campo “Fatores adicionais” (por exemplo bitonalidade) (Almeida, 2016).

2.3. Avaliação acústica da voz

A avaliação acústica da voz, ou análise acústica da voz, refere-se ao “estudo de caracterização das propriedades acústicas dos sons da voz” (SPTF, 2020). Este tipo de avaliação faz medições do sinal sonoro da voz (Behlau, Madazio, et al., 2001). É uma avaliação instrumental, usada como ferramenta complementar da prática clínica, permitindo obter dados quantitativos na avaliação da qualidade vocal e na monitorização do progresso no decorrer do processo terapêutico, com a vantagem de ser não invasiva (Parsa & Jamieson, 2006). Na avaliação acústica da voz são considerados vários parâmetros que constituem o sinal acústico: periodicidade, amplitude, duração e composição espectral (Guimarães, 2007).

No sentido de aumentar a qualidade do diagnóstico e da intervenção, a evidência científica continua a demonstrar que, nos procedimentos de diagnóstico não invasivos, tanto as medidas perceptivas como as instrumentais contribuem para uma avaliação mais integrada. Para além do mais, os métodos instrumentais não invasivos (onde se inclui a avaliação acústica da voz) permitem um uso mais frequente no momento da avaliação e durante a intervenção, em especial pelos terapeutas da fala (Roy et al., 2013). As medidas acústicas podem atribuir objetividade na avaliação da qualidade vocal. Para além disso, certos parâmetros traduzem componentes fisiológicos relacionados com o comportamento vibratório das pregas vocais (Lowell, Colton, Kelley, & Hahn, 2011).

A literatura identifica vários parâmetros vocais caracterizáveis pela análise acústica. Contudo, para garantir a objetividade e validade das avaliações é fundamental ter em conta vários cuidados, nomeadamente ao nível de: condições de registo (por exemplo, ruído de fundo ambiente, posição e distância ao microfone, postura do indivíduo), protocolos de gravação (tipo, forma de produção e duração das amostras de fala), sistemas de gravação (escolha do microfone, armazenamento), análise qualitativa do sinal acústico e sistema de extração dos parâmetros de análise (diferentes instrumentos e/ou algoritmos utilizados não permitem comparação de medidas obtidas). (Behlau, Madazio, Feijó, et al., 2001; Patel, Eadie, et al., 2018; Švec & Granqvist, 2010, 2018)

2.3.1. Medidas acústicas de curto termo

De entre as diversas medidas acústicas de curto termo existentes destacam-se algumas das mais utilizadas: frequência fundamental, intensidade, medidas de perturbação (*jitter*, *shimmer*) e ruído (*Harmonic-to-Noise Ratio -HNR*).

Frequência fundamental (f_0) – é o parâmetro físico que mede a velocidade de vibração das pregas vocais por segundo (Vieira & Rosa, 2006). São várias as unidades de medida utilizadas para expressar a f_0 : ciclos por segundo (cps), hertz (Hz), semitons (ST) e oitavas (Freitas, 2009). O valor de f_0 varia de acordo com o género, idade, treino vocal, tarefa vocal avaliada, hábitos pessoais e estilo de vida e presença de alteração vocal (Freitas, 2009; Guimarães, 2007).

Os mecanismos fisiológicos responsáveis pela alteração da f_0 resultam, sobretudo, da interação de dois músculos intrínsecos da laringe (cricotiróideo e tiroaritenóideo) que modificam o comprimento e tensão das pregas vocais e alteram a sua velocidade de vibração e a massa colocada em vibração. Além disto há uma ação complementar dos músculos extrínsecos da laringe (Zemlin, 2005) que influencia a tensão do sistema (Behlau, Azevedo, & Madazio, 2001; Zemlin, 2005).

De forma a quantificar a dinâmica vocal deverão ser usados os valores de f_0 média, desvio-padrão de f_0 , f_0 máxima e f_0 mínima, permitindo descrever medidas centrais e intervalos de variação (gama tonal e extensão tonal) (Patel et al., 2018; Baken & Orlikoff, 2000). Nas alterações da qualidade vocal está muitas vezes presente a diminuição da extensão vocal (Zemlin, 2005). A extensão vocal corresponde ao número de notas que um indivíduo é capaz de emitir, desde a mais grave à mais aguda. Nas perturbações da voz, é frequente a redução da extensão vocal (especialmente nos diagnósticos de lesões de massa, fendas glóticas ou após abuso vocal) ou o seu deslocamento para zonas mais graves (casos de Edema de Reinke) ou mais agudas (na muda vocal incompleta, falsete) (Behlau, Madazio, et al., 2001).

As medidas de f_0 podem ser calculadas em diferentes tarefas vocais. Se forem determinadas a partir de tarefas de curta-duração (por exemplo, vogais prolongadas) poderão ser consideradas medidas de curto termo. No caso de serem calculadas a partir de tarefas de maior duração (por exemplo, fala espontânea ou leitura), poderão ser consideradas medidas de longo-termo.

Em relação à variação de f_0 nas tarefas, é expectável que as variações sejam diferentes consoante as tarefas avaliadas. Na sustentação de vogais espera-se menor variação (até 2Hz) e caso tal não aconteça pode significar desvio. No caso de tarefas de fala encadeada, é esperada maior variabilidade, indicando capacidade de modulação da voz (Behlau, Madazio, et al., 2001).

Na Tabela 2 são apresentados valores de referência para a f_0 .

TABELA 2 - VALORES DE FREQUÊNCIA FUNDAMENTAL NA POPULAÇÃO ADULTA COM E SEM PATOLOGIA VOCAL EM DIFERENTES TAREFAS VOCAIS

Estudo	Amostra	Presença disfonia	f_0 Vogal [a] Hz (dp)	f_0 Leitura (dp)	f_0 Discurso espontâneo (dp)
Vários estudos in Guimarães & Abberton (2005)	Mulheres [18–52 anos]	Não	199–215 Hz	180–242 Hz	187–210 Hz
	Homens [17–49 anos]	Não	111–125 Hz	106–129 Hz	109–123 Hz
Guimarães & Abberton (2005)	Mulheres [19–67 anos]	Não	210.5 (33)	190 (21)	186.6 (19)
		Sim	200 (37)	183 (35)	179.7 (36)
	Homens [20–63 anos]	Não	118 (18)	110 (20)	110 (13.6)
		Sim	113 (37)	111 (36.5)	108 (27.7)

Legenda: dp – desvio-padrão

Na patologia vocal, os valores de f_0 apresentam-se frequentemente alterados. Por exemplo, podem apresentar-se valores mais reduzidos nas lesões de massa e nos aumentos de massa; valores mais elevados nos casos de rigidez de mucosa das pregas vocais; e valores irregulares nas perturbações da voz de causa neurológica ou psicogénica, por exemplo (Behlau, Madazio, et al., 2001)

Intensidade – A intensidade sonora refere-se a um fluxo de energia. É definida como a potência que passa por uma área perpendicular à direção de propagação da onda sonora (Baken & Orlikoff, 2000; Gomes, 2005). A intensidade vocal é medida na unidade nível de pressão sonora (em inglês *sound pressure level* – SPL) em decibéis (dB) (Baken & Orlikoff, 2000).

A intensidade vocal pode ser ajustada a nível glótico (pelo controlo da resistência glótica realizado pelos músculos adutores da laringe, sobretudo o músculo tiroarínoides), subglótico (o aumento da potência aerodinâmica faz aumentar a pressão subglótica e, conseqüentemente, a intensidade) e supraglótico (a ressonância do trato vocal supraglótico interfere na distribuição espectral da energia acústica, funcionando como um filtro que modula o som glótico) (Imamura, Tsuji, & Sennes, 2006).

A variação da intensidade vocal no ser humano é grande, podendo chegar aos 140dB (Behlau, Azevedo, et al., 2001). Voz com valor de intensidade inferior a 50 dB é percebida como voz de intensidade subjetiva (Int_{subj}) fraca. Vozes com valor de intensidade acima dos 80 dB são percebidas como vozes de Int_{subj} forte ou muito forte. (Behlau, Madazio, et al., 2001). A percepção de intensidade é maior para frequências médias e diminui nas frequências mais graves e mais agudas (Jackson-Menaldi, 2002).

Deverão ser usados, para além do valor da intensidade vocal habitual, os valores médios de intensidade, desvio-padrão, máximo e mínimo (Patel et al., 2018). Obtêm-se elementos de valor diagnóstico da disfonia a partir das medidas de intensidade mínima e máxima durante a produção de uma vogal sustentada (Jackson-Menaldi, 2002). A intensidade de fala é frequentemente reduzida nas perturbações da voz e fala, especialmente naquelas de origem neurológica ou associadas a patologias respiratórias (Baken & Orlikoff, 2000).

A intensidade também varia segundo a tarefa vocal, sendo bastante diferente caso se trate de fala encadeada ou de sílabas isoladas (Baken & Orlikoff, 2000). Tal como nas medidas de f_0 , se as medidas de intensidade forem determinadas a partir de tarefas de curta-duração, poderão ser consideradas de curto termo. Por outro lado, serão consideradas medidas de longo-termo, se forem calculadas a partir de tarefas de maior duração. A sua medição é complexa, depende da distância da fonte sonora ao microfone, sendo importante a normalização das condições de gravação e a calibração dos equipamentos (Jackson-Menaldi, 2002; Švec & Granqvist, 2018).

A **extensão dinâmica** diz respeito à faixa de variação de intensidade que um indivíduo é capaz de emitir, desde a mais fraca (excluindo o sussurro) à mais forte (excluindo o grito). É expressa em diferença de decibéis e medida com um medidor de nível de pressão sonora. A extensão dinâmica será maior nas f_0

médias e menor nos extremos grave e agudo. Nas perturbações da voz os limites superiores poderão estar diminuídos e os limites inferiores aumentados (Behlau, Madazio, et al., 2001).

Quando os registos dos níveis de intensidade que um indivíduo consegue produzir ao longo da sua extensão vocal são combinados obtém-se o perfil de extensão vocal, cuja representação gráfica é o fonetograma (*Voice range profile*). Ao considerar estas duas dimensões – f_0 e intensidade – o gráfico obtido representa um conjunto de dados mais propício a descrever a alteração vocal, quantificar resultados da intervenção ou do treino vocal (um aumento da área do fonetograma pode indicar melhoria da capacidade vocal), e orientar parte do treino vocal, do que quando consideradas as medidas de intensidade ou f_0 isoladas (Behlau, Madazio, et al., 2001; Hallin, Fröst, Holmberg, & Södersten, 2012; Ma & Yiu, 2006; Siupsinskiene & Lycke, 2011).

As variações de entoação e a ênfase dada às palavras são qualidades da voz normal, permitidas pelas variações de f_0 e intensidade da voz. No caso de vozes hiperfuncionais, por exemplo, Stepp et al. (in Lowell, Kelley, Awan, Colton, & Chan, 2012) mostraram variações reduzidas de f_0 para o início e fim do vozeamento em tarefa de fala encadeada. Os autores explicam que a hiperfunção dos músculos intrínsecos da laringe amortece as variações naturais de f_0 associadas às variações de voz. Os mesmos autores sugeriram que a hiperativação dos músculos intrínsecos da laringe pode produzir uma tensão generalizada nas pregas vocais que amortece as flutuações normais de f_0 associadas às transições de vozeamento.

As variações na qualidade vocal envolvem modificações em todo o trato vocal, em que os aspetos de ressonância contribuem para a modificação da curva espectral do som. Como exemplo, vozes de qualidade rouca apresentam ruído entre harmónicos e, inclusive, ruído que substitui alguns harmónicos (Behlau, Azevedo, & Madazio, 2001).

Medidas de perturbação (*jitter*, *shimmer*) e ruído (*HNR*) – O *jitter* é uma medida que quantifica as variações da f_0 ciclo a ciclo. O *shimmer* é uma medida que tenta quantificar flutuações mínimas na amplitude do sinal ciclo a ciclo (Vieira & Rosa, 2006). Estas medidas devem ser obtidas através da análise de vogais sustentadas, mantendo a altura tonal e Int_{subj} estáveis, e excluindo o início e o final da emissão (Freitas, 2012).

O *jitter* e o *shimmer* possuem medidas relativas que podem ser apresentadas de várias formas, por exemplo: *jitter* (jitt %), *Pitch Perturbation Quotient* (ppq %), *jitter* absoluto (μ s), *shimmer* (shim %), *shimmer* (ShdB dB) (Freitas, 2012).

Para o *jitter* (jitt %), valores acima de 1% podem ser considerados indicadores de alteração. No caso do *shimmer* (shim %), expresso em percentagem, qualquer valor acima de 3% é indicador de alteração (Freitas, 2012).

O *jitter* e o *shimmer* procuram identificar mudanças na f_0 , o que se afigura possível nas vozes periódicas (vozes normais ou ligeiramente alteradas). Contudo, o aumento da perturbação vocal torna as

vozes menos periódicas, o que diminui a fiabilidade destas medidas neste contexto (Heman-Ackah et al., 2003; Lowell et al., 2011).

A análise acústica da voz com disфония apoiada em parâmetros de perturbação, dependentes da extração da periodicidade do sinal acústico, é considerada por alguns autores sem fiabilidade (Titze, 1995; Yiu, 1999), dado que a aperiodicidade naturalmente presente em vozes com disфония não é passível de ser estimada por muitos dos programas existentes. Acresce que as medidas de perturbação extraídas de amostras de fala encadeada são influenciadas pela entoação e efeitos de modulação, reduzindo a sua sensibilidade para identificação e caracterização da patologia vocal. Ainda não foi possível estabelecer um consenso sobre a capacidade de estas medidas distinguirem amostras vocais normais de patológicas ou de predizer a qualidade vocal ou o grau de severidade da disфония (Parsa & Jamieson, 2006). As medidas de perturbação quando comparadas com outras medidas que não dependem da identificação ciclo a ciclo, revelam-se inferiores (Maryn, Roy, De Bodt, Van Cauwenberge, & Corthals, 2009). De forma a terem significado clínico, estas medidas devem ser aplicadas com cuidado, sobretudo se usadas de forma isolada (Eadie & Doyle, 2005; Ma & Yiu, 2006).

O *HNR* (em português, Proporção Harmónico-Ruído) é uma medida que relaciona a componente periódica (correspondente à vibração das pregas vocais) com a componente não periódica (ruído glótico, sobretudo energia inter-harmónica), dando indicação da eficiência do processo de fonação.

A componente ruído é tanto maior quanto menores forem as componentes harmónicas da emissão (Freitas, 2012). O seu valor é expresso em dB e valores abaixo de 7dB serão indicativos de alteração, dependendo do programa usado. Existe correlação significativa entre valores de *HNR* e os parâmetros rouquidão e sopro (Freitas, 2012).

2.3.2. Medidas acústicas de longo termo

Recentemente, o interesse pelas medidas de longo termo baseadas nos espectros tem sido crescente, motivado pela necessidade de ultrapassar as limitações das medidas de curto termo. As medidas obtidas pelo LTAS (*Long-Term Average Spectrum ou* Espectro médio de longo termo) e pelo *Cepstrum* são exemplos de medidas de longo-termo.

Para se compreender melhor a forma de obtenção destas medidas, será importante rever alguns conceitos. O espectro de amplitude (amplitude em função da frequência), Figura 1, revela a complexidade do sinal vocal: é composto por uma f_0 e por suas múltiplas, os harmónicos. Determinados harmónicos são reforçados pelo trato vocal (Zemlin, 2005) e, quanto mais rígida for uma estrutura, menos harmónicos serão reforçados (Pinho, 1998). Por exemplo, no caso de rigidez nas pregas vocais de causa orgânica com reduzida flexibilidade da mucosa, o resultado será uma voz com poucos harmónicos (Pinho, 1998). As formantes são definidas pela forma e comprimento do trato vocal; as suas frequências são inversamente proporcionais ao comprimento do trato vocal (Zemlin, 2005).

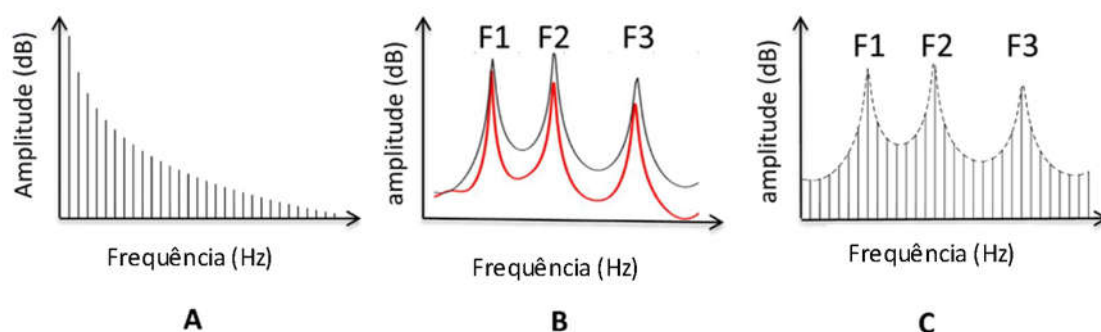


FIGURA 1 - ESPECTROS

A - Espectro do sinal da fonte glótica; B - Influência do trato vocal; C - Espectro acústico resultante da vogal

Imagem adaptada de Boë et al. (2019)

O espectro possibilita observar a distribuição da frequência e da amplitude num dado momento. Por sua vez, o espectrograma permite observar a evolução da frequência (eixo vertical) e da amplitude em função do tempo (eixo horizontal); a amplitude é representada pelo grau de escurecimento na gama de cinzento. O espectrograma é, assim, uma sequência de espectros alinhados (Mateus, Falé, & Freitas, 2005). O espectrograma permite estudar a estrutura dos harmônicos (relativos à fonte sonora) e da ressonância (relativa ao filtro). Para analisar a estrutura dos harmônicos utiliza-se o espectrograma de banda estreita em que se observam bandas horizontais que representam os harmônicos. As bandas mais escuras dizem respeito aos harmônicos que estão mais próximos dos picos de ressonância do trato vocal; este tipo de espectrograma permite identificar a f_0 e a entoação. Por outro lado, os espectrogramas de banda larga permitem distinguir as formantes (Jackson-Menaldi, 2002).

A análise acústica tradicional fundamenta-se na linearidade da Teoria Acústica da Produção de Fala, modelo fonte-filtro para a produção de vogais (Fant, 1970 in Zemlin, 2005). Segundo esta teoria, é assumida a independência da fonte-filtro e é declarado que a onda de fala é a resposta do trato vocal a uma ou mais fontes sonoras. Isto representa uma limitação (Freitas, 2012), dado que fonte e filtro na verdade interagem. A teoria não linear fonte-filtro (Titze, Riede, & Popolo, 2008) considera que existe interação dinâmica entre os vários componentes do sistema e os fenômenos responsáveis pela produção vocal passam pela interação entre a energia acústica e a mecânica, sendo que a pressão acústica do trato vocal influenciará o modo de vibração das pregas vocais. O aumento de energia do fluxo glótico poderá ser consequência da energia armazenada no trato vocal. Contudo, ainda hoje, os métodos de avaliação baseados na teoria linear fonte-filtro são os mais utilizados (Freitas, 2012).

A análise espectral baseia-se na análise computadorizada de Fourier através do método *Fast Fourier Transform* (FFT), que possibilita decompor um sinal complexo em cada uma das suas frequências parciais. Assim, os dados do domínio temporal passam ao domínio da frequência (Jackson-Menaldi, 2002).

A análise cepstral (*cepstrum*) resulta da aplicação de duas transformadas de Fourier, permitindo calcular a f_0 a partir da periodicidade do componente harmónico de um som, representado pelo pico cepstral da região alta do *cepstrum* (Jackson-Menaldi, 2002).

As medidas obtidas pelo LTAS e pelo *Cepstrum* têm revelado maior precisão nas vozes com disfonia, uma vez que não se baseiam no tempo, nem necessitam de identificação ciclo a ciclo. Além disso, são aplicáveis na avaliação acústica da fala encadeada, o que providencia uma amostra mais representativa dos padrões de fala com disfonia do que uma vogal sustentada consegue. Acresce que estas medidas não são influenciadas pela entoação (Heman-Ackah et al., 2003; Lowell et al., 2011).

As medidas do LTAS (Figura 2) dão informação sobre a distribuição de energia ao longo da frequência, e, por outro lado, as medidas obtidas a partir do *cepstrum* realçam a força relativa da energia dos harmónicos no sinal acústico (Lowell et al., 2011).

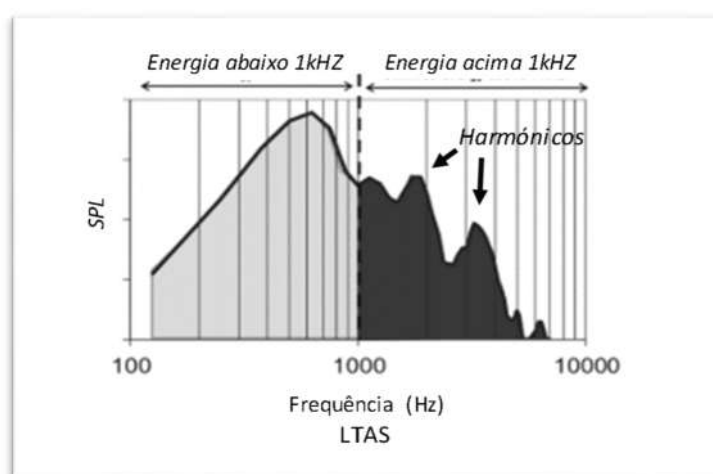


FIGURA 2- LTAS
Imagem adaptada de (Araújo, 2015)

Segundo Nordenberg e Sundberg (2003), o LTAS “reflete a contribuição tanto da fonte glótica como do trato vocal” na qualidade de uma voz. Permite analisar a média dos vários espectros de um segmento de fala de longa duração num só espectro. Considera, no eixo vertical, o nível de pressão sonora (dB) e, no eixo horizontal, a frequência (Hz). O tempo não é incluído nesta análise. A variação na intensidade tem impacto nos valores do LTAS.

