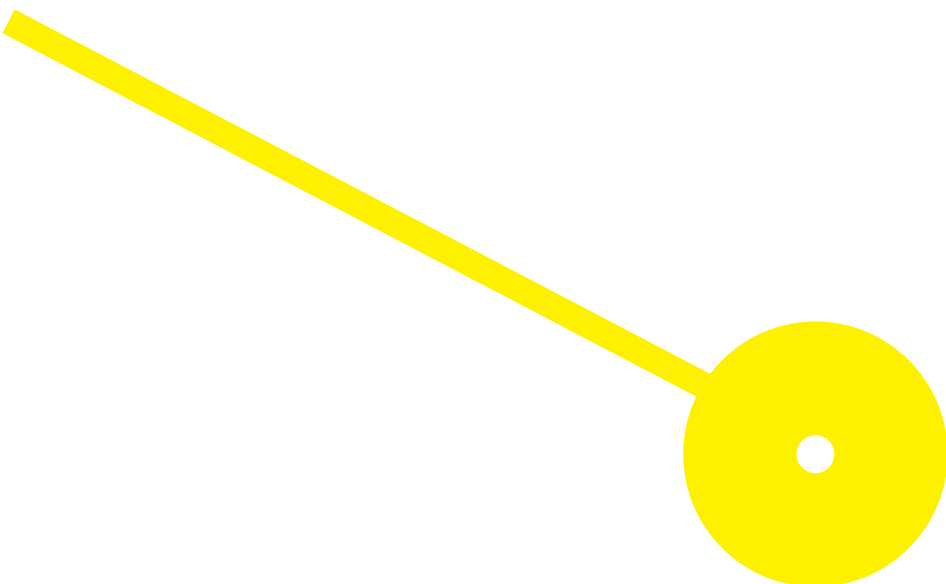




Análise Multivariada de Propriedades Acústicas da Voz para a Caracterização do Género

Teresa Dias Pereira

12/2020





**ESCOLA
SUPERIOR
DE SAÚDE**



Centro de Investigação em Reabilitação
Center for Rehabilitation Research

Análise Multivariada de Propriedades Acústicas da Voz para a Caracterização do Género

Autor

Teresa Dias Pereira

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.

Dedicatória

Ao meu avô e a todos os avós que se perderam sem ter os que amam por perto.

Agradecimentos

Ao meu orientador, Professor Doutor Pedro André Araújo, **por acreditar** em mim e nas minhas competências, **por valorizar** o meu trabalho, **por estar** sempre disponível, **por me proporcionar** todas as oportunidades, **por partilhar** o seu conhecimento e **por nunca me deixar** desistir.

À Professora Carla Oliveira pela **disponibilidade** e pelo **tempo** dispensado para ajudar nos procedimentos de análise estatística.

À Ana e à Rita pela **partilha**, que não é de agora, por todos os **momentos vividos** e por todas as palavras de **força** e **perseverança**.

À Ticha e à Andreia pela **amizade**, **motivação** e pela **imparcialidade** nas suas sugestões.

Aos amigos e à família pela **força**.

Ao Pedro por estar **presente**, por **compreender**, por **ajudar**, por **amar**.

Resumo

Introdução: A análise acústica contribui com dados objetivos na caracterização da voz normal e patológica. O Terapeuta da Fala intervém em quadros clínicos com comprometimento da perceção do género na voz dos falantes. A caracterização acústica do género pode revelar-se importante na avaliação e intervenção nestes casos. **Objetivo:** Analisar as propriedades acústicas da voz associadas à caracterização do género em falantes de Português Europeu (PE). **Metodologia:** Realizou-se um estudo observacional de carácter analítico transversal com 15 homens e 15 mulheres (18–25 anos) normofalantes de PE. Analisaram-se diferentes amostras de voz e extraíram-se medidas aerodinâmicas e acústicas de curto e longo termo. Os resultados foram obtidos a partir da estatística inferencial e multivariada. **Resultados:** A f_0 média, DP, α e $(H1-H2)_{LTAS}$ são superiores nas mulheres, a Extensão Vocal e CPP são superiores nos homens. As combinações de medidas representativas nos homens incluem f_0 , f_0DP , gama, HNR, *jitter* e *shimmer*, enquanto nas mulheres incluem f_0 , CPPS, *jitter* e *shimmer*. As medidas com mais peso na amostra total são mais próximas das obtidas nos homens. **Conclusões:** Este estudo demonstra a pertinência de estudar combinações de parâmetros acústicos na caracterização do género e aponta hipóteses para estudos futuros.