As frequências das formantes F1 e F2 são representadas por uma média, sendo dada maior relevância às formantes F3, F4 e F5, mais relacionados com a qualidade vocal. Os picos ou regiões de maior concentração de energia do LTAS correspondem à extensão da variação da f_0 e das formantes. Os picos do LTAS estão fortemente relacionados com a percepção de diferentes qualidades de vozes (Hammarberg, Fritzell, Gauffin, Sundberg, & Wedin, 1981). Além da análise visual do próprio traçado, o LTAS permite a quantificação das medidas: *leq*, *alpha ratio* (α), *spectral slope*, relação entre a região de H1 e de H2 [$(H1-H2)_{LTAS}$] (Master, Biase, Pedrosa, & Chiari, 2006), que serão descritas a seguir. Para além

disso, o LTAS possibilita a análise de fala encadeada e estabiliza ao fim de 40 segundos (Fritzell et al., 1974 in Nordenberg & Sundberg, 2003).

A medida *leg* diz respeito ao nível de pressão sonora constante ao longo do tempo que é medido em dB e representa uma medida da intensidade média (da Silva et al., 2011)

O *alpha ratio* (α -alfa_{LTAS}) corresponde às diferenças de nível de energia entre 50-1000 Hz e entre 1-5 kHz. Tais diferenças refletem a inclinação (*slope*) do espectro, que, por sua vez, está relacionada com a velocidade de encerramento glótico. Pode relacionar-se diretamente com a intensidade (da Silva et al., 2011; Guzman, Correa, Muñoz, & Mayerhoff, 2013).

A medida *spectral slope* (*slope*_{LTAS}) refere-se à inclinação espectral, correspondendo à intensidade média de inclinação por unidade de frequência (Tamarit, 2008 in Araújo, 2015). Pode ser afetada pelo volume e por diferentes tipos de fonação, correlacionando-se com a qualidade vocal. O formato do fluxo aéreo transglótico, durante a fase de encerramento glótico, é importante na amplitude dos harmônicos superiores. Se a fase de encerramento for lenta ou insuficiente, o que pode ser indicativo de alteração, o nível espectral na região da formante superior poderá ser reduzido (da Silva et al., 2011). Por outro lado, uma maior velocidade da fase de encerramento origina uma inclinação menor. Uma fase de encerramento glótico mais rápida poderá ser indicativa de uma adução mais adequada e/ou melhor mobilidade das pregas vocais. (Leino, 2009).

(H1-H2)_{LTAS} (em inglês *L1-L0*) corresponde à diferença de energia entre o primeiro e o segundo harmônico, dando informação sobre a intensidade vocal e a fonte glótica, e relacionando-se com a qualidade vocal. Um valor de H1 superior a H2 está relacionado com vozes hipofuncionais, percebidas como soprosas e asténicas; por outro lado, um valor de H2 superior a H1 está relacionado com vozes hiperfuncionais e fortes percebidas como tensas e fortes (da Silva et al., 2011; Guzman et al., 2013).

Segundo Wendler et al. (1986 in Wolfe & Martin, 1997), o LTAS é capaz de predizer o grau de *roughness*, soproidade e rouquidão mas não discrimina entre qualidades vocais. O LTAS demonstrou ser sensível às diferenças de género, idade, vozes profissionais (tanto falada quanto cantada), à variação da qualidade vocal no pré e pós tratamento da disfonia (quer através da intervenção por Terapia da Fala, quer através da realização de cirurgia) (Master et al., 2006) e, ainda, apropriado para distinguir vozes com disfonia e vozes sem disfonia (Lowell et al., 2011).

Na Tabela 3 são apresentados valores de referência das medidas do LTAS:

TABELA 3- VALORES DE REFERÊNCIA PARA MEDIDAS DO LTAS

Estudo	Língua	Amostra	Tarefa	MEDIDAS		
				α_{LTAS} dB (dp)	$(H1-H2)_{LTAS}$ dB (dp)	leq_{dB}
da Silva et al. (2011)	Português (Brasil)	F	Leitura Int _{subj} média	-14.3	0.05	72
			Leitura Int _{subj} forte	-10	4.1	80
Leino (2009)	Finlandês	M	Leitura Int _{subj} média	-10.5		71
Araújo (2015)	Português (cantoras tradicionais)	F (cantadeiras)	DE	-8.83	0.4	79
			Leitura Int _{subj} média	-8.29	4.8	80
			Canto	0.32	-12.5	89.5
Guzman et al. (2013)	Espanhol (Atores)	M (atores)	DE	-18.4 ± 3.3	-4.2 ± 0.7	
		F (atrizes)		-17.5 ± 3.3	0.96 ± 5.6	
Ng, Chen, & Chan (2012)	Cantonês	M	Leitura Int _{subj} média	-15.8 ± 1.6		
		F		-16.4 ± 2.3		
	Inglês	M		-16.5 ± 1.9		
		F		-17.4 ± 2.3		

Legenda: F-género feminino; M-género masculino; DE-Discurso espontâneo; dp – desvio-padrão

O *Cepstral Peak Prominence* (CPP) - Figura 3 - é uma medida acústica da qualidade vocal, que tem sido considerada a mais promissora e, talvez, a mais robusta para medir o grau de severidade da disfonia (Maryn et al., 2009), sendo incluída nas mais recentes orientações de avaliação instrumental da voz (Patel et al., 2018).

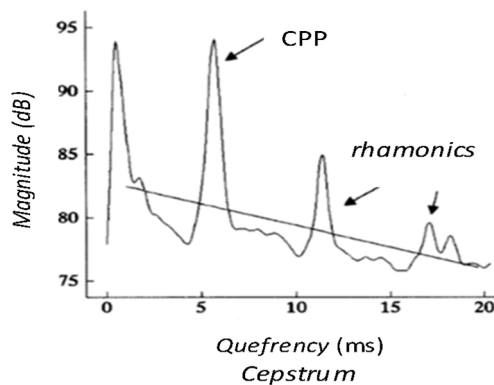


FIGURA 3- REPRESENTAÇÃO DO CEPSTRUM PARA VOZ SEM PERTURBAÇÃO
Imagem adaptada de Heman-Ackah et al. (2003)

O *cepstrum* é obtido a partir da aplicação da FFT no espectro; ou seja, o *cepstrum* é uma representação espectral do próprio espectro. O espectro passa a ser considerado como uma onda complexa, composta por várias ondas sinusoidais. Cada uma destas ondas tem características de amplitude e *quefrecy*³(para facilitar a terminologia, renomearam alguns dos termos). O pico

³ *Quefrecy* é a frequência de ocorrência da frequência no espectro de potência

predominante do *cepstrum* é o período fundamental do espectro que corresponde à *quefrequency* da onda sinusoidal dominante (tal como a f_0 é a frequência da onda sinusoidal da onda complexa do sinal vocal). Os pequenos picos de amplitude são denominados *rhamonics* (Heman-Ackah et al., 2003).

A partir do *cepstrum* obtêm-se várias medidas, entre elas o CPP e o *Cepstral Peak Prominence Smoothed* (CPPS), ambas expressas em dB. O CPP diz respeito à diferença entre o pico cepstral mais proeminente, ou primeiro *rhamonic*, e o ponto com a mesma *quefrequency* sobre a linha de regressão (Delgado-Hernández, León-Gómez, & Jiménez-Álvarez, 2019). O CPP fornece uma medida da diferença entre a energia da f_0 e a energia média do sinal derivado da regressão linear (Lowell et al., 2011) e reflete a relação global de energia periódica *versus* energia aperiódica de um sinal (Patel et al., 2018). Quanto mais periódico for o sinal acústico, maior o grau de harmonia e maior o valor de CPP. Vozes com grande alteração da qualidade apresentam estrutura menos harmónica e, por isso, a representação do *cepstrum* é tendencialmente plana e o valor de CPP é inferior (P S & Pebbili, 2020). Esta medida não depende da deteção da f_0 , logo tende a apresentar dados mais válidos mesmo nos sinais mais aperiódicos (Heman-Ackah et al., 2014). Fazendo a média de cepstras individuais, antes e depois do ponto de interesse, obtêm-se uma medida CPP suavizada – CPPS. Assim, calcula-se a média dos dados do sinal no *cepstrum*, reduzindo os artefactos, o que origina uma representação menos ruidosa da periodicidade relativa no sinal acústico (Lowell et al., 2011; P S & Pebbili, 2020).

Os valores de CPP e CPPS podem variar conforme a tarefa vocal, em parte devido às diferenças no grau de vozeamento. Os valores de CPP e CPPS obtidos a partir de vogais⁴ sustentadas são superiores aos obtidos em tarefas de leitura, provavelmente devido à maior estabilidade na fonação. A fala encadeada apresenta mais variações relacionadas com transições entre vogais-consoantes-vogais, com padrões de entoação e presença de sons não harmónicos no sinal acústico (Núñez-Batalla et al., 2019; P S & Pebbili, 2020). Os valores de CPP e CPPS também podem variar devido aos algoritmos utilizados pelos programas de análise acústica e pelas diferentes formas como podem tratar os segmentos não vozeados (Murton, Hillman, & Mehta, 2020). Na Tabela 4 são apresentados valores das medidas CPP e CPPS recolhidos de vários estudos, que foram obtidas no *Praat* em indivíduos sem disфонia. Estes valores podem ser usados como referência, apesar da variedade nas especificações ser diferente. Optou-se por apresentar apenas os valores obtidos no *Praat* dado que será o programa utilizado no presente estudo para obtenção destas medidas.

⁴ Apesar de o CPP e CPPS calculados a partir de vogal sustentada serem medidas de curto termo, o seu enquadramento teórico neste trabalho será feito no subcapítulo das medidas de longo termo

TABELA 4 - VALORES DE REFERÊNCIA DAS MEDIDAS CPP E CPPS (PRAAT) EM INDIVÍDUOS SEM DISFONIA

Estudo	Amostra	Medida	TAREFA			
			Vogal [a]	Leitura em Int _{subj} média	Discurso espontâneo	Fala em Int _{subj} forte
Phadke et al., 2018	Mulheres (n=40)	CPP	23.8	19	---	19.7
		CPPS	13.9	10.5	---	11.5
Mahalingam et al., 2020 (usaram definições Phadke 2018)	Homens e mulheres (n=50; 7+43)	CPP	26.4	---	19.71	---
		CPPS	15.04	---	9.49	---
Batthyany et al., 2019	Mulheres (n=52)	CPPS	15.08	7.59	---	---
	Homens (n=31)	CPPS	17.18	7.22	---	---

Alguns estudos estabelecem linhas de corte para valores de CPPS que poderão ser indicativos da presença de disfonia. No estudo de Heman-Ackah et al. (2014), valores de CPPS inferiores a 4 na tarefa de fala encadeada poderão significar presença de perturbação vocal (utilizando o algoritmo de Hillenbrand). O estudo de Delgado-Hernández et al. (2019) refere valores de CPSS (obtidos no programa *Praat*) para a tarefa de leitura, iguais ou inferiores a 8.37 dB, e para a tarefa de vogal [a] sustentada consideram-se valores de CPPS iguais ou inferiores a 13.96 dB. O estudo de Murton et al. (2020) apresenta resultados semelhantes: 14.45 dB para a tarefa de leitura e 9.33 dB para a tarefa de vogal [a] sustentada.

Vários estudos apontam correlação entre medidas cepstrais e a avaliação áudio-perceptiva. Está demonstrado que as medidas de CPP se correlacionam de forma significativa com o grau de disfonia, quer em tarefa de vogal sustentada, quer na fala encadeada (Núñez-Batalla et al., 2019). Nos casos com maior alteração da qualidade vocal, em que o sinal vocal é caracterizado por reduzida periodicidade, verifica-se redução da amplitude relativa do pico cepstral, isto é, redução do CPP (Watts & Awan, 2015). Lowell et al. (2012) referem vários estudos que indicam que a medida CPP se correlaciona fortemente com a soproidade, para além do grau de severidade global da disfonia percebidos auditivamente; por outro lado, a correlação é fraca com o correlato *roughness*. Lowell et al. (2012) encontraram valores de CPP significativamente baixos num grupo de indivíduos em que o parâmetro tensão era a característica mais evidente da alteração vocal, quando comparado com o grupo de indivíduos sem alteração da qualidade vocal. No estudo de Heman-Ackah, Michael, e Goding (2002), a partir da análise de vozes com disfonia causada por paralisia unilateral da prega vocal, foi encontrada correlação forte e inversa entre a medida CPPS e o grau de disfonia e o parâmetro soproidade. Núñez-Batalla et al. (2019) encontraram correlação significativa entre os valores de CPPS da vogal e fala encadeada (obtidos no *Praat*), o grau de disfonia e o grau de soproidade (escala GRBAS).

A comparação dos valores de CPPS obtidos a partir dos programas *Praat* e *Speech Tool*, revelou semelhança quase perfeita (92.4% da variância de CPPS do programa *SpeechTool* foi explicada pela variância do CPPS do programa *Praat*). Assim, ficou demonstrado que é possível usar qualquer um dos

programas para obter o CPPS (Maryn & Weenink, 2015). Apesar de os programas se basearem no mesmo algoritmo descrito por Hillenbrand, apresentam diferenças na forma como as medidas são calculadas e conseqüentemente nos valores absolutos obtidos (Sauder, Bretl, & Eadie, 2017).

Maryn et al. (2009) conduziram uma meta-análise sobre a correlação entre o grau de severidade da disфонia percebido auditivamente e várias medidas acústicas de avaliação da qualidade vocal. Foram analisados 25 estudos, correspondendo a 69 medidas acústicas obtidas a partir da análise de vogais sustentadas e 26 de fala encadeada. Destas medidas, apenas sete foram associadas à percepção da alteração vocal (ficaram acima da medida de corte determinada como grau forte de correlação), sendo consideradas as mais promissoras na avaliação acústica da qualidade vocal global. Quatro dessas medidas destinavam-se às vogais sustentadas e três à fala encadeada. Apenas uma medida – CPPS – apresentou validade suficiente quer para as vogais sustentadas, quer para a fala encadeada, sendo considerada a medida acústica com mais peso na medição acústica da severidade da disфонia.

Persistem muitas limitações na caracterização acústica de vozes roucas, apesar da utilização de métodos mais robustos como a análise cepstral (Awan & Awan, 2020).

2.3.3. Índices vs. medidas isoladas

Várias medidas acústicas têm-se revelado promissoras, de forma isolada, na maneira como se relacionam com o grau de gravidade da disфонia. No entanto, os modelos multidimensionais ou multiparamétricos, que combinam vários fatores com diferentes pesos relativos, apresentam correlações maiores ou maior capacidade para prever a gravidade da disфонia (Barsties & De Bodt, 2015; Hakkesteegt, Brocaar, Wieringa, & Feenstra, 2008; Lowell et al., 2012). São exemplos destes modelos multidimensionais: *Dysphonia Severity Index* (DSI), *Acoustic Voice Quality Index* (AVQI) e *Cepstral Spectral Index of Dysphonia* (CSID). Os dois últimos distinguem-se do primeiro por incluírem medidas obtidas a partir de tarefas de vogal sustentada e fala encadeada (Barsties & De Bodt, 2015).

O DSI (Wuyts et al., 2000) resulta da combinação de quatro medidas: tempo máximo de fonação (TMF), f_0 máxima, intensidade mínima e *jitter*. Estas quatro medidas foram selecionadas após uma análise multivariada a partir de 13 parâmetros. As medidas são obtidas através de tarefas de avaliação aerodinâmica (TMF da vogal [a]) e parâmetros da análise acústica de curto termo (*jitter* obtido da produção da vogal [a], pico de frequência f_0 e valor mínimo de intensidade que o indivíduo é capaz de produzir). O DSI é calculado através da fórmula:

$$0.13 \cdot \text{Tempo Máximo Fonação} + (0.0053 \cdot f_0 \text{ máxima}) - (0.26 \cdot \text{Intensidade mínima}) - (1.18 \cdot \text{jitter}) + 12.4$$

O resultado está relacionado com a avaliação áudio-perceptiva do grau global de disфонia (considerando a escala GRBAS em que o grau varia de 0 a 3). Ao grau 0 de alteração vocal (voz normal) percebido auditivamente corresponde o valor de DSI +5; ao grau 3 de alteração vocal (disфонia de

grau severo) corresponde o valor de DSI – 5. São possíveis valores de DSI inferiores a –5 e superiores a +5 (Wuyts et al., 2000). O estudo de Barsties v. Latoszek, Ulozaite-Staniene, Petrauskas, Uloza, & Maryn (2018) determinou o valor 3,05 a partir do qual o DSI terá precisão diagnóstica, distinguindo com base na avaliação laringoscópica.

Hakkesteegt, Brocaar, Wieringa, & Feenstra (2008) realizaram um estudo com 294 indivíduos com queixas vocais e patologia orgânica confirmada por ORL e 118 indivíduos no grupo controlo (sem queixas vocais). Não foi possível calcular o DSI para vozes percebidas auditivamente com grau 3 porque não foram obtidas medidas válidas relacionadas com a frequência (devido ao elevado grau de severidade de disфонia). Os dados foram concordantes com o facto de o resultado do DSI estar baseado no grau disфонia: valores de DSI significativamente reduzidos para valores elevados de grau de severidade de disфонia. Ainda foi determinado o ponto de corte a partir do qual o resultado do DSI discrimina entre controlos e pacientes: 3.

O género não parece influenciar os valores de DSI, muito provavelmente devido à compensação e oposição entre o pico de f_0 (maior nas mulheres) e o TMF (maior nos homens) (Barsties, Ulozaite-Staniene, Maryn, Petrauskas, & Uloza, 2017; Nemr, Simões-Zenari, de Souza, Hachiya, & Tsuji, 2016). Pelo contrário, os valores de DSI diminuem significativamente com a idade, o que pode ser explicado pela diminuição de f_0 , pela redução do TMF e pelo aumento do *jitter* (Baken, 2005; Barsties et al., 2017).

São também descritas outras influências sobre os valores de DSI: geográficas e étnicas (Jayakumar & Savithri, 2012), diferenças na fiabilidade entre equipamentos (Aichinger, Feichter, Aichstill, Bigenzahn, & Schneider-Stickler, 2012), e estado psicológico do sujeito durante a avaliação (Wuyts et al., 2000). De salientar que o DSI não mede a qualidade vocal; concentra uma combinação de medidas que covaria com o grau de severidade da disфонia (Freitas, 2012).

O *Acoustic Voice Quality Index (AVQI)* (Maryn, Corthals, Van Cauwenberge, Roy, & De Bodt, 2010) resulta de uma abordagem multiparamétrica que pretende medir a disфонia com recurso a tarefas de fala encadeada e vogal sustentada. Tem o objetivo de aumentar a validade da avaliação acústica da qualidade vocal e melhorar a distinção entre diferentes graus de gravidade de disфонia. Varia entre 0 e 10, representando um contínuo de gravidade desde normal (sem alterações) até qualidade vocal profundamente alterada. O AVQI foi validado usando-se na avaliação áudio-percetiva o parâmetro G da escala GRBAS. Entretanto, a fórmula foi sendo refinada, mas também ajustada a proporção entre a duração da vogal sustentada e da amostra de fala encadeada, de modo a aumentar a validade ecológica e a consistência interna. São agrupados 3 segundos centrais da vogal [a] sustentada e 34 sílabas da leitura de um texto foneticamente equilibrado em altura tonal e Int_{subj} considerados confortáveis (sendo retirados os segmentos não vozeados). A última versão 03.01 da fórmula de cálculo do AVQI é a seguinte (Barsties & Maryn, 2016):

$$\text{AVQI} = [4.152 - (0.177 + \text{CPPs}) - (0.006 * \text{HNR}) - (0.037 * \text{Shim}) + (0.941 * \text{ShdB}) + (0.01 * \text{Slope}) + (0.093 * \text{Tilt})] * 2.8902$$

Contempla parâmetros acústicos dos domínios tempo, frequência e *quefrequency*. O facto de incluir CPPS (medida acústica atualmente mais robusta) torna este índice uma base sólida para quantificar a gravidade da disфонia (Maryn & Weenink, 2015); também apresenta a vantagem de não sofrer influência do fator idade (Barsties e Maryn, 2013 in Pommée, Maryn, Finck, & Morsomme, 2018) ou género (Barsties et al., 2017).

Quando foi criado o AVQI era necessário recorrer a dois programas para o cálculo das diferentes medidas (*Praat* e *SpeechTool*); entretanto tornou-se possível calcular o AVQI apenas com o programa *Praat*, através da utilização de um *script*. Os valores de AVQI obtidos totalmente a partir do *Praat* são semelhantes à versão original (96% da variância dos valores de AVQI são explicados pela variância dos valores de AVQI obtidos exclusivamente a partir do *Praat*) (Maryn & Weenink, 2015).

O AVQI já foi validado em diferentes línguas (Delgado Hernández, León Gómez, Jiménez, Izquierdo, & Barsties v. Latoszek, 2018; Englert, Barsties v. Latoszek, Maryn, & Behlau, 2019; Hosokawa et al., 2019; Kim, von Latoszek, & Lee, 2019; Pommée et al., 2020) apresentando grande precisão de diagnóstico, com valores elevados de sensibilidade e especificidade; no entanto, em todas foram encontrados falsos positivos e falsos negativos. O limiar a partir do qual se considera qualidade vocal alterada varia nos diferentes estudos. Um estudo japonês definiu valor de 2,06, um estudo espanhol apresentou valor de 2,3 (Delgado Hernández et al., 2018), e um estudo alemão apresentou valor de 2,4 (Barsties & Maryn, 2016). No entanto, o AVQI para português europeu ainda não foi validado, não havendo valores de referência.

O *Cepstral Spectral Index of Dysphonia (CSID)* também é obtido a partir da análise da vogal sustentada e de fala encadeada, mas, ao contrário do AVQI, a análise das duas tarefas é feita de forma separada e contempla medidas espectrais e cepstrais. Varia entre 0 e 100 (embora possa, em alguns casos, apresentar valores fora deste intervalo). Já foi demonstrada a sua capacidade em distinguir vozes normais de vozes com qualidade alterada. Contempla as medidas CPP, *ratio of low-to-high frequency energy* e respetivos desvios-padrão. É calculado no programa *Analysis of Dysphonia in Speech and Voice* (não gratuito) (Watts, Awan, & Maryn, 2016).

Quando comparada a relação dos índices AVQI e CSID e a avaliação áudio-percetiva, utilizando as mesmas amostras vocais, verifica-se que as associações são mais fortes entre os valores de CSID e o grau de alteração vocal avaliado de forma áudio-percetiva (medido numa escala visual analógica) (Lee, Roy, Peterson, & Merrill, 2017).

2.3.4. Tarefas vocais

As tarefas vocais devem servir a obtenção de dados para melhor caracterizar a qualidade vocal do indivíduo. Devem ser adaptadas consoante o que se pretende avaliar: voz falada, voz cantada e as suas

especificidades. Em 2012, a ASHA aprovou a criação de um painel de especialistas para desenvolver um protocolo para avaliação instrumental da função vocal, publicado em 2018 (Patel et al.). Neste protocolo foram definidas tarefas para a avaliação endoscópica da laringe, análise acústica e avaliação aerodinâmica. Este protocolo deve ser complementado com a avaliação áudio-percetiva (é sugerida a utilização da CAPE-V) e escalas de auto-percepção do impacto da perturbação vocal.

Em relação à tarefa vocal – fala encadeada ou vogal sustentada – parece não haver consenso na literatura sobre a influência que a tarefa terá na avaliação perceptiva do tipo e gravidade da disфонia. Para os dois tipos de tarefa vocal não foi encontrada diferença estatisticamente significativa na percepção do grau de severidade da disфонia (Maryn & Roy, 2012; Maryn & Weenink, 2015). Noutro estudo, comparou-se a avaliação áudio-percetiva de três tipos de amostras vocais: fala encadeada, vogal sustentada completa e outra com parte da vogal sustentada (sem início nem final, tal como é utilizada para obtenção de medidas acústicas de perturbação). Sete avaliadores com experiência revelaram consistência nas avaliações. As avaliações acerca das vogais sustentadas sem início nem final atribuíram qualidade vocal menos alterada do que em relação à fala encadeada; as avaliações das vogais sustentadas completas mostraram-se semelhantes às amostras de fala encadeada (Revis, Giovanni, Wuyts, & Triglia, 1999).

Moers et al. (2012) compararam a avaliação áudio-percetiva (através da escala RBH e do parâmetro Grau) feita por 5 avaliadores a 73 casos com alterações vocais, com as medidas CPP e CPPS. Os resultados obtidos revelaram correlações mais fortes entre os parâmetros perceptivo-auditivos e acústicos na tarefa de fala encadeada, comparativamente com a tarefa de vogal sustentada.

Sabe-se que há vantagens na avaliação áudio-percetiva de vogais sustentadas, dado que permite que o avaliador se foque mais facilmente no sinal acústico. Este sinal à partida é mais estável (Revis et al., 1999) e não é influenciado por aspetos articulatórios ou prosódicos, tornando-se mais fácil estabelecer uma relação entre a qualidade vocal percebida e a fisiologia do movimento das pregas vocais. No entanto, as vogais sustentadas não são representativas do discurso espontâneo (Delgado Hernández et al., 2018); serão, por outro lado, mais representativas do canto (Parsa & Jamieson, 2006). Comparativamente à produção de vogais sustentadas, a fala encadeada é considerada mais natural, representando melhor os padrões de fala habituais de um indivíduo no dia-a-dia (Lowell et al., 2012), possui maior variabilidade na qualidade vocal e permite avaliar o efeito da coarticulação na qualidade vocal (Barsties & De Bodt, 2015). Deste modo, medidas baseadas na fala encadeada serão mais relevantes para a compreensão da fala habitual (Parsa & Jamieson, 2006).

Há acordo de que o protocolo de avaliação da voz falada deve contemplar tanto tarefas de vogal sustentada como de fala encadeada, importantes tanto para a avaliação áudio-percetiva como para a avaliação acústica (Patel et al., 2018).

Em relação às diferentes tarefas de fala encadeada – discurso espontâneo, semi-espontâneo e leitura – no estudo realizado por Pereira (2018), constatou-se que os resultados obtidos em todas as

tarefas de fala encadeada em Int_{subj} média são, com um nível de significância de 0.05, semelhantes ao discurso espontâneo para as medidas *alpha ratio*, $(H1-H2)_{LTAS}$ e *leq*. Deste modo, a avaliação poderá contemplar apenas uma tarefa de fala encadeada em Int_{subj} média e mais duas tarefas para avaliação de fala encadeada em Int_{subj} forte e fraca.

Halberstam (2004), no estudo que realizou com ouvintes treinados, concluiu que para o parâmetro rouquidão existe forte correlação com valores de CPP em tarefas de fala encadeada. Para as tarefas de vogal sustentada, a correlação entre o parâmetro rouquidão e os valores de *jitter*, *shimmer* e CPP foi menor. No estudo realizado por Eadie e Baylor (2006), o parâmetro *roughness* apresentou níveis de correlação maiores com o grau de severidade da disfonia em fala encadeada comparativamente às medidas obtidas a partir da vogal sustentada.

Lowell et al. (2012) encontraram diferenças significativas para um maior número de variáveis acústicas em situação de fala encadeada, comparando as mesmas variáveis para a produção de vogal sustentada. Nesse estudo, em que participaram indivíduos com disfonia com predomínio tensão e indivíduos sem disfonia, a fala encadeada foi sensível às variações das medidas desvio-padrão de CPP e desvio-padrão de *ratio of low-to-high frequency energy*, (medidas que se relacionam com a flexibilidade vocal). Os autores consideraram que, para os indivíduos com qualidade vocal tensa, o nível de flexibilidade vocal intrínseca à fala encadeada poderia ser exigente, explicando assim os resultados. Concluíram, deste modo, que a tarefa de fala encadeada, mais do que as vogais sustentadas, poderá ser particularmente reveladora da redução da flexibilidade vocal nos indivíduos com disfonia.

2.4. Objetivos

Considerando tudo o que foi apresentado anteriormente, a presente investigação parte dos seguintes pressupostos:

- a) A avaliação da qualidade vocal considerando apenas a análise de vogais sustentadas é insuficiente porque não são incluídos os aspetos próprios da dinâmica da fala encadeada (Hammarberg et al., 1981).
- b) As medidas acústicas de longo termo revelam potencial para a quantificação de alguns parâmetros da qualidade vocal, considerando todo o trato vocal, e tornam possível a obtenção de medidas acústicas em tarefas de fala encadeada, o que não se verifica nas medidas de curto termo. A obtenção das medidas de longo termo não é tão fácil e rápida, condicionando a sua utilização na prática clínica (Master et al., 2006).
- c) Na literatura são poucos os estudos realizados utilizando as medidas de longo termo na patologia vocal, considerando o grau de disfonia e relacionando-as com medidas acústicas de curto termo. Os estudos mais recentes apontam estas medidas de longo termo como promissoras. São exemplo as

medidas acústicas CPP e CPPS como marcadores acústicos da disфонia (Brinca et al., 2014; Maryn et al., 2009).

Neste sentido, o estudo proposto integra uma linha de investigação, já em curso na Escola Superior de Saúde do Politécnico do Porto (ESS-PP), sob o tema “Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação”. Os estudos já realizados neste âmbito caracterizaram a voz de falantes do Português Europeu sem alterações vocais (Pereira, 2018), testaram a aplicabilidade do LTAS na prática clínica, nomeadamente a sua sensibilidade à identificação do género (Portela, 2018) e da idade (Magalhães, 2018) e, ainda, exploraram diferenças entre a voz normal e as vozes de personagens-tipo produzidas por atores de dobragem (Daniel, 2018).

De forma a refletir sobre quais as tarefas e medidas que melhor servem a avaliação e caracterização da qualidade vocal, o presente estudo tem como objetivo central estudar a contribuição de tarefas de fala encadeada, medidas acústicas de longo termo e medidas combinadas para complementar a avaliação da qualidade vocal, tendo como referência medidas de curto termo e áudio-percetivas. Assim, definiram-se como objetivos específicos:

1. Investigar as correlações entre medidas acústicas de longo termo avaliadas em tarefas de leitura, em diferentes Int_{subj} , (média, forte e fraca) e a avaliação áudio-percetiva;
2. Verificar se existe correlação entre as medidas acústicas de longo termo e medidas acústicas de curto termo;
3. Verificar se existe correlação entre os índices AVQI e DSI e diferentes graus de disфонia avaliados de forma percetiva.

3. Métodos

Neste capítulo serão descritos os métodos escolhidos para levar a cabo a investigação e alcançar os objetivos propostos.

3.1. Tipo e desenho do estudo

Tendo em conta os objetivos anteriormente elencados, pretendeu-se avaliar e descrever vários parâmetros vocais de pessoas com diferentes graus de alteração vocal num único momento, analisando estatisticamente diferentes parâmetros de caracterização acústica. Esta análise pretende também aferir que tarefas e medidas melhor servem a avaliação e caracterização da qualidade vocal. Desta forma, foi realizado um estudo analítico do tipo observacional de carácter quantitativo transversal (Ribeiro, 1999).

Trata-se de um estudo observacional porque, de forma a descrever os acontecimentos sem a intervenção da investigadora sobre as variáveis, foram desenvolvidos vários procedimentos. O facto de se pretender encontrar relações entre variáveis e fazer descrições através do recurso ao tratamento

estatístico dos dados permite qualificar o estudo como quantitativo. Por fim, considerando que os dados serão recolhidos num único momento, este é um estudo transversal. (Ribeiro, 1999).

3.2. População e amostra

Este estudo foi orientado a uma população alargada de pessoas, com ou sem perturbações da voz, residentes em Portugal e cuja língua materna é o Português Europeu. A amostra, por questões de facilidade de acesso, foi limitada a utentes e funcionários do Hospital Pedro Hispano, tendo-se constituído uma amostra por conveniência, que é um tipo de amostragem não probabilística (Ribeiro, 1999). Como critérios de inclusão consideraram-se: ter mais de 18 anos, ser falante de Português Europeu como língua materna e saber ler. Como critérios de exclusão consideraram-se: ter alterações ao nível da audição e visão que impedissem ou influenciassem a realização do protocolo de avaliação; ter sido submetido a laringectomia total (ausência de fonte vocal glótica); ter alterações de voz que impedissem a análise acústica (nomeadamente grau extremo ou ausência de sinal vocal glótico); ter alterações ao nível da fala, nomeadamente perturbações motoras da fala, perturbações de fluência, perturbações articulatórias ou outras que pudessem afetar a perceção da qualidade vocal ou serem variáveis que pudessem induzir desvios aos resultados obtidos na ausência destas perturbações. Estes critérios foram referidos e/ou observados no momento do primeiro contacto. Foram consideradas válidas para a amostra do estudo pessoas sem disfonia e com disfonia, independentemente do seu grau ou tipo, pois só assim seria possível obter um espectro suficientemente alargado de vozes que permitiriam relacionar as variáveis em estudo.

De acordo com o cronograma da investigação, foi definido um período de três meses para a recolha das amostras vocais, com início em fevereiro de 2020, no fim do qual se avaliaria a necessidade de solicitar um prolongamento do prazo de recolha, com o objetivo de alcançar uma amostra próxima de 50 indivíduos. Este número de indivíduos está de acordo com o que seria espectável recolher e posteriormente permitir uma análise mais robusta. No entanto, a recolha foi interrompida quatro semanas depois, na sequência da suspensão da atividade não urgente do hospital, com a decretação de pandemia pela Organização Mundial de Saúde. Dois meses depois, após o início da retoma parcial da atividade não urgente do hospital, as circunstâncias não permitiam prosseguir a recolha com as mesmas condições de gravação. Assim, o período de recolha de dados foi substancialmente inferior ao previsto, o que afetou a dimensão final da amostra - 12 indivíduos - que foi por isso bastante inferior ao inicialmente planeado.

3.3. Instrumentos

Para o presente estudo desenvolveu-se o seguinte conjunto de instrumentos/protocolos de avaliação: a) ficha de caracterização do participante (Apêndice 1); b) protocolo de gravação de voz (Apêndice 2); c) protocolo de avaliação áudio-percetiva (Apêndice 3); d) *software* de análise de dados.

Foi criada uma ficha de caracterização dos participantes (Apêndice 1) considerando as informações acerca do género, idade e avaliação do médico ORL.

A recolha das amostras vocais consistiu na realização e gravação de diferentes tarefas vocais (Apêndice 2), adaptadas do protocolo de recolha de dados utilizado nos restantes estudos incluídos na linha de investigação já referida. Foi realizado um teste piloto para avaliar a aplicabilidade do Protocolo de Gravação da Voz. Solicitou-se a um participante voluntário, com características semelhantes aos indivíduos que participariam no estudo, que desse a sua opinião acerca dos procedimentos realizados. A partir das suas observações foi reestruturada a ordem das tarefas, assim como reformuladas algumas das estratégias usadas para orientar o protocolo, facilitando a sua compreensão e execução.

Para a gravação foi usado o sistema *lingWAVES Voice Suite Pro* (WEVOSYS®, versão 3.0), cedido pela área técnico-científica de Terapia da Fala da Escola Superior de Saúde – Politécnico do Porto (ESS-PP). O sistema *lingWAVES* inclui microfone com captação direta com sonómetro *lingWAVES* SPL meter II e, como tal, calibração automática do sinal. As amostras vocais foram gravadas em formato wav com frequência de amostragem de 44.1 kHz. Foi utilizado um computador portátil HP Intel® Core™ i3 64bits. Todas as gravações foram realizadas numa cabine insonorizada com um nível de ruído inferior a 40dB SPL. Nem todas as gravações de todas as tarefas foram utilizadas neste estudo, tendo ficado em base de dados para utilização em estudos futuros integrados na mesma linha de investigação.

Numa das fases do trabalho recorreu-se à avaliação áudio-percetiva por parte de um painel de peritos. Para esta avaliação foi desenvolvida uma adaptação da 2ª versão CAPE-V EP (Almeida, 2016), tendo-se limitado a análise à vogal [a] sustentada, de forma a evitar um procedimento demasiado longo, que viesasse os resultados por fadiga do avaliador. Foram mantidos os parâmetros grau de severidade global, rouquidão, sopro e tensão. Não sendo as tarefas de fala encadeada sujeitas a análise, os parâmetros altura tonal, intensidade e ressonância foram excluídos. A avaliação da consistência dos parâmetros entre tarefas também foi excluída. Como esta escala contempla gradações da gravidade (escala contínua), ao invés de pontos discretos, este foi o principal aspeto que motivou a escolha desta escala para o presente trabalho.

A avaliação áudio-percetiva realizada pelo painel de peritos foi registada numa página Web construída para o efeito (Apêndice 3). No sentido de medir a consistência interna de cada avaliador, houve necessidade de duplicar amostras de voz de forma aleatória para se calcular a fiabilidade *intra* e *inter*avaliadores.

As amostras vocais (vogais [a] sustentadas) foram agrupadas em grupos de cinco, e divididas em secções para facilitar a gestão do processo pelo participante. A cada amostra vocal correspondia um ficheiro áudio (com os botões “play” e “pause” que permitiam iniciar e interromper a reprodução), quatro cursores reguláveis (um para cada parâmetro contemplado na avaliação: Grau de severidade global, rouquidão, sopro e tensão) numa escala visual analógica e um campo de resposta aberta (Fatores

adicionais). As amostras vocais podiam ser ouvidas repetidamente, era possível avançar ou retroceder nas várias secções, assim como corrigir a avaliação já realizada.

Para a obtenção dos dados acústicos foram usados os programas: *lingWAVES Voice Suite Pro* (WEVOSYS®, versão 3.0), *Praat* (versão 6.1.09) e *Sopran* (versão 1.0.20). Para o cálculo de medidas específicas secundárias foi usado o *software* Microsoft Excel (Microsoft® Excel® 2016).

Em relação aos sistemas e programas utilizados, há algumas considerações a referir. Relativamente ao sistema *lingWAVES*, a grande vantagem para este estudo foi a calibração do sinal. A informação sobre resultados das medidas acústicas é visualizada de forma rápida. No caso das tarefas de fala encadeada, são fornecidas várias medidas de frequência (em Hz e semitons) e de intensidade. Também é gerado um gráfico que conjuga os valores de f_0 e intensidade da voz, semelhante ao fonetograma, e a frequência de ocorrência dos valores de f_0 e intensidade. Além disso, o programa permite a comparação direta da f_0 com notas musicais, fornecendo a indicação tanto em Hz quanto em semitons. Este facto auxilia, sobretudo na área do canto, na maior exatidão dos limites da função vocal. Também fornece *feedback* visual através de gráficos e permite comparar várias avaliações do mesmo indivíduo, o que pode ser interessante na prática clínica (Mendes, 2012). A grande desvantagem poderá estar no seu preço de aquisição.

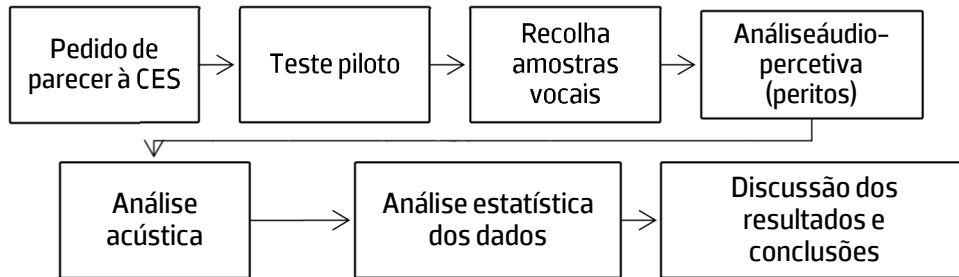
O programa *Praat* também reúne várias vantagens. Desde logo, o facto de ser disponibilizado de forma gratuita. Além disso, é atualizado de forma frequente, permitindo que na atualidade já seja possível a obtenção de muitas medidas de forma acessível. É já utilizado em muitos estudos e com a utilização de *scripts* é possível obter várias medidas de forma rápida e quase automática, bastando adaptar e renomear os ficheiros áudio. São disso exemplo os índices DSI e AVQI. Para o cálculo do índice AVQI foi utilizado o *script* de Maryn e Weenink (2015). O resultado apresentado pelo *Praat* é constituído pelos valores de cada uma das seis medidas que compõem o AVQI, pelo valor final do índice e por várias representações gráficas (entre elas, espectrograma, LTAS e *Cepstrum*) que também poderão ser alvo de análise. Além disso, O índice DSI pode ser obtido utilizando o *script* disponibilizado por Maryn, Morsomme e De Bodt (2017). No entanto, no presente estudo, o índice DSI foi calculado de forma "manual", dado ter havido um erro em dois dos indivíduos com a utilização do *script*. De qualquer forma, o índice DSI é relativamente fácil de obter e utilizar na prática clínica e apresenta boa consistência *inter* observador (Hakkestegt et al., 2008). No entanto existe inconsistência nos valores de DSI entre programas, consequência das diferenças nos algoritmos de deteção de f_0 ou da calibração insuficiente das medidas de nível de som (Aichinger et al., 2012).

Para a análise estatística dos dados foram utilizados os programas Microsoft Excel (Microsoft® Excel® 2016) e o Statistical Package for the Social Sciences (IBM® SPSS® Statistics, Versão 26).

3.4. Procedimentos

O presente estudo seguiu a ordem de procedimentos representada na Figura 4.

FIGURA 4 – PROCEDIMENTOS DO ESTUDO



O estudo iniciou-se com o pedido de parecer à Comissão de Ética em Saúde do Hospital Pedro Hispano que deu parecer positivo (Apêndice 4).

A primeira seleção dos indivíduos com patologia vocal foi realizada pelos médicos do Serviço de ORL do hospital, que, durante o período de recolha definido (3 meses), identificaram na sua consulta utentes que se enquadravam nos critérios de inclusão. Caso aceitassem, e depois de avaliados pelo médico ORL, os indivíduos eram contactados pela investigadora (telefonicamente ou, caso fosse possível, presencialmente no dia da consulta de ORL) que apresentou o estudo, explicando os objetivos e procedimentos. Também foram convidados a participar no estudo os utentes com patologia vocal encaminhados como habitualmente para Terapia da Fala no serviço de ORL, que se enquadrassem nos critérios de inclusão, antes de iniciarem tratamento. Por uma questão de conveniência, os indivíduos sem alterações vocais foram funcionários do hospital que aceitaram participar no estudo; também estes foram sujeitos a avaliação pelo médico ORL e todos negaram queixas vocais.

A cada indivíduo que aceitou de livre e esclarecida vontade participar no estudo foi apresentada e assinada a declaração de consentimento livre e esclarecido (Apêndice 5). A caracterização dos participantes (Apêndice 1) foi realizada a partir da recolha de dados extraídos do processo clínico e fornecidos pelo próprio indivíduo, considerando as informações acerca do género, idade e avaliação do médico ORL. Os fatores que pudessem determinar a não inclusão no estudo, e que já foram mencionados anteriormente, foram obtidos no contacto prévio (por telefone/presencial) e no momento da avaliação vocal. A declaração de consentimento livre e esclarecido (Apêndice 5), na qual constam os objetivos do estudo e os procedimentos de recolha, manutenção e tratamento dos dados recolhidos, foi entregue aos participantes.

Atribuiu-se um código a cada participante, de forma a garantir a confidencialidade dos dados fornecidos; este código foi usado tanto nas folhas de registo como nos ficheiros áudio relativos às amostras vocais. A gestão dos dados de identidade e respetiva relação com a codificação atribuída foi feita pela investigadora responsável num documento independente. A informação relativa ao

diagnóstico clínico, decorrente da avaliação feita pelo médico ORL, foi obtida através do processo clínico. Depois, estes diagnósticos foram agrupados segundo a classificação de Pontes, Behlau e Brasil (2000 in Behlau, Azevedo, et al., 2001).