Palavras-chave: género, análise acústica, frequência fundamental, medidas acústicas de curto-termo, medidas acústicas de longo-termo

Abstract

Introduction: Acoustic analysis provides objective data to normal and pathological voice characterization. In clinical context, the Speech Therapist works with compromised gender perception in speakers voice. In these cases, the acoustic gender voices characterization can be important in assessment and intervention. **Objective:** To analyze voice acoustic properties associated with gender characterization in European Portuguese (PE) speakers. **Methodology:** An observational cross-sectional analytical study was carried out with 15 men and 15 women (18–25 years old) PE speakers. Aerodynamic, short and long term acoustic measurements were extracted from different voice samples. The results were obtained from inferential and multivariate statistics. **Results:** Mean f_0 , f_0SD , $alpha$ and $(H1-H2)_{LTAS}$ are higher in women, Vocal Extension and CPP are higher in men. Representative measurement combinations in men include f_0 , f_0DP , $gamma$, HNR , $jitter$ and $shimmer$, while in women they include f_0 , $CPPS$, $jitter$ and $shimmer$. Most relevant combination measures in the total sample are closer to those obtained in males. **Conclusions:** This study demonstrates the relevance of studying acoustic parameters' combinations in gender characterization and points out future studies hypotheses.

Keywords: gender, acoustic analysis, fundamental frequency, short-term acoustic measures, long-term acoustic measures

Índice

1. Introdução.....	1
2. Enquadramento Teórico.....	2
2.1. Avaliação da voz	5
2.1.1. Avaliação Perceptiva da Voz.....	5
2.1.2. Avaliação Aerodinâmica da Voz.....	7
2.1.3. Análise Acústica – Frequência Fundamental	8
2.1.4. Análise acústica – Medidas de Curto Termo.....	9
2.1.5. Análise acústica – Medidas de Longo Termo	10
2.2. Caracterização do género vocal.....	14
2.3. Relevância do género vocal para a Terapia da Fala	19
2.4. Objetivos da investigação	23
3. Métodos.....	24
3.1 Tipo e Desenho de estudo	24
3.2. Amostra	24
3.3. Instrumentos.....	25
3.4. Procedimentos.....	27
3.4.2. Recolha do Sinal.....	27
3.4.2. Procedimentos de Análise Acústica.....	28
3.4.3. Procedimentos de Análise Estatística	30
4. Resultados.....	31
5. Discussão.....	37
6. Conclusão.....	43
Referências Bibliográficas	45
Apêndices e Anexos.....	52

Índice de Tabelas

Tabela 1- Valores de Referência de f_0	16
Tabela 2- Valores de Referência para as Medidas de Curto Termo obtidos através do software Praat® para a vogal [a].....	17
Tabela 3 - Valores de referência para as medidas obtidas a partir do LTAS.....	18
Tabela 4 - Valores de Referência CPP e CPPS obtidos através do software Praat®	19
Tabela 5 – Procedimentos para extração dos valores de CPP e CPPS, para versão 6.1.34 do Praat, baseado em Phadke et al. (2018).....	29
Tabela 6 – Estatística Descritiva e Testes de Hipóteses para Medidas Aerodinâmicas	31
Tabela 7- Estatística Descritiva e Testes de Hipóteses para valores de f_0	32
Tabela 8 – Estatística Descritiva e Testes de Hipóteses para Parâmetros de Curto Termo	33
Tabela 9 – Estatística Descritiva e Testes de Hipóteses para medidas de LTAS	34
Tabela 10 – Estatística Descritiva e Testes de Hipóteses para medidas cepstrais.....	35
Tabela 11- CATPCA – Homens.....	36
Tabela 12 – Alpha de Cronbach Homens.....	36
Tabela 13 – CATPCA – Mulheres.....	36
Tabela 14 – Alpha de Cronbach Mulheres.....	36
Tabela 15 – CATPCA – Homens+Mulheres.....	37
Tabela 16- Alpha de Cronbach Homens+Mulheres	37

Índice de Figuras

Figura 1- Representação visual dos componentes usados no cálculo do alpha e do $(H1-H2)_{LTAS}$. Adaptado de Araújo (2015).....	12
Figura 2 - Representação gráfica do Cepstro e do cálculo da CPPS (Delgado-Hernandez et al., 2017).....	14
Figura 3- Scree Plot Eigenvalues Homens.....	36
Figura 4- Scree Plot Eigenvalues Mulheres.....	36
Figura 5- Scree Plot Eigenvalues Homens+Mulheres	37

Índice de Abreviaturas

(H1-H2) _{LTAS} – Diferença entre o primeiro e o segundo harmónico no LTAS.....	12
<i>alpha</i> – Rácio entre a energia acima e abaixo dos 1000Hz.....	12
C – Canto.....	24
CAPE-V – Escala de avaliação perceptiva da voz.....	6
CATPCA – Categorical Principal Components Analysis.....	29
CPP – Cepstral Peak Proeminence.....	12
CPPS – Smoothed Cepstral Peak Proeminence.....	13
dB – Decibéis.....	10
DE – Discurso espontâneo.....	16
DP – Desvio padrão.....	8
DSE – Discurso semi-espontâneo.....	24
ESS – Escola Superior de Saúde.....	23
ExtV – Extensão Vocal.....	25
f_0 – Frequência fundamental.....	8
FFT – Fast Fourier Transform.....	10
GA – Glissando ascendente.....	25
GD – Glissando descendente.....	25
GRBAS – Escala de avaliação perceptiva da voz.....	6
H1 – Primeiro harmónico, correspondente à frequência fundamental.....	8
H2 – Segundo harmónico.....	12
HNR – Harmonic to Noise Ratio.....	10
<i>jitter</i> – Medida de perturbação da frequência.....	9
Leq – Equivalent Sound Level.....	11
LFo – Leitura em intensidade forte.....	18
LFr – Leitura em intensidade fraca.....	24
LN – Leitura em intensidade normal.....	16
LTAS – Long Term Average Spectrum.....	11
PE – Português Europeu.....	1
<i>shimmer</i> – Medida de perturbação da intensidade.....	9
ST – Semitons.....	8
TMF – Tempo máximo de fonação.....	7
TpF – Terapeuta da fala.....	1
Va – Vogal [a].....	16
Vi – Vogal [i].....	24
VoiSS – Voice Symptom Scale.....	3
V-RQOL – Voice-Related Quality of Life.....	3
Vu – Vogal [u].....	24