Deu-se início à recolha das amostras vocais que consistiu na gravação de diversas tarefas de exploração vocal frequentes na avaliação da qualidade vocal (Apêndice 2). Todas as gravações foram realizadas numa cabine insonorizada e os indivíduos realizaram a gravação de pé. O microfone foi colocado a uma distância de 30 cm da boca dos indivíduos (a régua acoplada ao microfone permite verificar a manutenção da distância). O computador ficou em cima de uma mesa, colocado nas costas do participante evitando distrações que afetassem a sua produção. As frases, o texto e a letra da música "Parabéns a Você" foram impressos. Os ficheiros áudio foram guardados em formato wav.

3.4.1. Análise áudio-percetiva

As amostras vocais foram depois avaliadas de forma áudio-percetiva. Esta avaliação foi realizada por um painel de peritos composto por três terapeutas da fala com mais de 5 anos de experiência na área da voz (Almeida, 2016), que aceitaram de livre vontade participar nesta fase do estudo. Os contactos foram realizados pela investigadora/orientador e confirmados através de e-mail. Os peritos receberam informações acerca da escala CAPE-V, respetiva adaptação e breve descrição dos parâmetros a avaliar, servindo de orientação para a classificação áudio-percetiva.

Os peritos tiveram acesso à escuta das amostras vocais (apenas vogal [a] sustentada) e ao formulário para registo da sua avaliação através da página Web construída para o efeito.

Foram acrescentadas oito amostras vocais (a partir de uma base de dados existente) às doze amostras dos doze sujeitos do estudo, de modo a evitar os efeitos de familiaridade. Deste conjunto de vinte amostras vocais, foram duplicadas dez, mantendo a proporção em relação ao género e em relação ao número de amostras do estudo vs. amostras de voz de sujeitos que não pertenciam ao estudo. A ordem das amostras foi determinada aleatoriamente, salvaguardando-se que vozes iguais não ficaram no mesmo grupo de cinco. Depois, acrescentaram-se mais cinco amostras, colocadas no início do teste, com o objetivo de os peritos se familiarizarem com a escala. A avaliação destas cinco amostras não contou para efeitos estatísticos. No total, os peritos avaliaram um conjunto de 35 vogais sustentadas. Foram dadas indicações aos peritos sobre as condições mínimas para garantir a qualidade da avaliação, entre elas o uso de auscultadores, realizar a avaliação de todas as vozes no mesmo período e num ambiente calmo, usar os mesmos equipamentos e evitar mudanças no controlo do volume de áudio.

3.4.2. Análise acústica

A análise acústica foi realizada pela investigadora recorrendo a diferentes programas, para determinação das várias medidas acústicas de longo e curto termo.

Os valores das medidas de TMF, f_0 e intensidade das tarefas de leitura e respetivos desvios-padrão (dp) foram obtidas através do *lingWAVES Voice Suite Pro* (WEVOSYS®, versão 3.0). Os valores das medidas de f_0 média da vogal, desvio-padrão de f_0 da vogal, intensidade média da vogal e respetivos dp *jitter*, *shimmer* e *HNR* foram calculados no programa *Praat*, utilizando janela de 75-500Hz. Foi escolhido o *mid-vowel method*: foram selecionados três segundos centrais da vogal a analisar, sendo retirados o início e o final da emissão (Maryn & Weenink, 2015). O cálculo de DSI foi feito no Excel (Microsoft® Excel® 2016), considerando a equação já apresentada, e os valores de cada medida foram obtidos através do *Praat*.

Para extração do LTAS foi utilizado o *software Sopran®*, que permite retirar valores exatos num certo ponto e obter a lista dos principais valores de todo o espectro, entre outras funcionalidades. Foi preparado um ficheiro Excel (Microsoft® Excel® 2016) para registo dos valores obtidos no *Sopran®* e cálculo de medidas através de fórmulas previamente inseridas. O valor de *leq* médio foi obtido no *software Sopran®* selecionando a parte do ficheiro correspondente à tarefa e depois registado no ficheiro Excel (Microsoft® Excel® 2016). Para a extração dos espectros, e de forma a permitir uma melhor visualização do espectro, utilizaram-se os seguintes critérios de visualização: janela *Hamming*, um tempo de resolução de 40 milissegundos e uma banda de 50 Hz. Retiraram-se todos os valores de frequência e intensidade espectral de longo termo correspondente até aos 6000 Hz e inseriram-se no ficheiro Excel (Microsoft® Excel® 2016), obtendo-se, assim, através das fórmulas inseridas previamente, os valores finais de *leq*, *alpha ratio*, $(H1-H2)_{LTAS}$ e $slope_{LTAS}$.

Foi encontrada diferença qualitativa na análise dos gráficos do LTAS, nomeadamente o espaçamento entre as linhas das diferentes leituras, sobretudo entre aquelas em Int_{subj} forte e fraca. Por este motivo, fez-se o cálculo da diferença de intensidades médias das leituras em Int_{subj} forte e fraca. A esta diferença chamou-se **dinâmica de intensidades**.

Para obtenção das medidas CPP e CPPS na análise das tarefas de fala encadeada e da vogal [a] sustentada, foi utilizado o programa *Praat*. Para distinguir, no presente trabalho, as medidas CPP e CPPS nas tarefas de fala e vogal, decidiu-se utilizar as siglas CPP_{fala} e CPP_{vogal} , $CPPS_{fala}$ e $CPPS_{vogal}$.

Para obter as medidas CPP_{fala} e CPP_{vogal} no programa *Praat* seguiram-se as especificações definidas por Hillenbrand, Cleveland, e Erickson (1994) segundo Weenink (2019):

1. Sound object: Analyse Spectrum seleccionar To Spectrum, depois ok.
2. Spectrum object: Analyse, seleccionar To PowerCesptrum
3. Power Cepstrum object Query, seleccionar Get Peak Prominence com as seguintes definições: Search peak in pitch range: 60-333.3; Interpolation: Cubic; Trend line quefrequency range: 0.001-0; Trend Type: straight; Fit method: robust.

Para a obtenção da medida CPPS no programa *Praat* seguiram-se os três passos indicados por Brockmann-Bauser et al. (2019) (baseados em Maryn & Weenink, 2015):

1. Sound object: Filter (stop Hann Band). Settings: 0 to 34 Hz, Smoothing 0.1 Hz;
2. Filtered sound object: Analyse Periodicity, seleccionar To Power Cepstrogram (standard settings);
3. Power Cepstrogram object: Query, seleccionar Get CPPS. Definições: Subtract tilt before smoothing: "no," Time averaging window: 0.01 s, Quefreny window: 0.001 s, Peak search pitch range: 60-330 Hz, Tolerance: 0.05, Interpolation: Parabolic, Tilt line quefreny range 0.001-0.0 s, Line type: Straight, Fit method: robust

Para a obtenção dos valores de AVQI foi adicionado um *script* (versão 03.01) no programa *Praat* e adaptados os ficheiros áudio para que pudesse ser realizada a análise. Foi criado um ficheiro com 3 segundos centrais da vogal sustentada [a] e outro ficheiro com a leitura da primeira frase do texto lida em Int_{subj} média (Maryn & Weenink, 2015).

As informações recolhidas foram guardadas em formato digital no computador portátil da investigadora e os ficheiros protegidos com palavra-passe. Os registos foram arquivados na base de dados do Centro de Investigação em Reabilitação, estando a sua consulta reservada apenas aos membros da equipa de investigação do projeto. A base de dados existente foi criada no âmbito do projeto "Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação", da responsabilidade do Prof. Doutor André Araújo. Os dados serão armazenados durante 6 anos em disco rígido externo, à guarda do investigador principal, na ESS-PP. Após este período, caso não se verifique necessidade de voltar a recolher novos parâmetros acústicos, os ficheiros em formato áudio serão destruídos, mantendo-se apenas as bases de dados estatísticas.

3.4.3. Análise de dados

Foram registados os dados no Excel (Microsoft® Excel® 2016) para organização dos dados e cálculo de algumas medidas (médias, desvio-padrão, mínimo e máximo, índice DSI e medidas do LTAS).

Para a análise estatística dos dados foi realizada a análise descritiva e inferencial, com o programa SPSS (IBM® SPSS® *Statistics*, Versão 26). Para avaliar a normalidade das variáveis em análise foi usado o teste *Shapiro-Wilk*. No caso de não se verificar a normalidade, foi usado o quociente de assimetria (Brooks, 2014 in (Alabdullah, Yahya, & Ramayah, 2014). No caso de ambos os critérios não se verificarem, as correlações não foram consideradas. Depois, foi aplicado o coeficiente de correlação de Pearson de forma a analisar e verificar as relações entre variáveis. Foram consideradas apenas as correlações estatisticamente significativas (significativas a, pelo menos, 0.05).

4. Resultados

No presente capítulo serão expostos os resultados do trabalho. Primeiro, será apresentada a caracterização da amostra em termos de distribuição de género, idade, avaliação por ORL e dados das avaliações perceptiva e acústica da voz. Depois, serão apresentados os resultados relativos às correlações testadas entre variáveis.

4.1. Caracterização da amostra

Os dados demográficos que caracterizam a amostra podem ser observados na Tabela 5.

TABELA 5- CARACTERIZAÇÃO DEMOGRÁFICA DA AMOSTRA

Género	n	Idade (anos)			
		Média	Mínimo	Máximo	Desvio – padrão
Feminino	7	42	20	74	22.2
Masculino	5	60	34	70	15.2

A amostra é constituída por doze indivíduos, cinco do género masculino (M) e sete do género feminino (F). As idades dos indivíduos da amostra estão compreendidas entre os 20 e os 74 anos. A média das idades é 50 anos. No género masculino a média das idades é 60 anos e no género feminino a média das idades é 42 anos. A distribuição da amostra, considerando o diagnóstico ORL e o género, está representada na Tabela 6.

TABELA 6- DISTRIBUIÇÃO DA AMOSTRA CONSIDERANDO DIAGNÓSTICO ORL

Origem da perturbação vocal	n	Género	n	Código
Orgânica	3	F	1	Or5f
		M	2	Or6m; Or7m
Orgânico-funcional	2	F	2	OF8f; OF9f
		M	0	-
Funcional	3	F	2	Fu10f; Fu11f
		M	1	Fu12m
Sem alterações (voz normal)	4	F	2	VN1f; VN2f
		M	2	VN3m; VN4m

Legenda: VN – Voz Normal; Or – Disfonia Orgânica; OF – Disfonia Orgânico Funcional; Fu – Disfonia Funcional; F – Género feminino; M – Género masculino

Considerando o diagnóstico clínico feito pelos médicos otorrinolaringologistas, e segundo a classificação de Pontes, Behlau e Brasil (2000 in Behlau, Azevedo, et al., 2001), oito indivíduos apresentaram alterações sendo três de origem orgânica – Or- (parésia da prega vocal esquerda, parésia da prega vocal direita, e edema de Reinke extenso com degenerescência polipóide na prega vocal direita), dois de origem orgânico-funcional – OF- (Nódulos bilaterais na transição 1/3 anterior -1/3 médio das pregas vocais) e três de origem funcional – Fu- (disfonia tensão muscular, um deles com refluxo faringolaríngeo associado). Quatro indivíduos não apresentaram alterações (VN). Atribuíram-se códigos a cada indivíduo, nos quais contam a identificação do diagnóstico (VN, Or, OF, Fu), número e depois género (f/m).

De forma a avaliar a fiabilidade da avaliação áudio-percetiva das amostras vocais realizadas pelos três peritos avaliadores, calcularam-se os coeficientes de correlação intra (Tabela 7) e inter avaliador (Tabela 8).

TABELA 7- COEFICIENTES DE CORRELAÇÃO *INTRA* AVALIADOR

Peritos	Coeficientes de correlação <i>intra</i> avaliador (<i>r</i> Pearson)			
	Grau disfonia	Grau rouquidão	Grau soproisidade	Grau tensão
Perito 1	0.98	0.93	0.8	0.80
Perito 2	0.90	0.80	0.3	0.96
Perito 3	0.89	0.91	0.8	0.93

A fiabilidade *intra* avaliadores foi medida a partir da aplicação do coeficiente de correlação de Pearson (*r*) entre as amostras de voz repetidas presentes no teste de percepção auditiva. Os avaliadores foram considerados consistentes e fiáveis dado que, nas suas avaliações, apresentaram coeficiente de correlação de Pearson (*r*) superior a 0,8 entre a primeira e segunda avaliação da mesma amostra de voz. Apenas o parâmetro soproisidade para um dos avaliadores apresentou coeficiente de correlação reduzido. (Kreiman et al., 1992).

O cálculo da fiabilidade *inter* avaliadores permitiu determinar a consistência das avaliações feitas pelos diferentes peritos sobre as mesmas amostras vocais. Essa consistência foi calculada através do coeficiente de correlação *intra* classe (ICC) com modelo aleatório a dois fatores, tipo concordância absoluta com intervalo de confiança de 95%. Foi considerado o valor de medidas médias (Koo & Li, 2016)

TABELA 8 - COEFICIENTE DE CORRELAÇÃO INTRACLASSE (ICC)

Coeficiente de correlação intraclass (ICC)				
	Grau Disfonia	Grau Rouquidão	Grau Soproisidade	Grau Tensão
Medidas médias	0.96	0.87	0.89	0.93

O ICC é uma medida de fiabilidade, indicando a semelhança entre as respostas dadas dentro de um determinado conjunto. O valor de ICC varia entre 0 e 1 e, quanto mais próximo de 1 se encontra o resultado, maior a consistência dos resultados (West, Welch & Galecki, 2006). Perante os resultados encontrados no cálculo do ICC, verifica-se que existe elevada fiabilidade *inter* avaliadores (peritos).

Para a avaliação áudio-percetiva dos sujeitos da presente amostra, considerou-se o valor médio das avaliações realizadas pelos peritos. Na Tabela 9 são apresentadas as médias de todos os indivíduos.

TABELA 9 – RESULTADOS DA AVALIAÇÃO ÁUDIO-PERCETIVA (MÉDIA PERITOS)

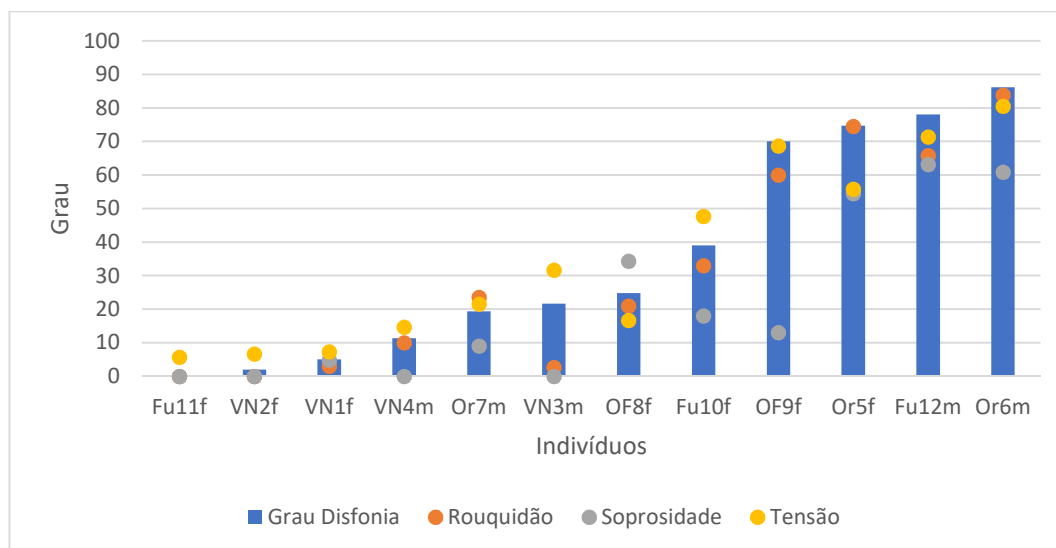
Medida	Média (dp)	Mínimo	Máximo
Grau de disfonia	36 (32)	0	86
Grau de rouquidão	31 (31)	0	84
Grau de soproisidade	21 (25)	0	63
Grau de tensão	36 (28)	6	81

Legenda: dp= desvio-padrão

Verifica-se que os indivíduos da amostra se distribuíram ao longo da escala (0-100), sobretudo para os parâmetros grau de disфонia, rouquidão e tensão.

Na Figura 5 está representada a distribuição do resultado da avaliação áudio-percetiva para cada indivíduo da amostra, agrupados segundo o diagnóstico clínico (VN, Or, OF, Fu) e ordenados segundo o grau de disфонia percebido auditivamente.

FIGURA 5- GRÁFICO DA AVALIAÇÃO ÁUDIO-PERCETIVA



Indivíduos ordenados segundo diagnóstico de disфонia. Legenda: VN - Voz Normal; Or - Disфонia Orgânica; OF - Disфонia Orgânico Funcional; Fu - Disфонia Funcional

É possível observar que, tal como esperado, os maiores graus de alteração vocal correspondem a indivíduos com diagnóstico de disфонia. Também se verifica variedade na distribuição dos parâmetros perceptivos, quer na presença/ausência dos parâmetros, quer na sua frequência. O parâmetro tensão foi aquele que mais vezes foi pontuado/classificado.

Considerando os graus de disфонia de 0 a 100, os indivíduos foram agrupados em quatro níveis: voz sem alterações - SA, desvio ligeiro - DL (11-50), desvio moderado - DM (51-90) e desvio severo - DS (91-100). Os indivíduos distribuíram-se pelos níveis voz normal, desvio ligeiro e desvio moderado, tal como indicado na Tabela 10. Apenas o nível desvio severo não teve nenhum indivíduo.

TABELA 10 - DISTRIBUIÇÃO DA AMOSTRA CONSIDERANDO O GRAU DE DISFONIA

	Voz sem alterações (<11)	Desvio ligeiro (11-50)	Desvio moderado (51-90)	Desvio severo (91-100)
n	3	5	4	0

Em relação ao primeiro objetivo do presente estudo, para as três tarefas de fala encadeada realizadas com Int_{sub} diferentes (média, forte e fraca), foram consideradas as seguintes medidas de longo termo: α_{LTAS} , $(H1-H2)_{LTAS}$, $slope_{LTAS}$, leg , f_0 média, desvio-padrão de f_0 , f_0 mínima e máxima, Gama,

CPP_{fala} e CPPS_{fala}. Para as medidas de f_0 , dada a influência que o gênero pode ter nos resultados, estes são apresentados de forma separada para o gênero feminino e masculino.

Para a tarefa de leitura em *Int_{subj}* média, obtiveram-se os valores apresentados na Tabela 11, agrupados pelos três níveis de grau de disфония (SA, DL e DM) da amostra

TABELA 11- RESULTADOS DAS MEDIDAS ACÚSTICAS NA TAREFA DE LEITURA EM INTENSIDADE SUBJETIVA MÉDIA

	Geral (n=12)	SA (n=3)		DL (n=5)		DM (n=4)	
Medida	Média (dp)	Média	Min-Max	Média	Min-Max	Média	Min-Max
α_{LTAS}	6 (0.3)	6.2	6.1 a 6.2	5.8	5.7 a 6	6.05	5.5 a 6.4
$(H1-H2)_{LTAS}$	-3.5 (3.4)	-1.2	-4.3 a 1	-2.8	-6 a 3.3	-6.1	-7.5 a -3.6
$slope_{LTAS}$	-0.0024 (0.0009)	-0.0025	-0.003 a -0.002	-0.002	-0.004 a -0.0008	-0.003	-0.004 a -0.001
leq (dB)	66 (5)	66	64 a 69	66	62 a 70	65	53 a 72
CPP _{fala} (dB)	36 (2)	34	33 a 35	38	35 a 40	35	34 a 35
CPPS _{fala} (dB)	8 (2)	9.5	9 a 10	9	7 a 10	6	4 a 9
f_0 média (Hz)	F=200 (37) M=159 (69)	F= 216	F=192 a 247	F=199; M=109	F=188 a 209; M=94 a 128	F=176; M=233	F=129 a 222; M=222 a 243
f_0 dp	F=24; M=22	F=27	F=12 a 50	F=22; M=9	F=19 a 25; M=8 a 10	F=20; M=41	F=12 a 27; M=31 a 50

Legenda: SA - Sem Alterações; DL - Desvio Ligeiro; DM - Desvio Moderado; dp - desvio-padrão; F - Gênero feminino; M - Gênero masculino

Os valores médios de $(H1-H2)_{LTAS}$ foram inferiores nos indivíduos do grupo SA. Os resultados médios de leq foram semelhantes para todos os indivíduos, nesta tarefa, independentemente do grau de disфония. Os valores de f_0 média foram superiores nas mulheres SA e inferiores para os homens com DL comparativamente a homens com DM.

Para a tarefa de leitura em *Int_{subj}* fraca, obtiveram-se os valores apresentados na Tabela 12 agrupados pelos três níveis de grau de disфония (SA, DL e DM) da amostra.

TABELA 12– RESULTADOS DAS MEDIDAS ACÚSTICAS NA TAREFA DE LEITURA EM INTENSIDADE SUBJETIVA FRACA

Medida	Geral (n=11)	SA (n=3)		DL (n=5)		DM (n=3)	
	Média (dp)	Média	Min-Max	Média	Min-Max	Média	Min-Max
α_{LTAS}	5.9 (0.3)	5.8	5.6 a 6.1	5.9	5.7 a 6.2	5.9	5.6 a 6.4
$(H1-H2)_{LTAS}$	1 (4.3)	5.3	3.5 a 7.2	0.53	-2.8 a 8.1	-2.4	-3.9 a 0.21
$slope_{LTAS}$	0.57 (0.2)	0.44	0.8 a 0.64	0.57	0.39 a 0.74	0.71	0.63 a 0.82
leq (dB)	57 (3.9)	53	51 a 56	58	55 a 63	58	53 a 61
CPP _{fala} (dB)	36 (2.1)	35.7	34.7 a 37.5	35.6	33.2 a 40.3	37	36 a 38
CPPS _{fala} (dB)	6 (1.4)	5.6	5.1 a 6.6	6	4.9 a 8.5	4	3.2 a 4.7
f_0 média (Hz)	F=187 (29); M=137 (76)	F=194	F=128 a 250	F=193; M=100	F=187 a 200; M=89 a 111	F=167; M=250	F=128 a 205; M=250
f_0 dp	F=20; M=9	F=18	F=16 a 21	F=24; M=6	F=22 a 26; M=5 a 7	F=19; M=17	F=16-21 M=17

Legenda: SA – Sem Alterações; DL – Desvio Ligeiro; DM – Desvio Moderado; dp – desvio-padrão; F – Género feminino; M – Género masculino

Os valores médios da medida $(H1-H2)_{LTAS}$, ao contrário da tarefa de leitura em Int_{subj} média, foram superiores no grupo SA e inferiores nos grupos com alteração vocal. A medida $slope_{LTAS}$ apresentou valores médios mais elevados e todos positivos nesta tarefa e aparentou ter uma tendência crescente com o grau de disфонia. Em relação à medida leq , os valores médios encontrados foram superiores nos grupos DL e DM. De destacar ainda que, para as medidas de f_0 nas mulheres, na tarefa em Int_{subj} fraca, todos os resultados foram inferiores quando comparados com os resultados da tarefa em Int_{subj} média. Os valores de CPP nesta tarefa foram semelhantes nos três grupos e os valores de CPPS foram inferiores no grupo DM.

Para a tarefa leitura em Int_{subj} forte, obtiveram-se os valores apresentados na Tabela 13 agrupados pelos três níveis de grau de disфонia (SA, DL e DM) da amostra.

TABELA 13– RESULTADOS DAS MEDIDAS ACÚSTICAS NA TAREFA DE LEITURA EM INTENSIDADE SUBJETIVA FORTE

	Geral (n=12)	SA (n=3)		DL (n=5)		DM (n=4)	
Medida	Média (dp)	Média	Min-Max	Média	Min-Max	Média	Min-Max
α_{LTAS}	6.21 (0.37)	6.7	6.6 a 6.8	6	5.7 a 6.2	6.13	5.6 a 6.3
$(H1-H2)_{LTAS}$	-2.4 (5.7)	0.5	-3.8 a 7.5	-2.2	-9.41 a 7.64	-4.8	-9 a 0.5
$slope_{LTAS}$	1.24 (0.5)	1.6	0.8 a 2.1	1.3	0.95 a 1.5	0.95	0 a 1.67
leq (dB)	74 (6)	77	77 a 78	76	73 a 80	69	59 a 76
CPP _{fala} (dB)	37 (3.35)	38	35 a 42	38	32 a 41	35	39 a 32
CPPS _{fala} (dB)	10 (2.92)	12	12 a 13	11	10 a 12	6	4 a 9
f_0 média (Hz)	F=227 (44); M=176 (44)	F=253	F=225 a 290	F=218; M=147	F=209 a 228; M=118 a 174	F=196; M=218	F=147 a 244; M=209 a 227
f_0 dp	F=29; M=30	F=34	F=28 a 40	F=25; M=18	F=23 a 27 M=	F=26; M=47	F=15 a 38; M=39 a 55

Legenda: SA – Sem Alterações; DL – Desvio Ligeiro; DM – Desvio Moderado; dp – desvio-padrão; F – Género feminino; M – Género masculino

A medida α_{LTAS} apresentou valores semelhantes para os três grupos. Tal como nas tarefas anteriores, a medida $(H1-H2)_{LTAS}$ também apresentou grande variação de valores. Nesta tarefa, a medida $slope_{LTAS}$ apresentou tendência decrescente com o grau de disфонia e valores mais elevados quando comparada com a tarefa leitura em Int_{subj} média. A medida leq , apresentou valores inferiores nos indivíduos com desvio moderado. Em relação à medida f_0 média, nas mulheres, os valores mais elevados verificaram-se no grupo SA, diminuindo à medida que o grau de disфонia aumenta; nos homens, os valores mais elevados registaram-se no grupo DM. Os valores de CPP foram semelhantes nos três grupos. Os valores de CPPS foram inferiores no grupo DM.

Comparando as tarefas de leitura nas três Int_{subj} diferentes, em termos médios, a medida CPP_{fala} (dB) não apresentou grande variação, ao contrário do CPPS_{fala} que apresentou valores mais elevados na tarefa em Int_{subj} forte e valores inferiores na tarefa em Int_{subj} fraca. Os valores médios de leq , nas três tarefas de fala encadeada, variaram de acordo com a Int_{subj} . Na tarefa de fala em Int_{subj} forte obtiveram-se os valores mais elevados de leq médio, na tarefa em Int_{subj} fraca foram encontrados os valores de leq mais reduzidos e, finalmente, na tarefa em Int_{subj} média, obtiveram-se valores de leq intermédios. Os valores médios de f_0 apresentam tendência decrescente com o aumento do grau de disфонia.

Em relação à dinâmica de intensidades, calculada pela diferença entre o leq nas tarefas de Int_{subj} fraca e forte, os valores obtidos encontram-se na Tabela 14 agrupados pelos três níveis de grau de disфонia (SA, DL e DM) da amostra.

TABELA 14- RESULTADOS DA DINÂMICA DE INTENSIDADES

	Geral (n=12)	SA (n=3)		DL (n=5)		DM (n=4)	
Medida	Média (dp)	Média	Min-Max	Média	Min-Max	Média	Min-Max
Dinâmica intensidades (dB)	18 (6)	24	21 a 28	18	12 a 23	13	6 a 16

Legenda: SA – Sem Alterações; DL - Desvio Ligeiro; DM - Desvio Moderado; dp – desvio-padrão

O valor médio da amostra foi de 18dB. O valor da dinâmica de intensidades foi superior no grupo SA e inferior no grupo DM. Foi feita observação e análise qualitativa dos gráficos. A Figura 6 diz respeito à representação gráfica do maior valor de dinâmica de intensidades da amostra, 28 dB, correspondendo ao indivíduo VN2f, com grau de disфонia igual a 2 (SA).

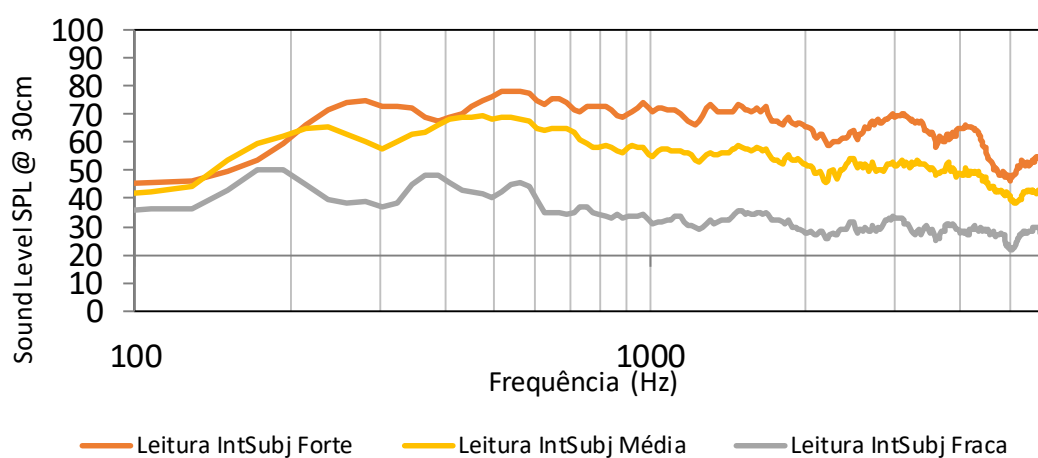


FIGURA 6 - GRÁFICO LTAS DO INDIVÍDUO VN2f

Verificou-se que a um maior valor na dinâmica de intensidades, correspondeu um maior espaçamento entre as linhas das intensidades médias das leituras em Int_{subj} forte e fraca.

A Figura 7 corresponde à representação do menor valor na dinâmica de intensidades da amostra, 15dB, correspondente ao indivíduo OF9f, com grau de disфонia de 70 (DM).

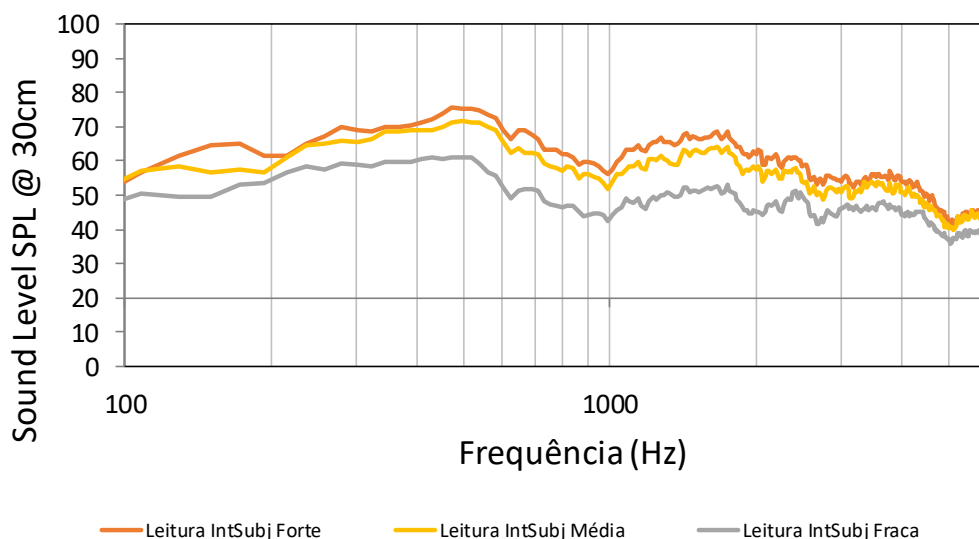


FIGURA 7 - GRÁFICO LTAS DO INDIVÍDUO OF9f

É possível observar que a um menor espaçamento entre as linhas das intensidades média, fraca e forte correspondeu um menor valor na dinâmica de intensidades.

Considerando o segundo objetivo do estudo, referente à relação entre as medidas acústicas de longo e curto termo, são apresentados na Tabela 15 os valores obtidos na análise acústica da vogal [a] sustentada. Os valores estão agrupados pelos três níveis de grau de disфонia (SA, DL e DM) da amostra.

TABELA 15- RESULTADOS DAS MEDIDAS ACÚSTICAS NA TAREFA DE VOGAL [a] SUSTENTADA

	Geral (n=12)	SA (n=3)	DL (n=5)	DM (n=4)			
Medida	Média (dp)	Média	Min-Max	Média	Min-Max	Média	Min-Max
<i>Jitter</i> local (%)	1 (0.95)	0.5	0.4 a 0.5	1.3	0.3 a 3.5	1.4	0.8 a 1.9
<i>Shimmer</i> local (%)	9 (4.8)	5	4 a 8	8	3 a 14	13	9.5 a 16
<i>HNR</i> (dB)	11 (5.8)	17	14 a 19	12	5 a 18	6	3 a 8
Intensidade média (dB)	54 (7)	54	48 a 62	57	50 a 65	51	42 a 63
CPP _{vogal} (dB)	38 (4)	42	40 a 45	38	35 a 41	34	32 a 36
CPPS _{vogal} (dB)	13 (4.7)	14	11 a 18	17	10 a 19	8	6 a 10
f ₀ média (Hz)	F=178 (53); M=105 (30)	F=209	F=193 a 236	F=206; M=111	F=206 a 207 M=98 a 122	F=104 M=149	F=93 a 116 M=121 a 176
f ₀ desvio-padrão	F=7; M=2	F=1.4	F=1 a 2	F=3; M=1	F=3; M=1	F=18; M=4	F= 6 a 31; M=4

Legenda:SA–Sem Alterações; DL– Desvio Ligeiro; DM–Desvio Moderado; dp–desvio-padrão; F – Género feminino; M – Género masculino

Salienta-se deste conjunto de resultados os seguintes: a medida *shimmer* apresentou valores médios mais elevados no grupo com maior grau de disфонia; a medida *HNR* apresentou valores médios mais elevados no grupo SA, diminuindo com o aumento do grau de disфонia; o desvio-padrão de f₀ média foi elevado, sobretudo no género feminino e no grupo com maior grau de disфонia; os valores de CPP_{vogal} foram diminuindo à medida que o grau de disфонia aumenta. Verifica-se também que os valores de CPP_{vogal} são semelhantes aos valores de CPP_{fala} encontrados nas diferentes Int_{subj}. Os valores de CPPS_{vogal}

foram superiores aos de $CPPS_{fala}$, apresentando quebras mais acentuadas para os indivíduos com DM do que nas tarefas de fala encadeada.

Para o terceiro objetivo deste estudo foi necessário calcular os índices AVQI e DSI. Na Figura 8 está o resultado do cálculo do DSI para o indivíduo VN3m, com grau de disfonia 22 (DL), calculado no *Praat* utilizando o *script* de Maryn et al. (2017).

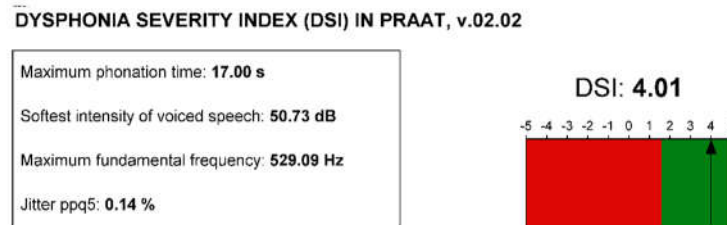


FIGURA 8 - RESULTADO DO DSI NO PRAAT - INDIVÍDUO VN3m

À esquerda é apresentado um quadro com as medidas que compõem o DSI e à direita o valor final do DSI com uma representação gráfica que indica se o valor obtido se encontra dentro ou fora da norma.

Para o índice AVQI, a informação dada pelo programa *Praat* corresponde ao valor final do índice, aos valores individuais das seis medidas que compõem este índice e às representações gráficas dos espectros e cepstra. Nas figuras seguintes são apresentados dois exemplos que representam os valores mínimo e máximo do índice AVQI obtidos neste estudo. Na Figura 9 é apresentado o menor resultado do AVQI cujo valor é 0.99, correspondente ao indivíduo VN3m com um grau de disfonia 22 (DL).

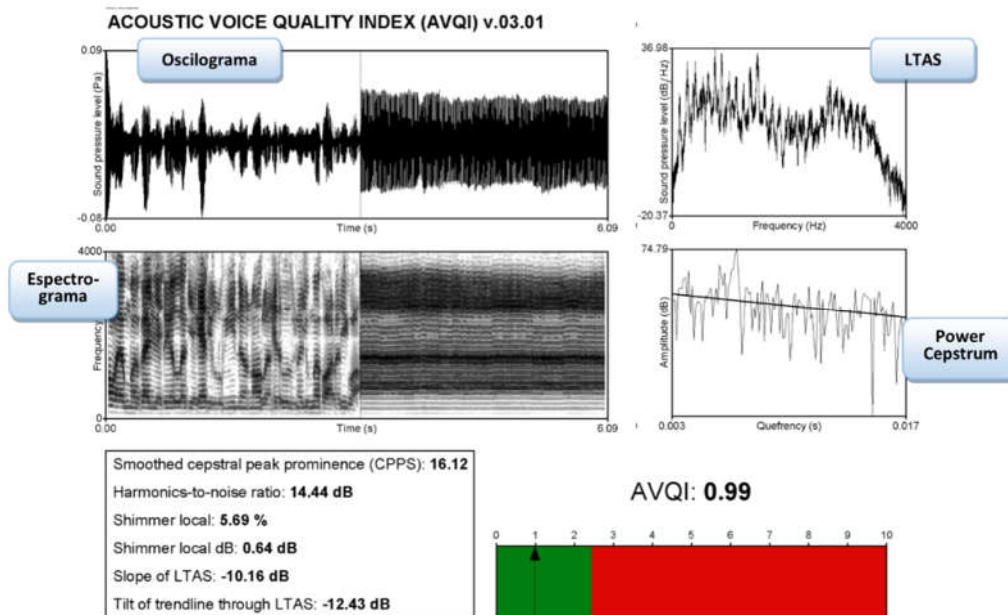


FIGURA 9 - ÍNDICE AVQI INDIVÍDUO VN3m

Na Figura 10 é apresentado o maior resultado do AVQI com o valor é 8.83, correspondente ao indivíduo Fu12m com um grau de disfonia 78 (DM).

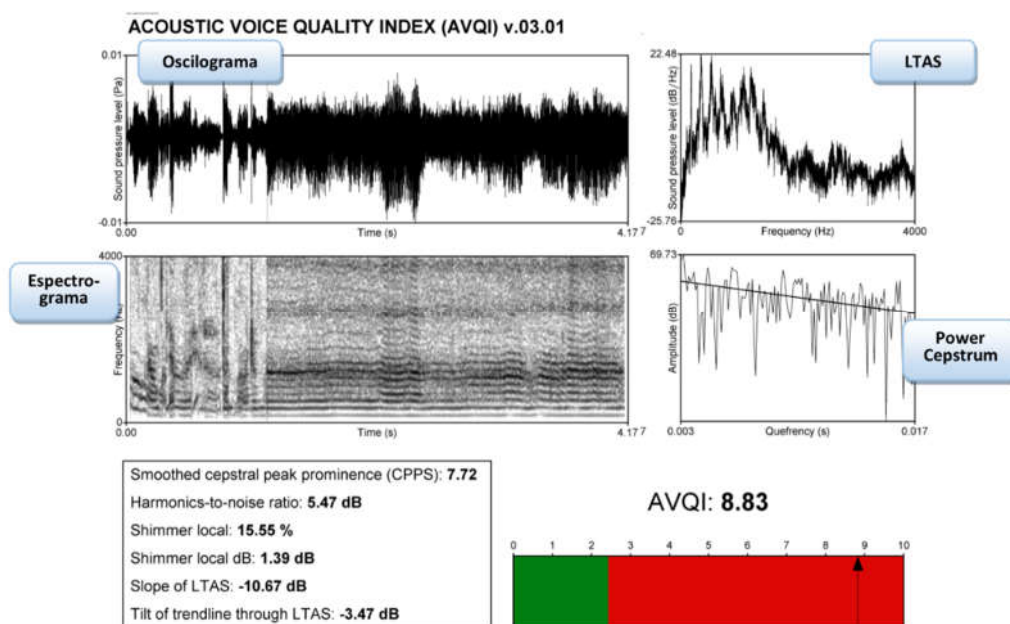


FIGURA 10 - ÍNDICE AVQI INDIVÍDUO FU12m

Os resultados médios dos valores dos índices AVQI e DSI obtidos são apresentados na Tabela 16 agrupados pelos três níveis de grau de disfonia (SA, DL e DM) da amostra.

TABELA 16- RESULTADOS DOS ÍNDICES DSI E AVQI

	Geral (n=12)		SA (n=3)		DL (n=5)		DM (n=4)	
Medida	Média	DP	Média	Min-Max	Média	Min-Max	Média	Min-Max
AVQI	4	2.6	2.7	2 a 3.5	2.7	0.9 a 5.6	7	5.2 a 8.8
DSI	2	1.5	3.7	3 a 4	3	0.7 a 4	0.8	0.3 a 1.4

O valor médio do índice DSI foi menor no grupo DM. O valor médio do índice AVQI, apesar de igual para os indivíduos SA e DL, apresentou maior variação de valores no grupo DL; além disso, o valor médio do índice AVQI foi superior nos indivíduos DM. Na Figura 11 são apresentados os valores de DSI e AVQI, com os indivíduos ordenados pelo grau crescente de disfonia.

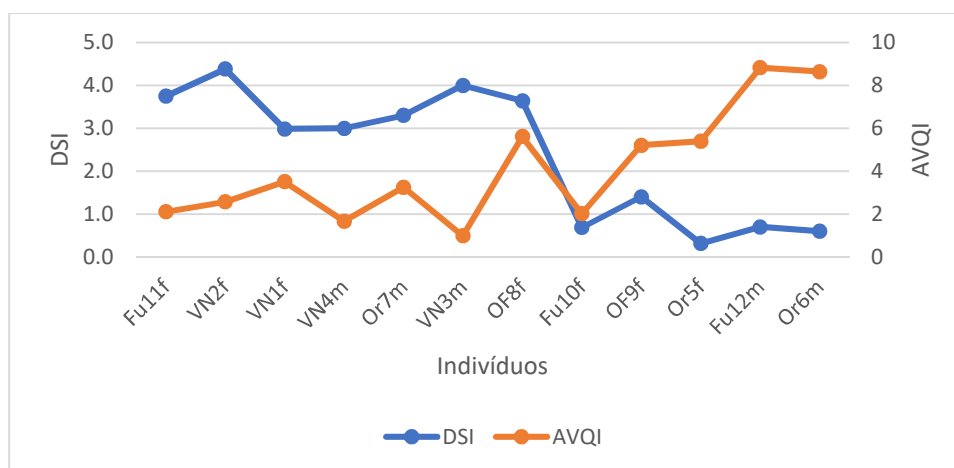


FIGURA 11- RESULTADOS DOS ÍNDICES DSI E AVQI

Legenda: VN - Voz Normal; Or - Disfonia Orgânica; OF - Disfonia Orgânico Funcional; Fu - Disfonia Funcional

A escala dos dois índices é diferente, sendo que para o DSI o limite superior é 5 (escala à esquerda) e para o AVQI é 10 (escala à direita). É possível observar que estes índices variam de forma inversa: a valores elevados no índice AVQI correspondem valores menores no índice DSI e a valores elevados do DSI correspondem valores inferiores no índice AVQI. Quanto maior o grau de disfonia, maior o valor do índice AVQI e menor o valor do índice DSI.

4.2. Relações entre variáveis

Para o objetivo 1 - investigar as correlações entre medidas acústicas de longo termo avaliadas em tarefa de fala encadeada em Int_{subj} média, forte e fraca, e a avaliação áudio-perceptiva - apresentam-se apenas as correlações significativas na Tabela 17.

Foi encontrado maior número de correlações significativas entre a avaliação áudio-perceptiva e medidas acústicas das tarefas de leitura em Int_{subj} forte e fraca.

Para a medida $CPPS_{fala}$ obtiveram-se correlações significativas com todos os parâmetros perceptivos. Nas três tarefas de leitura, o $CPPS_{fala}$ correlacionou-se inversamente com o grau de disfonia; na tarefa de leitura em Int_{subj} forte, correlacionou-se inversamente com os parâmetros rouquidão, tensão e sopro; na tarefa de leitura fraca, correlacionou-se com o parâmetro rouquidão.

Nas tarefas de em Int_{subj} média e fraca foi encontrada correlação inversa entre o Grau de disfonia e tensão e a medida $(H1-H2)_{LTAS}$, ou seja, quanto maior o grau de disfonia e tensão, menor $(H1-H2)_{LTAS}$. A medida α_{LTAS} foi a única medida do LTAS que não obteve correlação significativa com medidas perceptivas.

TABELA 17 – COEFICIENTE DE CORRELAÇÃO DE PEARSON – MEDIDAS ACÚSTICAS DE LONGO TERMO E AVALIAÇÃO ÁUDIO-PERCETIVA

Coeficiente de correlação de Pearson (valor-p)	Tarefa		
	Leitura Int _{subj} Média	Leitura Int _{subj} Forte	Leitura Int _{subj} Fraca
<i>alpha</i> _{LTAS} e Grau Disfonia	0.143 (0.657)	-0.306 (0.333)	0.106 (0.757)
(H1-H2) _{LTAS} e Grau Disfonia	-0.594 (0.042)	-0.319 (0.312)	-0.634 (0.036)
(H1-H2) _{LTAS} e Grau Tensão	-0.677 (0.016)	-0.328 (0.297)	-0.709 (0.015)
<i>slope</i> _{LTAS} e Grau Disfonia	-0.148 (0.647)	-0.403 (0.194)	0.653 (0.029)
<i>slope</i> _{LTAS} e Grau Rouquidão	-0.104 (0.748)	-0.363 (0.246)	0.659 (0.028)
<i>leq</i> e Grau Soprosidade	-0.077 (0.812)	*	0.641 (0.034)
CPPS _{fala} e Grau Disfonia	-0.728 (0.007)	-0.898 (0.000)	-0.624 (0.040)
CPPS _{fala} e Grau Rouquidão	-0.735 (0.006)	-0.893 (0.000)	-0.722 (0.012)
CPPS _{fala} e Grau Soprosidade	-0.758 (0.004)	-0.870 (0.000)	-0.562 (0.070)
CPPS _{fala} e Grau Tensão	-0.714 (0.009)	-0.890 (0.000)	-0.562 (0.072)
<i>f</i> ₀ média e Grau Soprosidade	0.550 (0.064)	0.249 (0.436)	0.612 (0.046)
<i>f</i> ₀ mín e Grau Soprosidade	0.535 (0.073)	0.429 (0.164)	0.620 (0.042)
dp <i>f</i> ₀ e Grau Soprosidade	0.659 (0.020)	0.647 (0.023)	0.400 (0.223)
Gama e Grau Soprosidade	0.485 (0.110)	0.623 (0.030)	0.130 (0.703)
Int mín e Grau Soprosidade	0.115 (0.723)	-0.547 (0.066)	0.721 (0.012)
Dinâmica de intensidades e Grau Disfonia	n.a.	-0.789 (0.002)	
Dinâmica de intensidades e Grau Rouquidão	n.a.	-0.774 (0.003)	
Dinâmica de intensidades e Soprosidade	n.a.	-0.763 (0.004)	
Dinâmica de intensidades e Tensão	n.a.	-0.780 (0.003)	

Legenda: * *leq* não seguiu distribuição normal para leitura Int_{subj} forte; n.a. – não aplicável

Nota: na tabela são assinaladas a cor diferente as correlações estatisticamente significativas (significativas a, pelo menos, 0.05): a verde as correlações diretas e a vermelho as correlações inversas.

O parâmetro soprosidade foi o parâmetro da avaliação áudio-percetiva que obteve maior número de correlações significativas com medidas acústicas de longo termo. Estas correlações verificaram-se nas tarefas de leitura em Int_{subj} forte e fraca. Foram encontradas correlações inversas entre a dinâmica de intensidades e todos os parâmetros percecionados auditivamente: grau de disfonia, rouquidão, soprosidade e tensão. Estes dados significam que os valores de dinâmica de intensidades mais elevados correspondem a indivíduos com menor grau de alteração vocal percecionado auditivamente e, no sentido inverso, que os menores valores de dinâmica de intensidades correspondem aos indivíduos com maior grau de alteração vocal.

Para o segundo objetivo, em que se pretendia investigar as correlações entre medidas acústicas de longo termo (obtidas na tarefa de Leitura em Int_{subj} Média) e medidas de curto termo (obtidas a partir da vogal [a] sustentada), obtiveram-se as correlações significativas apresentadas na Tabela 18:

TABELA 18– COEFICIENTE DE CORRELAÇÃO DE PEARSON – MEDIDAS ACÚSTICAS DE LONGO E CURTO TERMO

Coeficiente de correlação de Pearson(valor-p)	Medidas vogal [a]						
	Medidas leitura Int _{subj} média	<i>shimmer</i> local	<i>HNR</i>	<i>f</i> ₀ média	Intensidade média	dp Intensidade	CPP _{vogal}
<i>alfa</i> _{LTAS}	0.275 (0.387)	-0.142 (0.660)	0.455 (0.138)	0.029 (0.929)	-0.398 (0.200)	-0.115 (0.722)	-0.371 (0.234)
(H1-H2) _{LTAS}	-0.139 (0.667)	0.240 (0.452)	0.666 (0.018)	-0.197 (0.540)	-0.556 (0.060)	0.229 (0.474)	0.090 (0.781)
<i>slope</i> _{LTAS}	0.184 (0.568)	-0.121 (0.708)	0.339 (0.281)	-0.398 (0.200)	-0.071 (0.826)	-0.201 (0.531)	0.019 (0.954)
<i>leq</i> _{LTAS}	-0.229 (0.475)	0.078 (0.810)	-0.121 (0.707)	0.716 (0.009)	-0.106 (0.742)	0.233 (0.466)	0.092 (0.777)
<i>f</i> ₀ média	0.453 (0.140)	-0.370 (0.237)	0.562 (0.057)	-0.118 (0.715)	-0.242 (0.448)	-0.286 (0.367)	-0.666 (0.018)
desvio-padrão de <i>f</i> ₀	0.523 (0.081)	-0.383 (0.219)	0.407 (0.189)	-0.334 (0.289)	-0.178 (0.580)	-0.389 (0.211)	-0.639 (0.025)
<i>f</i> ₀ mínima	0.394 (0.205)	-0.311 (0.326)	0.544 (0.068)	-0.261 (0.412)	-0.240 (0.453)	-0.287 (0.365)	-0.698 (0.012)
<i>f</i> ₀ máxima	0.430 (0.163)	-0.318 (0.314)	0.602 (0.038)	-0.239 (0.453)	-0.258 (0.418)	-0.249 (0.436)	-0.659 (0.020)
Gama	0.441 (0.151)	-0.308 (0.330)	0.627 (0.029)	-0.206 (0.522)	-0.262 (0.411)	-0.197 (0.538)	-0.585 (0.046)
CPP _{fala}	-0.222 (0.488)	0.155 (0.631)	-0.206 (0.521)	0.151 (0.640)	-0.150 (0.641)	-0.018 (0.957)	0.446 (0.146)
CPPS _{fala}	-0.761 (0.004)	0.689 (0.013)	0.207 (0.518)	0.285 (0.369)	-0.426 (0.168)	0.784 (0.003)	0.683 (0.014)
Dinâmica de Intensidades	-0.864 (0.000)	0.841 (0.001)	0.168 (0.602)	0.179 (0.577)	-0.438 (0.154)	0.846 (0.001)	0.423 (0.170)

Legenda: dp – desvio-padrão

Nota: na tabela são assinaladas a cor diferente as correlações estatisticamente significativas (significativas a, pelo menos, 0.05): a verde, os coeficientes de correlação positivos (indicando correlação direta) e, a vermelho, os coeficientes de correlação negativos (indicando correlação inversa).

Todas as variáveis seguem distribuição normal à exceção do *jitter* e desvio-padrão *f*₀, e, por isso, estas não são apresentadas na tabela. As medidas CPPS_{fala} e CPPS_{vogal} apresentaram o maior número de correlações significativas. O CPPS_{fala} correlacionou-se com as medidas *HNR*, CPP_{vogal}, e CPPS_{vogal} e inversamente com o *shimmer*. O CPPS_{vogal} correlacionou-se significativamente com todas as medidas de longo termo relacionadas com a *f*₀ (média, dp, máximo, mínimo e gama). Também foram encontradas correlações significativas entre a dinâmica de intensidades e medidas acústicas de curto termo. A dinâmica de intensidades correlacionou-se com as medidas de curto termo *HNR* e CPP_{vogal} e correlacionou-se inversamente com a medida *shimmer*.

Apesar de não se ter investigado as correlações entre medidas de curto termo e grau de disфония, pela observação dos resultados obtidos nas várias medidas de curto termo e considerando o grau de disфония, verificou-se que os indivíduos da amostra com maior grau de disфония obtiveram valores superiores nas medidas *jitter*, *shimmer* e desvio-padrão de f_0 . Por outro lado, verificou-se que os indivíduos da amostra com maior grau de disфония apresentam valores inferiores nas medidas CPP_{vogal} , $CPPS_{vogal}$ e *HNR* quando comparados com os indivíduos com menor grau de disфония.

Para o terceiro objetivo, em que se pretendia investigar as correlações entre os índices AVQI e DSI e diferentes graus de disфония, o coeficiente de correlação encontrado é positivo entre o grau de disфония e o índice AVQI, indicando uma correlação direta, e negativo entre o grau de disфония e o índice DSI, indicando uma correlação inversa, tal como apresentado na Tabela 19:

TABELA 19– COEFICIENTE DE CORRELAÇÃO DE PEARSON – GRAU DE DISFONIA E ÍNDICES AVQI E DSI

Medidas	Coefficiente de correlação de Pearson (valor-p)
Grau de disфония – AVQI	0.804(0.002)
Grau de Disфония – DSI	-0.883(< 0.001)

Nota: na tabela é apresentado a verde o coeficiente de correlação positivo (indicando correlação direta) e a vermelho o coeficiente de correlação negativo (indicando correlação inversa).

Numa análise exploratória dos dados, foram encontradas outras correlações significativas com alguma pertinência. Foi encontrada correlação entre a dinâmica de intensidades e o índice DSI, indicando que quanto maior a dinâmica de intensidades, maior o valor de DSI. Foi também descoberta correlação inversa entre a dinâmica de intensidades e o índice AVQI, indicando que quanto maior a dinâmica de intensidades, menor o valor de AVQI. Estando os índices DSI e AVQI relacionados com grau de disфония, e tendo sido estabelecida correlação inversa entre a dinâmica de intensidades e o grau de disфония, é natural que estas medidas se relacionem. As medidas CPP_{vogal} e $CPPS_{vogal}$ correlacionaram-se de forma inversa com todos os parâmetros perceptivos.

5. Discussão

A avaliação da qualidade vocal deve ser multidimensional, atendendo à sua natureza também multidimensional. A avaliação acústica, e, em particular, a avaliação através de medidas de longo termo, é apenas uma parte da avaliação da qualidade vocal. Servirá para complementar a restante avaliação acústica, a avaliação aerodinâmica, a avaliação áudio-perceptiva, a avaliação por ORL (incluindo imagens endoscópicas da laringe) e a autoavaliação do impacto da perturbação vocal (Patel et al., 2018). As medidas acústicas, ainda que objetivas, não anulam as medidas áudio-perceptivas. Estas caracterizam a qualidade vocal com maior número de atributos e apreciações possíveis pelo ouvido humano (Awan & Lawson, 2009; Eskenazi, Childers, & Hicks, 1990)

Na avaliação acústica da voz, de natureza quantitativa e, por isso, considerada objetiva, deverão ser incluídas várias medidas, dado que não há uma medida que represente todas as variantes e os vários desvios possíveis, tanto da voz normal, como perturbada. Além disso, a análise acústica deverá ser baseada em tipos de medidas diferentes (por exemplo, medidas de perturbação e medidas espectrais), dada a natureza dos parâmetros de qualidade vocal ser diversa (Wolfe & Martin, 1997).

Na avaliação acústica da voz deverão contemplar-se diferentes tarefas, porque destas depende a correta análise da voz em todos os seus domínios. É habitual a utilização das vogais sustentadas na avaliação da qualidade vocal pela simplicidade na sua realização e pela facilidade na extração das medidas. No entanto, nestas tarefas não estão presentes os aspetos próprios da dinâmica da fala encadeada (Hammarberg et al., 1981). Por outro lado, as medidas de longo termo, nas quais se incluem medidas do LTAS e medidas cepstrais, permitem a análise acústica da fala encadeada, mais representativa daquilo que a voz naturalmente é (Lowell et al., 2011). De uma forma geral, estas medidas são independentes do material de fala usado, sendo as medidas CPP e CPPS apontadas como promissoras (Maryn et al., 2009).

Tarefas de emissão de vogais sustentadas e fala encadeada em Int_{subj} média são habituais na avaliação clínica da voz (Barsties & De Bodt, 2015; Dejonckere et al., 2001; Oates, 2009; Patel et al., 2018). No entanto, nas perturbações da voz, os limites da emissão vocal (sejam de altura tonal ou intensidade) estão diminuídos, correspondendo a queixas frequentes nos indivíduos com disфонia (Behlau, Madazio, Feijó, et al., 2001). Surge daí a importância de avaliar a qualidade vocal em Int_{subj} forte e fraca, que foi explorada neste trabalho.

Têm surgido índices que combinam várias medidas acústicas e que são utilizados na investigação. Trata-se de formas de combinar diferentes aspetos da voz, no sentido de corresponder à multidimensionalidade da qualidade vocal. Neste estudo foram considerados os índices DSI e AVQI.

Dos resultados encontrados para a correlação entre variáveis, são vários os concordantes com a literatura. De salientar, em primeiro lugar, a variação de algumas das medidas acústicas com o grau de alteração vocal (grau de disфонia) percecionado auditivamente.

A medida CPPS, considerada como uma medida acústica bastante robusta do grau global de disфонia (Patel et al., 2018), e correlacionada de forma inversa com o grau de alteração vocal, apresentou resultados semelhantes na presente investigação. O CPPS_{fala} correlacionou-se de forma inversa com o grau de disфонia: quanto maior o grau de alteração vocal, menor o valor de CPPS_{fala}, tal como referida na literatura (Lowell et al., 2012). Esta mesma correlação verificou-se para as leituras realizadas em Int_{subj} média, forte e fraca. Também as medidas CPP_{vogal} e CPPS_{vogal} se correlacionaram inversamente com o grau de disфонia. As vozes com melhor qualidade vocal e, portanto, mais periódicas, apresentam uma configuração harmónica bem definida dentro do espectro e, desta forma, o pico cepstral é mais proeminente, a que correspondem valores mais elevados de CPP e CPPS (Delgado-Hernández et al.,

2019). Sabendo que estas medidas traduzem o grau de harmonia e periodicidade do sinal acústico, é esperado que, quanto maior o seu valor, melhor a qualidade vocal e, portanto, menor o grau de alteração vocal. As medidas cepstrais utilizadas neste trabalho revelaram-se de fácil obtenção, dado que o seu cálculo já é possível no programa *Praat* (Phadke et al., 2018). O CPPS obtido no *Praat* em tarefas de fala encadeada teve uma grande taxa de precisão na previsão da presença de perturbação vocal com excelente valor de sensibilidade de 90% (Sauder et al., 2017).

No entanto, para o CPP e CPPS, existiu dificuldade em encontrar valores de referência comparáveis, obtidos no mesmo programa e usando as mesmas especificações. Este facto demonstra a necessidade de valores de referência para as medidas acústicas, que devem ser específicos de diferentes tarefas (vogais, fala encadeada em diferentes intensidades, canto), diferentes programas e diferentes populações. Assim, esta é uma das áreas a que importa dar continuidade em investigações futuras, tanto em Portugal como em outros países.

Os valores de CPPS_{fala} são, em média, maiores na leitura em Int_{subj} forte. O estudo de Brockmann-Bauser et al. (2019) chegou a um resultado semelhante para a medida CPPS_{vogal}. Nesse estudo, verificou-se aumento do CPPS_{vogal} com o aumento da intensidade da voz em mulheres com e sem perturbação vocal (de causa hiperfuncional). Segundo os autores, a relação direta entre CPPS e intensidade pode ser uma consequência da fonte harmónica mais forte com o aumento da intensidade de voz (Brockmann-Bauser et al., 2019). Phadke et al. (2018) também encontraram correlação moderada entre SPL e as medidas CPP_{fala} e CPPS_{fala} em tarefa em Int_{subj} forte, o que já não se verificou na tarefa em Int_{subj} média. Os autores explicam que a relação encontrada pode ter a ver com o facto de o aumento da intensidade vocal diminuir o grau de alteração vocal. Isto acontece porque o acréscimo da compressão medial das pregas vocais aumenta, melhorando o encerramento glótico, o que leva à diminuição do ruído glótico e consequente aumento dos valores de CPP e CPPS.

Na tarefa de leitura em Int_{subj} forte do presente estudo, a medida CPPS_{fala} correlacionou-se inversamente com os parâmetros rouquidão, sopro e tensão. Considerando que estas correlações se verificaram na tarefa Int_{subj} forte, seria de esperar que, tal como descrito anteriormente, o aumento da intensidade fizesse diminuir o grau de alteração vocal. Na literatura estão descritas relações entre o CPPS, a sopro e a tensão (para além do grau de disфонia), em tarefas em Int_{subj} média. Sendo o CPPS uma medida relacionada com a qualidade vocal, parece natural que surjam outras associações, tal como a rouquidão (Wolfe & Martin, 1997).

Os valores de CPP_{vogal} e CPPS_{vogal} são superiores aos obtidos nas tarefas de leitura, tal como era esperado. Isto será explicado pelo facto de a vogal sustentada ser mais estável durante a fonação, não apresentando variações próprias da fala encadeada, tais como transições entre vogais-consoantes, variações de frequência e intensidade, pausas, entre outras. Estas variações na fala encadeada diminuem a proeminência dos harmónicos sobre o ruído, medido pelo CPP e CPPS (Núñez-Batalla et al.,

2019; P S & Pebbili, 2020; Phadke et al., 2018). Tal como no estudo de Delgado-Hernández et al. (2019), comparando o CPPS_{fala} e o CPSS_{vogal}, verificou-se que o CPPS_{fala} revelou maior poder discriminatório entre voz normal e voz alterada.

Relativamente a medidas de intensidade de longo termo, verificaram-se diferenças nos valores de *leq* médio nas tarefas em Int_{subj} forte e fraca considerando o grau de disфонia. Apesar de não ter sido encontrada correlação significativa entre *leq* médio e o grau de disфонia, pela observação dos resultados foi possível verificar diferenças nos valores médios de *leq* entre os grupos naquelas duas tarefas. Os indivíduos com maior grau de disфонia apresentaram valores mais elevados de *leq* na tarefa em Int_{subj} fraca e valores inferiores na tarefa em Int_{subj} forte, quando comparados com os indivíduos com menor grau de disфонia. Estes dados poderão revelar dificuldades em produzir voz em Int_{subj} fraca e forte. Estes são resultados esperados uma vez que nas perturbações da voz é habitual haver dificuldade nos extremos da emissão vocal, ou seja, a diminuição da dinâmica ao nível da intensidade da voz é comum (Behlau, Madazio, et al., 2001). São frequentes as queixas relacionadas com dificuldade na projeção vocal ou na emissão vocal em Int_{subj} fraca, mantendo a qualidade vocal ou sem realização de esforço vocal excessivo (Behlau, Madazio, et al., 2001). No presente estudo, os valores médios de *leq* foram semelhantes para a tarefa de fala em Int_{subj} média para todos os indivíduos, independentemente do seu grau de disфонia. Estes resultados reforçam a necessidade de a avaliação vocal incluir tarefas em diferentes intensidades, registando os valores de intensidade que o indivíduo é capaz de produzir ao longo do processo de reabilitação vocal. Esta é também uma prática que deverá ser promovida e integrada pelos terapeutas da fala na avaliação da voz.

Além disso, neste estudo, os resultados encontrados para a dinâmica de intensidades indicam relações desta medida com a qualidade vocal. Verificaram-se correlações significativas com o grau de disфонia percebido auditivamente e com medidas acústicas habitualmente relacionadas com a qualidade vocal. Concretizando, correlacionou-se de forma inversa com o grau de disфонia, com o índice AVQI, e ainda com o *shimmer*. O *shimmer* mede a perturbação da amplitude ciclo a ciclo e sofre influência da amplitude de vibração das pregas vocais e da pressão subglótica. Desta forma, sofre efeito do suporte respiratório inadequado ou do fechamento glótico incompleto.

A dinâmica de intensidades correlacionou-se ainda com o DSI, CPPS_{fala}, CPP_{vogal} e com o *HNR*, indicando que, quanto maior a dinâmica de intensidades da voz, maiores serão os valores de DSI, CPPS_{fala}, CPP_{vogal} e de *HNR*. A forma como se correlacionou com as medidas acústicas traduz como esta medida pode contribuir para a aferição da qualidade da voz na patologia vocal. Em estudos futuros seria interessante investigar se esta medida é sensível às mudanças de qualidade vocal que se verificam em processos de treino ou reabilitação vocal. Esta medida, mais simples de obter do que o fonetograma (*voice range profile*), pode ser um indicador relevante para os terapeutas da fala incluírem de forma generalizada nos seus processos de demonstração de resultados clínicos.

Nas tarefas em Int_{subj} média e fraca foi encontrada correlação inversa entre o grau de disфонia e tensão e a medida $(H1-H2)_{LTAS}$, ou seja, quanto maior o grau de disфонia e tensão, menor $(H1-H2)_{LTAS}$.

A medida $(H1-H2)_{LTAS}$ está relacionada com a qualidade vocal: um valor de H1 superior a H2 está relacionado com vozes hipofuncionais e, por outro lado, um valor de H1 inferior a H2 relaciona-se com vozes hiperfuncionais (da Silva et al., 2011). No entanto, o grau de disфонia pode ser elevado quer em casos de vozes hipofuncionais, quer em casos de vozes hiperfuncionais (estes dois aspetos podem representar alteração da qualidade vocal e, conseqüentemente, presença de disфонia). Por isso foi relevante constatar que, dos parâmetros perceptivos (da disфонia), a tensão foi aquele que se relacionou com a medida $(H1-H2)_{LTAS}$. Tal poderá significar que vozes com valores de $(H1-H2)_{LTAS}$ positivos corresponderão a vozes percecionadas como menos tensas e vozes com valores de $(H1-H2)_{LTAS}$ negativos corresponderão a vozes percecionadas como mais tensas. Este resultado corroborou a relevância e potencial desta medida de LTAS na avaliação da voz, especialmente se se pretender explorar o parâmetro tensão da qualidade vocal.

Neste estudo, a medida $alfa_{LTAS}$ não se correlacionou significativamente com nenhum parâmetro vocal perceptivo. Isto não exclui a sua relevância, uma vez que pode ser útil para representar fenómenos de ressonância (como a formante do cantor) ou desvios como ressonância cul-de-sac ou faríngea (Sundberg, 1987 in Master et al., 2006). Contudo, os parâmetros perceptivos neste estudo limitaram-se à descrição da qualidade de fonte glótica. Assim, será relevante voltar a estudar esta medida com outros critérios de avaliação perceptiva.

A medida $slope_{LTAS}$ correlacionou-se com o grau de disфонia e rouquidão na tarefa de leitura em Int_{subj} fraca. Esta medida refere-se à inclinação espectral e está relacionada com a velocidade da fase de encerramento glótico. A inclinação ($slope_{LTAS}$) será maior se aquela velocidade for lenta, e por sua vez, indicativo de uma pior adução ou pior mobilidade das pregas vocais, o que é frequente nas vozes com maior grau de disфонia e rouquidão (da Silva et al., 2011; Leino, 2009).

O parâmetro soprosidade foi o parâmetro da avaliação áudio-perceptiva com maior número de correlações significativas com medidas acústicas de longo termo, nomeadamente leg , f_0 média, f_0 mínima, f_0 dp, gama tonal, intensidade mínima. Também se relacionou com o CPPS de forma inversa, na leitura em Int_{subj} forte, como já foi explicado anteriormente. A correlação encontrada entre soprosidade e as medidas leg e intensidade da leitura em Int_{subj} fraca poderá ser justificada pela natural redução da tensão e aumento da fenda glótica quando em situação de Int_{subj} mais fraca.

Foi encontrado maior número de correlações significativas entre a avaliação áudio-perceptiva e medidas acústicas das tarefas de leitura forte e fraca. Isto poderá indicar e justificar a necessidade de a avaliação vocal contemplar tarefas em diferentes Int_{subj} . Diferentes tarefas vocais podem evidenciar características fisiológicas ou patológicas específicas (Watts, Ronshaugen & Saenz, 2015 in Brockmann-Bauser et al., 2019). No entanto, de acordo com a literatura, não está totalmente definido

que tipo de tarefa vocal poderá ser a que melhor represente a perturbação vocal (Brockmann-Bauser et al., 2019). No presente estudo foi encontrada relevância nas várias tarefas utilizadas, sendo por isso importante combinar diferentes medidas provenientes de cada uma, demonstrando o perfil vocal de cada indivíduo de forma mais completa.

Em relação aos índices AVQI e DSI, os resultados do presente estudo seguem vários resultados já descritos na literatura. O índice AVQI apresentou correlação direta com o grau de disфонia, em linha com os resultados de estudos anteriores (Delgado Hernández et al., 2018; Kim et al., 2019). Este índice tem a vantagem de conjugar medidas obtidas da fala encadeada e vogal sustentada. Por outro lado, o índice DSI apresentou correlação inversa com o grau de disфонia, indicando que, quanto maior o grau de alteração global da voz, menor o valor do índice, tal como já havia sido descrito (Hakkesteeft et al., 2008; Wuyts et al., 2000). O DSI é considerado uma medida da função vocal. Como tal, poderá traduzir algumas das capacidades ou limitações no funcionamento vocal e representar uma medida de desempenho vocal e/ou perturbação vocal (Barsties et al., 2018). Valores mais elevados do DSI indicarão uma função vocal geralmente melhor e mais saudável. Além disso, o DSI é sensível ao treino vocal, tal como apontam vários estudos, e é também capaz de medir efeitos a longo prazo (Nusseck et al., 2019). No entanto, a falta de dados normativos é uma limitação do AVQI e do CPPS, uma vez que são poucos os estudos sobre estas medidas – sobretudo do AVQI – em pessoas sem perturbação da voz (Batthyany et al., 2019).

Os índices DSI e AVQI usam medidas no domínio do tempo, nomeadamente *jitter*, *shimmer* e *HNR*. Nas vozes extremamente alteradas e roucas, e por isso, aperiódicas, existe a possibilidade de aquelas medidas contribuírem com erros para os valores dos índices (Barsties et al., 2018). Estes índices são de cálculo relativamente fácil na prática clínica, pois partem de tarefas e/ou medidas que habitualmente surgem nos protocolos de avaliação vocal. Além disso, são medidas objetivas e sensíveis à presença de alteração vocal, podendo também medir progressos da intervenção terapêutica. O facto de existirem *scripts* destes dois índices para utilização no programa *Praat* representa uma facilidade de utilização acrescida. O *Praat* é um *software* gratuito e disponível online. É muito usado na avaliação clínica de voz e de fala, porque, entre outros aspetos, é relativamente fácil de usar e permite o uso de *script*. Neste sentido, este é um dos contributos que o presente estudo, de forma indireta, poderá proporcionar. O facto de se corroborar pela primeira vez num estudo realizado em Portugal (tanto quanto a autora tem conhecimento), quer a validade destes índices na avaliação clínica da voz, quer a sua acessibilidade, poderá contribuir para um uso mais alargado destas medidas por terapeutas da fala em Portugal.

Wuyts et al. (2000) referem que, a graus de disфонia 2 e 3 na escala GRBAS, correspondem valores negativos no índice DSI. No entanto, no presente estudo, nenhum indivíduo obteve valor negativo no índice DSI, o que poderá ser explicado pelas diferenças nos sistemas e programas, referidas anteriormente. São necessários mais estudos no sentido de melhorar a fiabilidade das medidas

necessárias ao cálculo do DSI para que aumente o seu uso e aplicabilidade e seja possível a comparação dos resultados (Aichinger et al., 2012).

Por motivos externos à investigação, o n da amostra deste estudo é reduzido, o que traz naturalmente limitações aos resultados obtidos e à generalização das conclusões apresentadas. No entanto, os indivíduos da amostra distribuíram-se ao longo da escala dos parâmetros perceptivos da qualidade vocal, tal como era pretendido, ainda que não tenha havido representação do grau severo. Este facto permitiu perceber como se comportaram os indivíduos da amostra, tendo em conta o grau de alteração vocal e relacionando com medidas acústicas de longo e curto termo. A metodologia de recolha e análise dos dados vocais ficou assim demonstrada e será possível replicar o estudo e alargar a amostra no futuro, verificando se as conclusões se comprovam.

Uma descrição válida da qualidade vocal alterada deve contemplar a avaliação durante a fala encadeada (Awan, Roy, & Dromey, 2009). Num próximo estudo, será interessante incluir a tarefa de fala encadeada também na avaliação áudio-percetiva.

A relação entre a avaliação áudio-percetiva e as medidas acústicas da qualidade vocal é variável. Na revisão sistemática de Roy et al. (2013), a maioria dos estudos incluídos fez uso da análise acústica para classificar os indivíduos com presença ou ausência de disfonia. São poucos os estudos que consideram a gama de graus de gravidade de disfonia e se os resultados da análise acústica variam de acordo com a gravidade da disfonia. Em termos de análise acústica, será mais fácil distinguir indivíduos sem disfonia de indivíduos com grau extremo de disfonia, comparativamente a distinguir indivíduos com disfonia de grau ligeiro de indivíduos sem disfonia. No presente trabalho, o grau extremo de disfonia não foi considerado dado não haver indivíduos classificados dessa forma. Os resultados demonstram, para várias medidas, que não houve grande diferença entre os grupos SA e DL. As diferenças nos valores, na maioria das medidas, verificaram-se entre os grupos SA e DL e o grupo DM.

6. Conclusão

Sabendo da importância da avaliação da fala encadeada para a melhor caracterização da qualidade vocal, sobretudo na disfonia, a presente investigação pretendeu colaborar no estudo de medidas acústicas e tarefas vocais que permitam considerar os aspetos vocais próprios da dinâmica da fala encadeada, contribuindo para a melhoria das práticas dos terapeutas da fala na avaliação da qualidade vocal.

Foi possível estabelecer relações para as medidas acústicas de longo termo, considerando a avaliação áudio-percetiva da qualidade vocal – o *gold standard* da avaliação da voz – e a avaliação acústica mais frequente na prática clínica, baseada sobretudo nas medidas de curto-termo.

Pelos resultados obtidos verifica-se a relevância e a necessidade de a avaliação vocal contemplar tarefas de fala encadeada em diferentes Int_{subj} . Além disso, registar os valores de intensidade que o

indivíduo é capaz de produzir ao longo do processo de reabilitação vocal, poderá dar uma maior noção da evolução ao longo do processo de reabilitação vocal.

Foram várias as medidas de longo termo que se correlacionaram com o grau de disфонia e com parâmetros vocais perceptivos auditivos. As medidas: dinâmica de intensidades, CPPS_{fala} e CPPS_{vogal} correlacionaram-se com o grau de disфонia; a medida (H1-H2)_{LTAS} correlacionou-se inversamente com o grau de disфонia e a tensão percebida auditivamente. A medida CPPS_{fala} correlacionou-se ainda com todos os parâmetros perceptivos considerados: rouquidão, sopro e tensão. A dinâmica de intensidades correlacionou-se com os índices DSI e AVQI e ainda com as medidas CPPS_{fala}, *jitter*, *shimmer* e *HNR*.

Também foi verificada a validade e acessibilidade, na avaliação clínica da voz, dos índices DSI e AVQI, que também se correlacionaram com o grau de disфонia, o que poderá contribuir para um uso mais alargado destas medidas por terapeutas da fala em Portugal.

Tendo-se comprovado a relevância de várias medidas acústicas de longo termo na avaliação das perturbações da voz, em investigações futuras poderá ser relevante estudar a relação entre estas medidas e outras medidas de avaliação da voz, nomeadamente parâmetros laringoscópicos, aerodinâmicos e electroglotográficos, e medidas de autoavaliação vocal.

As medidas acústicas de longo termo são muitas vezes consideradas, à partida, de mais difícil extração e interpretação, o que pode reduzir a sua utilização de forma corrente na prática clínica. O presente estudo demonstra que a sua obtenção e interpretação já se tornaram, entretanto, mais fáceis, e que podem ser um contributo importante para a avaliação da qualidade vocal e para a demonstração de resultados de processos de reabilitação vocal. A formação e a divulgação de informação junto dos terapeutas da fala serão importantes na melhoria das práticas de avaliação acústica da voz.

Garantir a maior fiabilidade das medições realizadas, com a calibração do sinal acústico revela-se essencial. A calibração pode ser conseguida pela utilização de equipamento que faça automaticamente essa calibração, o que se revela como uma mais-valia, quer em termos de resultados, quer em termos do tempo despendido no momento da avaliação propriamente dita.

A utilização dos índices que, de forma (mais ou menos) automática devolvem informação quantitativa, pode ter relevância no uso clínico, podendo ser incluídos nos protocolos de avaliação acústica. Mais uma vez, reforça-se a importância de obter medidas de referência para o PE em indivíduos sem patologia vocal.

Em suma, constata-se que é possível avaliar acusticamente a fala encadeada, um comportamento representativo da voz falada em contexto natural. Tendo como referência medidas de curto termo e áudio-perceptivas, foi possível estudar a contribuição de tarefas de fala encadeada, de medidas acústicas de longo termo e de medidas combinadas na avaliação acústica da qualidade vocal,

contribuindo para uma mais correta caracterização da voz, com e sem perturbações, pelos terapeutas da fala.

Referências Bibliográficas

- Aichinger, P., Feichter, F., Aichstill, B., Bigenzahn, W., & Schneider-Stickler, B. (2012). Inter-device reliability of DSI measurement. *Logopedics Phoniatrics Vocology, 37*(4), 167–173.
- Alabdullah, T., Yahya, S., & Ramayah, T. (2014). Corporate governance mechanisms and Jordanian companies' financial performance. *Asian Social Science, 10*(22), 247.
- Almeida, S. C. (2016). *Validity and reliability of the 2nd European Portuguese version of the "Consensus Auditory-Perceptual Evaluation of Voice"*. Escola Superior de Saúde do Instituto Politécnico de Setúbal. Obtido de <http://comum.rcaap.pt/handle/10400.26/17609>
- Araújo, P. A. (2015). Describing Vocal Performance of Portuguese Cantadeiras of Alto Minho. *Departamento de Comunicação e Arte: Universidade de Aveiro*.
- Aronson, A. E., & Bless, D. M. (1990). *Clinical Voice Disorders* New York. Thieme Stratton.
- Awan, S. N., & Awan, J. A. (2020). A Two-Stage Cepstral Analysis Procedure for the Classification of Rough Voices. *Journal of Voice, 34*(1), 9–19.
- Awan, S. N., & Lawson, L. L. (2009). The effect of anchor modality on the reliability of vocal severity ratings. *Journal of Voice, 23*(3), 341–352.
- Awan, S. N., Roy, N., & Dromey, C. (2009). Estimating dysphonia severity in continuous speech: Application of a multi-parameter spectralcepstral model estimating dysphonia severity in continuous speech. *Clinical Linguistics and Phonetics, 23*(11), 825–841.
- Baken, R. J. (2005). The aged voice: a new hypothesis. *Journal of Voice, 19*(3), 317–325.
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. Cengage Learning.
- Barsties, B., & De Bodt, M. (2015). Assessment of voice quality: Current state-of-the-art. *Auris Nasus Larynx, 42*(3), 183–188.
- Barsties, B., & Maryn, Y. (2016). External validation of the Acoustic Voice Quality Index version 03.01 with extended representativity. *Annals of Otology, Rhinology & Laryngology, 125*(7), 571–583.
- Barsties, B., Ulozaitė-Stanienė, N., Maryn, Y., Petrauskas, T., & Uloza, V. (2017). The Influence of Gender and Age on the Acoustic Voice Quality Index and Dysphonia Severity Index: A Normative Study. *Journal of Voice, 33*(3), 340–345.
- Barsties, B., Ulozaitė-Stanienė, N., Petrauskas, T., Uloza, V., & Maryn, Y. (2018). Diagnostic Accuracy of Dysphonia Classification of DSI and AVQI. *Laryngoscope, 129*(3), 692–698.
- Batthyany, C., Maryn, Y., Trauwaen, I., Caelenberghe, E., van Dinther, J., Zarowski, A., & Wuyts, F. (2019). A case of specificity: how does the Acoustic Voice Quality Index perform in normophonic subjects? *Applied Sciences, 9*(12), 2527.
- Beck, J. M., & Schaeffler, F. (2015). Voice quality variation in Scottish adolescents: gender versus geography. Em *18th International Congress of Phonetic Sciences*. Glasgow.
- Behlau, M., Azevedo, R., & Madazio, G. (2001). Anatomia da Laringe e Fisiologia da Produção Vocal. Em M.

- Behlau (Ed.), *Voz: O Livro do Especialista*. Rio de Janeiro: Revinter.
- Behlau, M., Azevedo, R., & Pontes, P. (2001). Conceito de voz normal e classificação das disfonias. Em M. Behlau (Ed.), *Voz: O Livro do Especialista* (pp. 53–84). Rio de Janeiro: Revinter.
- Behlau, M., Madazio, G., Feijó, D., & Pontes, P. (2001). Avaliação de Voz. Em M. Behlau (Ed.), *Voz: O Livro do Especialista* (Vol. 1, pp. 86–180). Rio de Janeiro: Revinter.
- Behlau, Mara. (2001). *Voz: o livro do especialista. vol. 1*. Rio de Janeiro: Revinter.
- Behrman, A. (2005). Common practices of voice therapists in the evaluation of patients. *Journal of Voice*, 19(3), 454–469.
- Boë, L.-J., Sawallis, T. R., Fagot, J., Badin, P., Barbier, G., Captier, G., ... Schwartz, J.-L. (2019). Which way to the dawn of speech?: Reanalyzing half a century of debates and data in light of speech science. *Science Advances*, 5(12).
- Brinca, L., Batista, A. P. F., Tavares, A. I., Gonçalves, I. C., Moreno, M. L., In, A., & Gonc, C. (2014). Use of cepstral analyses for differentiating normal from dysphonic voices: A comparative study of connected speech versus sustained vowel in european portuguese female speakers. *Journal of Voice*, 28(3), 282–286.
- Brinca, L., Batista, A. P., Tavares, A. I., Pinto, P. N., & Araújo, L. (2015). The Effect of Anchors and Training on the Reliability of Voice Quality Ratings for Different Types of Speech Stimuli. *Journal of Voice*, 29(6), 776.e7–776.e14.
- Brockmann-Bausser, M., Van Stan, J. H., Carvalho Sampaio, M., Bohlender, J. E., Hillman, R. E., & Mehta, D. D. (2019). Effects of Vocal Intensity and Fundamental Frequency on Cepstral Peak Prominence in Patients with Voice Disorders and Vocally Healthy Controls. *Journal of Voice*, 1–7.
- Carding, P. N., Wilson, J. A., MacKenzie, K., & Deary, I. J. (2009). Measuring voice outcomes: state of the science review. *The journal of laryngology and otology*, 123(8), 823.
- da Silva, P. T., Master, S., Andreoni, S., Pontes, P., Ramos, L. R., Andreoni, S., ... da Silva, P. T. (2011). Acoustic and Long-Term Average Spectrum Measures to Detect Vocal Aging in Women. *Journal of Voice*, 25(4), 411–419.
- Daniel, E. (2018). *Caracterização Espectral da Voz de Atores de Dobragem*. Escola Superior de Saúde - Instituto Politécnico do Porto.
- Dejonckere, P. H., Bradley, P., Clemente, P., Cornut, G., Crevier-Buchman, L., Friedrich, G., ... Woisard, V. (2001). A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *European Archives of Oto-Rhino-Laryngology*, 258(2), 77–82.
- Delgado-Hernández, J., León-Gómez, N., & Jiménez-Álvarez, A. (2019). Precisión diagnóstica del pico cepstral de mayor prominencia en el cepstrum suavizado (CPPS) en la detección de la disfonía en español. *Loquens*, 6(1), 10–3989.

- Delgado Hernández, J., León Gómez, N. M., Jiménez, A., Izquierdo, L. M., & Barsties v. Latoszek, B. (2018). Validation of the Acoustic Voice Quality Index Version 03.01 and the Acoustic Breathiness Index in the Spanish language. *Annals of Otology, Rhinology and Laryngology*, *127*(5), 317–326.
- Eadie, T. L., & Baylor, C. R. (2006). The effect of perceptual training on inexperienced listeners' judgments of dysphonic voice. *Journal of voice*, *20*(4), 527–544.
- Eadie, T. L., & Doyle, P. C. (2005). Classification of dysphonic voice: acoustic and auditory-perceptual measures. *Journal of Voice*, *19*(1), 1–14.
- Englert, M., Barsties v. Latoszek, B., Maryn, Y., & Behlau, M. (2019). Validation of the Acoustic Voice Quality Index, Version 03.01, to the Brazilian Portuguese Language. *Journal of Voice*, 1–7.
- Eskenazi, L., Childers, D. G., & Hicks, D. M. (1990). Acoustic correlates of vocal quality. *Journal of Speech, Language, and Hearing Research*, *33*(2), 298–306.
- Fawcus, M. (2001). As causas e a classificação dos distúrbios vocais. Em M. Fawcus (Ed.), *Disfonias Diagnóstico e Tratamento* (2ª, pp. 19–38). Revinter.
- Freitas, S. (2009). A avaliação das alterações vocais: registo e análise áudio-perceptual e acústica da voz. Em *Metodologias de Intervenção em Terapia da Fala* (pp. 225–253). Porto: Edições Universidade Fernando Pessoa.
- Freitas, S. (2012). *Avaliação Acústica e Áudio Percetiva na Caracterização da Voz Humana*. Faculdade de Engenharia da Universidade do Porto. Faculdade de Engenharia da Universidade do Porto.
- Gomes, L. R. (2005). *Biofísica para Ciências da Saúde*. Porto: Universidade Fernando Pessoa.
- Guimarães, I. (2007). *A ciência e a arte da voz humana*. Alcoitão, Escola Superior de Saúde de Alcoitão.
- Guimarães, I., & Abberton, E. (2005). Fundamental Frequency in Speakers of Portuguese for Different Voice Samples. *Journal of Voice*, *19*(4), 592–606.
- Guzman, M., Correa, S., Muñoz, D., & Mayerhoff, R. (2013). Influence on spectral energy distribution of emotional expression. *Journal of Voice*, *27*(1), 129.e1-129.e10.
- Hakkesteegt, M. M., Brocaar, M. P., Wieringa, M. H., & Feenstra, L. (2008). The relationship between perceptual evaluation and objective multiparametric evaluation of dysphonia severity. *Journal of Voice*, *22*(2), 138–145.
- Halberstam, B. (2004). Acoustic and Perceptual Parameters Relating to Connected Speech Are More Reliable Measures of Hoarseness than Parameters Relating to Sustained Vowels. *ORL*, *66*(2), 70–73.
- Hallin, A. E., Fröst, K., Holmberg, E. B., & Södersten, M. (2012). Voice and speech range profiles and Voice Handicap Index for males—methodological issues and data. *Logopedics Phoniatrics Vocology*, *37*(2), 47–61.
- Hammarberg, B., Fritzell, B., Gauffin, J., Sundberg, J., & Wedin, L. (1981). Perceptual and Acoustic Correlates of Abnormal Voice Qualities. *Acta Otolaryngologica*, *90*, 441–451.

- Heman-Ackah, Y. D., Heuer, R. J., Michael, D. D., Ostrowski, R., Horman, M., Baroody, M. M., ... Sataloff, R. T. (2003). Cepstral peak prominence: A more reliable measure of dysphonia. *Annals of Otology, Rhinology and Laryngology*, *112*(4), 324–333.
- Heman-Ackah, Y. D., Michael, D. D., & Goding, G. S. (2002). The relationship between cepstral peak prominence and selected parameters of dysphonia. *Journal of Voice*, *16*(1), 20–27.
- Heman-Ackah, Y. D., Sataloff, R. T., Laureyns, G., Lurie, D., Michael, D. D., Heuer, R., ... Hillenbrand, J. (2014). Quantifying the Cepstral Peak Prominence, a Measure of Dysphonia. *Journal of Voice*, *28*(6), 783–788.
- Hillenbrand, J., Cleveland, R. A., & Erickson, R. L. (1994). Acoustic correlates of breathy vocal quality. *Journal of Speech, Language, and Hearing Research*, *37*(4), 769–778.
- Hosokawa, K., v Latoszek, B. B., Iwahashi, T., Iwahashi, M., Iwaki, S., Kato, C., ... Matsushiro, N. (2019). The acoustic voice quality index version 03.01 for the Japanese-speaking population. *Journal of Voice*, *33*(1), 125–e1.
- Imamura, R., Tsuji, D., & Sennes, L. (2006). Fisiologia da Laringe. Em S. M. R. . Pinho, D. H. Tsuji, & C. Bohadana (Eds.), *Fundamentos em Laringologia e Voz* (pp. 1–20). Revinter.
- Jackson-Menaldi, M. C. A. (2002). *La voz patológica*. Buenos Aires: Ed. Médica Panamericana.
- Jayakumar, T., & Savithri, S. R. (2012). Effect of geographical and ethnic variation on Dysphonia Severity Index: a study of Indian population. *Journal of Voice*, *26*(1), e11–e16.
- Kim, G.-H., von Latoszek, B. B., & Lee, Y.-W. (2019). Validation of Acoustic Voice Quality Index Version 3.01 and Acoustic Breathiness Index in Korean Population. *Journal of Voice*.
- Koo, T. K. ., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, *15*(2), 155–163.
- Lee, J. M., Roy, N., Peterson, E., & Merrill, R. M. (2017). Comparison of Two Multiparameter Acoustic Indices of Dysphonia Severity: The Acoustic Voice Quality Index and Cepstral Spectral Index of Dysphonia. *Journal of Voice*, *32*(4), 515.e1–515.e13.
- Leino, T. (2009). Long-Term Average Spectrum in Screening of Voice Quality in Speech: Untrained Male University Students. *Journal of Voice*, *23*(6), 671–676.
- Lowell, S. Y., Colton, R. H., Kelley, R. T., & Hahn, Y. C. (2011). Spectral- and cepstral-based measures during continuous speech: Capacity to distinguish dysphonia and consistency within a speaker. *Journal of Voice*, *25*(5), e223–e232.
- Lowell, S. Y., Kelley, R. T., Awan, S. N., Colton, R. H., & Chan, N. H. (2012). Spectral-and cepstral-based acoustic features of dysphonic, strained voice quality. *Annals of Otology, Rhinology & Laryngology*, *121*(8), 539–548.
- Ma, E. P.-M., & Yiu, E. M.-L. (2006). Multiparametric evaluation of dysphonic severity. *Journal of Voice*, *20*(3), 380–390.

- Magalhães, A. C. (2018). *Comparação da qualidade vocal de indivíduos jovens e idosos através do espectro médio de longa-duração (LTAS)*. Escola Superior de Saúde - Instituto Politécnico do Porto.
- Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice*, *24*(5), 540–555.
- Maryn, Y., Morsomme, D., & De Bodt, M. (2017). Measuring the Dysphonia Severity Index (DSI) in the Program Praat. *Journal of Voice*, *31*(5), 644.e29-644.e40.
- Maryn, Y., Roy, N., De Bodt, M., Van Cauwenberge, P., & Corthals, P. (2009). Acoustic measurement of overall voice quality: A meta-analysis. *The Journal of the Acoustical Society of America*, *126*(5), 2619–2634.
- Maryn, Y., & Weenink, D. (2015). Objective dysphonia measures in the program praat: Smoothed cepstral peak prominence and acoustic voice quality index. *Journal of Voice*, *29*(1), 35–43.
- Master, S., Biase, N. de, Pedrosa, V., & Chiari, B. M. (2006). O espectro médio de longo termo na pesquisa e na clínica fonoaudiológica. *Pró-Fono Revista de Atualização Científica*, *18*(1), 111–120.
- Mateus, M. H. M., Falé, I., & Freitas, M. J. (2005). *Fonética e fonologia do português*. Lisboa: Universidade Aberta.
- Mendes, A. (2012). Softwares e hardwares de análise acústica da voz e da fala. *Distúrbios da Comunicação*, *24*(3).
- Mendes, A., Costa, A. N. da, Martins, A. D., Fernandes, A. F. O., Vicente, S. M. D. do R., & Freitas, T. C. S. (2012). Contributos para a construção de um texto foneticamente equilibrado para o Português-Europeu. *Revista CEFAC*, *14*(5), 910–917.
- Murton, O., Hillman, R., & Mehta, D. (2020). Cepstral peak prominence values for clinical voice evaluation. *American Journal of Speech-Language Pathology*, *29*(3), 1596–1607.
- Nemr, K., Simões-Zenari, M., de Souza, G. S., Hachiya, A., & Tsuji, D. H. (2016). Correlation of the Dysphonia Severity Index (DSI), Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V), and Gender in Brazilians With and Without Voice Disorders. *Journal of Voice*, *30*(6), 765.e7-765.e11.
- Ng, M. L., Chen, Y., & Chan, E. Y. K. (2012). Differences in vocal characteristics between Cantonese and English produced by proficient Cantonese-English bilingual speakers - A long-term average spectral analysis. *Journal of Voice*, *26*(4), e171–e176.
- Nordenberg, M., & Sundberg, J. (2003). Effect on LTAS of vocal loudness variation. *Tmh-Qpsr*, *45*(1), 093–100.
- Núñez-Batalla, F., Cartón-Corona, N., Vasile, G., García-Cabo, P., Fernández-Vañes, L., & Llorente-Pendás, J. L. (2019). Validation of the Measures of Cepstral Peak Prominence as a Measure of Dysphonia Severity in Spanish-Speaking Subjects. *Acta Otorrinolaringologica (English Edition)*,

- 70(4), 222–228.
- Núñez-Batalla, F., Díaz-Fresno, E., Álvarez-Fernández, A., Muñoz Cordero, G., & Llorente Pendás, J. L. (2017). Application of the Acoustic Voice Quality Index for Objective Measurement of Dysphonia Severity. *Acta Otorrinolaringologica Espanola (English Edition)*, 68(4), 204–211.
- Nusseck, M., Immerz, A., Spahn, C., Echternach, M., & Richter, B. (2019). Long-Term Effects of a Voice Training Program for Teachers on Vocal and Mental Health. *Journal of Voice*.
- Oates, J. (2009). Auditory-perceptual evaluation of disordered voice quality. *Folia Phoniatica et Logopaedica*, 61(1), 49–56.
- P S, S., & Pebbili, G. K. (2020). Cepstral Analysis of Voice in Young Adults. *Journal of Voice*.
- Parsa, V., & Jamieson, D. G. (2006). Acoustic Discrimination of Pathological Voice. *Journal of Speech, Language, and Hearing Research*, 44(2), 327–339.
- Patel, R. R., Eadie, T., Paul, D., Hillman, R., Barkmeier-Kraemer, J., Awan, S. N., ... Švec, J. G. (2018). Recommended Protocols for Instrumental Assessment of Voice: American Speech-Language-Hearing Association Expert Panel to Develop a Protocol for Instrumental Assessment of Vocal Function. *American Journal of Speech-Language Pathology*, 27(3), 887–905.
- Pereira, T. D. (2018). *Caracterização de Falantes do Português Europeu através do Espectro Médio de Longa Duração: comparação entre diferentes tarefas de fala*. Escola Superior de Saúde - Politécnico do Porto.
- Phadke, K. V., Laukkanen, A. M., Ilomäki, I., Kankare, E., Geneid, A., & Švec, J. G. (2018). Cepstral and Perceptual Investigations in Female Teachers With Functionally Healthy Voice. *Journal of Voice*.
- Pinho, S. (1998). Avaliação e tratamento da voz. Em S. Pinho (Ed.), *Fundamentos em fonoaudiologia: tratando os distúrbios de voz* (Vol. 2, pp. 1–40). Editora Guanabara Koogan.
- Pinho, S., & Pontes, P. (2002). Escala de Avaliação Percetiva da Fonte Glótica: RASAT, 1–4.
- Pinho, S., & Pontes, P. (2008). Músculos Intrínsecos da Laringe e Dinâmica Vocal. *Desvendando os Segredos da Voz - Volume I*. Rio de Janeiro: Revinter.
- Pommée, T., Maryn, Y., Finck, C., & Morsomme, D. (2020). Validation of the Acoustic Voice Quality Index, Version 03.01, in French. *Journal of Voice*, 34(4), 646.e11-646.e26.
- Portela, A. R. (2018). *Caracterização de Falantes do Português Europeu através do Espectro Médio de Longa Duração: comparação entre géneros*. Escola Superior de Saúde - Politécnico do Porto.
- Revis, J., Giovanni, A., Wuyts, F., & Triglia, J.-M. (1999). Comparison of Different Voice Samples for Perceptual Analysis. *Folia Phoniatica et Logopaedica*, 51(3), 108–116.
- Ribeiro, J. L. P. (1999). *Investigação e avaliação em psicologia e saúde* (1ª edição). Lisboa: Climepsi Editores.
- Roy, N., Barkmeier-Kraemer, J., Eadie, T., Sivasankar, M. P., Mehta, D., Paul, D., & Hillman, R. (2013). Evidence-based clinical voice assessment: a systematic review. *American Journal of Speech-*

- Language Pathology*, 22(2), 212–226.
- Santos, R. J. F. (2009). *Avaliação de Pacientes com Paralisia Unilateral das Pregas Vocais*. Dissertação de Mestrado. Universidade de Aveiro.
- Sauder, C., Bretl, M., & Eadie, T. (2017). Predicting Voice Disorder Status From Smoothed Measures of Cepstral Peak Prominence Using Praat and Analysis of Dysphonia in Speech and Voice (ADSV). *Journal of Voice*, 31(5).
- Siupsinskiene, N., & Lycke, H. (2011). Effects of vocal training on singing and speaking voice characteristics in vocally healthy adults and children based on choral and nonchoral data. *Journal of Voice*, 25(4), e177–e189.
- SPTF. (2020). *Dicionário Terminológico de Terapia da Fala* (1ª). Lisboa: Papa-Letras.
- Švec, J. G., & Granqvist, S. (2010). Guidelines for selecting microphones for human voice production research. *American Journal of Speech–Language Pathology*, 19(4), 356–368.
- Švec, J. G., & Granqvist, S. (2018). Tutorial and Guidelines on Measurement of Sound Pressure Level in Voice and Speech. *Journal of Speech, Language, and Hearing Research*, 61(3), 441–461.
- Titze, I. R. (1995). *Workshop on acoustic voice analysis: Summary statement*. National Center for Voice and Speech.
- Titze, I., Riede, T., & Popolo, P. (2008). Nonlinear source–filter coupling in phonation: Vocal exercises. *The Journal of the Acoustical Society of America*, 123(4), 1902–1915.
- Vieira, M. N., & Rosa, L. L. C. (2006). Avaliação acústica na prática fonoaudiológica. Em *Pinho SMR, Tsuji DH, Bohadana SC, organizadores. Fundamentos em laringologia e voz*. Rio de Janeiro: Revinter (pp. 33–52).
- Watts, C. R., & Awan, S. N. (2015). An examination of variations in the cepstral spectral index of dysphonia across a single breath group in connected speech. *Journal of Voice*, 29(1), 26–34.
- Watts, C. R., Awan, S. N., & Maryn, Y. (2016). A Comparison of Cepstral Peak Prominence Measures From Two Acoustic Analysis Programs. *Journal of Voice*, 31(3), 387.e1–387.e10.
- Weenink, D. (2019). PowerCepstrum: Get peak prominence. Obtido 3 de Outubro de 2020, de https://www.fon.hum.uva.nl/praat/manual/PowerCepstrum__Get_peak_prominence___html
- Wolfe, V., & Martin, D. (1997). Acoustic correlates of dysphonia: Type and severity. *Journal of Communication Disorders*, 30(5), 403–416.
- Wuyts, F. L., Bodt, M. S. De, Molenberghs, G., Remacle, M., Heylen, L., Millet, B., ... Heyning, P. H. Van de. (2000). The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *Journal of speech, language, and hearing research*, 43(3), 796–809.
- Yiu, E. M. L. (1999). Limitations of perturbation measures in clinical acoustic voice analysis. *Asia Pacific Journal of Speech, Language and Hearing*, 4(3), 155–166.
- Zemlin, W. R. (2005). *Princípios de anatomia e fisiologia em fonoaudiologia* (4ª edição). Porto Alegre:

Artes Médicas Sul.

Apêndices

Apêndice 1 – Ficha de Caracterização do Participante

CÓDIGO _____

A – Dados extraídos do processo clínico

Sexo __ Idade ___

Diagnóstico clínico _____

B – Dados obtidos no contacto prévio (telefone/presencial)

Língua materna Português Europeu: Sim___ Não___

Literacia _____

Alt. acuidade auditiva: Sim___ Não___

Alt. acuidade visual: Sim___ Não___

C – Dados a observar no momento da avaliação vocal

Alterações da fala (pert. motoras da fala, da fluência, da articulação ou outras): Sim___ Não___

Apêndice 2 – Protocolo de Gravação da Voz – Tarefas

- 1. Índice s/z (Mara Behlau, 2001):**
 - a. Produção prolongada de fricativa não vozeada /s/, 3 vezes
 - b. Produção prolongada de fricativa vozeada /z/, 3 vezes

- 2. Produção de vogal sustentada [a] (Almeida, 2016)**
 - a. tempo máximo de emissão, 3 vezes
 - b. mínimo 4 segundos

- 3. Produção de vogal [a] na mínima intensidade possível (excluindo sussurro)**

- 4. Produção de Glissandos – vogal [a]- (Patel et al., 2018)**
 - a. Glissando ascendente
 - b. Glissando descendente

- 5. Leitura de Frases – CAPE-V EP 2ª versão (Almeida, 2016)**

- 6. Leitura do texto “O Sol” (Mendes et al., 2012) em três intensidades subjetivas diferentes: média, forte e fraca**

- 7. Discurso espontâneo (mínimo 20 seg.): “Fale-me do sítio onde cresceu”**

- 8. Canto da música “Parabéns a você”**

Apêndice 3 – Registo da avaliação áudio-percetiva (página Web)

Teste de percepção auditiva

Um conjunto de 35 vogais [a] produzidas de forma prolongada por diferentes indivíduos será sujeito à sua avaliação áudio-percetiva. A análise das cinco primeiras amostras tem como objetivo permitir a familiarização com a escala. A escala a utilizar é uma adaptação do *Consensus Auditory-Perceptual Evaluation of Voice* (CAPE -V).

Indicações: Em primeiro lugar deverá escutar a amostra vocal clicando sobre o botão "play" (▶).

Após a escuta deverá considerar os seguintes parâmetros da qualidade vocal: **grau de severidade global, rouquidão, soprosidade e tensão**. Para cada parâmetro existe uma linha que representa uma escala visual analógica. Ao limite à esquerda corresponde ausência de desvio (alteração) e à direita refere-se ao máximo de desvio.

Sob a linha estão indicadores qualitativos: DL - desvio ligeiro, DM – desvio moderado, DS – desvio severo. Deverá indicar o grau de alteração percebido para cada um dos parâmetros vocais deslizando o cursor ao longo da linha.

No final, poderá indicar outros parâmetros que julgue serem relevantes na avaliação da amostra de voz no campo "Fatores adicionais".

O teste encontra-se dividido em 7 páginas. No fundo de cada página, poderá encontrar três botões diferentes.

Para navegar entre páginas deverá utilizar os botões 'Página anterior' e 'Página seguinte'.

O botão 'Guardar' permite submeter os dados introduzidos numa página, relativos a um conjunto de 5 vozes. Poderá responder, alterar e submeter mais do que uma vez. Tenha em conta que uma nova submissão de dados de uma página irá substituir a submissão anterior.

Atenção: Se tiver introduzido novos dados numa página e navegar para outra sem ter carregado previamente no botão 'Guardar', os novos dados serão perdidos.

Sigla	Legenda
DL	desvio ligeiro
DM	desvio moderado
DS	desvio severo

(Página 1/7)

Voz 1

▶ 0:00 / 0:04 ————— ◀

Grau de severidade global DL DM DS

Rouquidão DL DM DS

Soprosidade DL DM DS

Tensão DL DM DS

Fatores adicionais

(por ex: diplofonia, aspereza, falsofona, astenia, afonia, bitonalidade, tremor, estridência, "glottal fry", outros aspetos relevantes)

Apêndice 4 – Autorização da Comissão de Ética da ULSM para realização do estudo

UNIDADE LOCAL DE SAÚDE DE MATOSINHOS HOSPITAL PEDRO HISPANO	INFORMAÇÃO	Nº 123/CE/IAS Data: 13-12-2019
--	------------	---------------------------------------

Para: Serviço de Gestão do Conhecimento
De: Comissão de Ética

Assunto: Apreciação ao pedido de autorização para realização de estudo intitulado "Avaliação complementar da voz através de medidas acústicas de longo termo em vozes disfónicas"


INFORMAÇÃO

Exmo. Senhor,

A Comissão de Ética analisou o pedido de autorização para realização de estudo intitulado "Avaliação complementar da voz através de medidas acústicas de longo termo em vozes disfónicas", cuja autora é Joana Costa Frutuoso Faria, Terapeuta da Fala na ULSM, no âmbito do Mestrado em Terapia da Fala na Escola Superior de Saúde do Instituto Politécnico do Porto.

A Comissão de Ética para a Saúde da ULSM deliberou, por unanimidade, nada opor à realização deste estudo.

Com os melhores cumprimentos


Dr. José Alberto Silva
Presidente da Comissão de Ética
da ULSM Matosinhos

Dr. José Alberto Silva
(Presidente da Comissão de Ética da ULSM, E.P.E.)

Apêndice 5 – Termo de Consentimento Informado

Declaração de consentimento informado

Conforme a lei 67/98 de 26 de Outubro e a “Declaração de Helsínquia” da Associação Médica Mundial (Helsínquia 1964; Tóquio 1975; Veneza 1983; Hong Kong 1989; Somerset West 1996, Edimburgo 2000; Washington 2002, Tóquio 2004, Seul 2008, Fortaleza 2013) – quando se aplicar

Designação do Estudo: Avaliação complementar da voz através de medidas acústicas de longo termo em vozes disfónicas

Eu, abaixo-assinado _____
fui informado de que o Estudo de Investigação acima mencionado, se destina à Avaliação da Qualidade Vocal através de Espectros Médios de Longa Duração na Caracterização da Voz com Disfonia.

Tomei conhecimento que se trata de um estudo que obteve aprovação pela Comissão de Ética e Saúde da Unidade Local de Saúde de Matosinhos.

Sei que neste estudo está prevista a realização e gravação de tarefas de natureza vocal, para posterior análise. A minha voz será gravada durante a produção de vogais, leitura de frases, discurso espontâneo e canto. Posteriormente, será realizada a análise áudio-perceptiva e acústica da minha voz nas diferentes tarefas gravadas.

Estes procedimentos não são invasivos e não têm quaisquer riscos associados. Os registos serão arquivados na base de dados do Centro de Investigação em Reabilitação, estando a sua consulta reservada apenas aos membros da equipa de investigação do projeto. A base de dados existente foi criada no âmbito do projeto "Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação", da responsabilidade do Prof. Doutor André Araújo. Os dados serão armazenados durante 6 anos em disco rígido externo, à guarda do investigador principal, na Escola Superior de Saúde do Politécnico do Porto. Após este período, caso não se verifique necessidade de voltar a recolher novos parâmetros acústicos, os ficheiros áudio serão destruídos, mantendo-se apenas as bases de dados estatísticos.

Foi-me garantido que todos os dados relativos à identificação dos Participantes neste estudo são confidenciais e que será mantido o anonimato, com a atribuição de um código a cada participante.

Compreendi a informação que me foi dada, tive oportunidade de fazer perguntas e as minhas dúvidas foram esclarecidas.

Sei que posso recusar-me a participar ou interromper a qualquer momento a participação no estudo, sem nenhum tipo de penalização por este facto.

Aceito participar de livre vontade no estudo acima mencionado. Também autorizo a utilização dos dados e a divulgação dos resultados obtidos para fins científicos e académicos, desde que seja garantido o anonimato.

Nome do Investigador e Contacto: Joana Costa Frutuoso, joanafrutuosotf@gmail.com

Data

___/___/___

Assinatura

Anexo – Tarefas de leitura

- Leitura de Frases – CAPE-V EP 2ª versão (Almeida, 2016)

Num domingo estive sol e fui com o avô António à esplanada “Évora” comer uma empada.

Segundo Simão, só Samuel sabe.

A Zé, mãe do Gabriel, deu-lhe um bolo de laranja e vinho velho de Runa.

É hora da Urraca ir à caça.

Onde eu brinco há um ninho de andorinhas encostado ao muro.

A Kika tapou a tua capa preta.

- Texto “O Sol” (Mendes et al., 2012)

“O Sol é uma velha estrela que aquece e ilumina o nosso planeta todos os dias mas de uma forma desigual. Possui um terço de hidrogénio, hélio e outros gases, mas nenhum diferente dos terrestres. O Sol é uma estrela de tamanho médio e cor amarela, que se encontra a metade da sua vida. É a única em todo o sistema solar e a mais próxima da Terra, muito importante para a existência de vida. Sem o brilhar do Sol, a Terra ficaria fria, sem plantas, mais pobre e menos bela. O Sol tem um aspeto granulado e os seus grãos são locais por onde emerge o calor. Nas zonas escuras ou manchas é onde se encontra a maior intensidade do campo gravitacional solar.”