1. Introdução

O ser humano utiliza a voz falada como meio privilegiado de transmissão de sentimentos e ideias. Isto é possível devido ao conteúdo emocional e à expressividade que a mesma é capaz de conferir a uma mensagem (Pinho, 2006). Qualquer desvio da normalidade da voz pode interferir na eficácia comunicativa do indivíduo. A aceitação das variações vocais dentro dos limites da normalidade depende de fatores culturais, do grau do impacto da qualidade vocal e da duração dos sintomas, sendo o ouvinte que determina os parâmetros que separam uma voz normal de uma voz alterada (Guimarães, 2007).

Na sociedade em que vivemos vários fatores podem ter influência na percepção da voz de um indivíduo. Uma das características mais vulneráveis ao julgamento por parte da sociedade é a adequação ao que é esperado para o sexo do indivíduo. Ou seja, falantes do género masculino ou feminino que apresentem traços típicos do género oposto, são comumente alvo de julgamento por parte da sociedade, gerando ruído no processo de comunicação. Indivíduos com edema de Reinke, com puberfonia e transgéneros são exemplos de casos que podem surgir na prática clínica do terapeuta da fala (TpF) cujas alterações vocais estão diretamente relacionadas com a percepção de género e que muitas vezes apresentam problemas psicoemocionais associados à voz, limitações na atividade e restrições na participação na sociedade.

Ao longo do tempo têm sido exploradas diferentes medidas e procedimentos de avaliação acústica, de forma a complementar a avaliação perceptiva habitualmente usada no meio clínico. Existem medidas de curto e longo termo desenhadas para caracterizar a voz e evidenciar aspetos relacionados com as alterações vocais. Embora vários estudos em todo o mundo utilizem medidas, procedimentos e *software* distintos, é importante perceber a sua robustez e pertinência, de forma a tornar a avaliação acústica mais simples e eficaz.

Para melhor compreender a voz na prática clínica é preponderante analisar as vozes normais e definir valores de referência, de forma a tornar a definição de normalidade associada à voz o mais objetiva possível para o Português Europeu (PE). Uma base de referências bem definida vem auxiliar também a atribuição de diagnósticos fundamentados na prática clínica, junto de vozes patológicas.

Na presente dissertação serão exploradas as formas de avaliação da voz, com principal destaque na avaliação acústica. Procura-se descrever e obter parâmetros que permitam caracterizar acusticamente a voz dos diferentes géneros e contribuir para os valores de referência do PE. Serão, também, explorados alguns quadros clínicos diretamente relacionados com a importância do género como forma de explicar a importância do estudo das vozes normais para prática clínica.

2. Enquadramento Teórico

A voz é o som produzido pela vibração das pregas vocais à passagem do ar pulmonar e pela modificação do trato vocal, caracterizado por parâmetros acústicos e perceptivos. Permite a comunicação humana por via da fala (voz falada), canto (voz cantada) ou de outras formas de expressão não verbal. É um importante biomarcador de identidade pessoal que permite expressar ideias, emoções e intenções variadas, pelo que se revela um dos mais importantes recursos para as atividades e participação a nível pessoal, social, profissional e artístico (SPTF, 2020).

A voz normal, assim como qualquer conceito ao qual é associada a designação “normal”, estando relacionada com o comportamento humano, é necessariamente definida de forma subjetiva, pois é influenciada por fatores civilizacionais e interpretações ideológicas (Behlau, 2004). A definição de voz normal depende da qualidade e adequação da voz, no entanto estas características são avaliadas consoante o gosto pessoal do ouvinte, tendo por base fatores sexuais, raciais e culturais (Guimarães, 2007).

Na prática clínica a voz pode ser avaliada tendo em conta as características fisiológicas, perceptivas e acústicas. Fisiologicamente, uma voz pode ser considerada normal quando existe um equilíbrio entre diferentes fatores tais como: a pressão e a velocidade do fluxo de ar expiratório, os padrões de vibração das pregas vocais e a configuração das estruturas do trato vocal (Guimarães, 2007). Perceptivamente pode ser avaliada tendo em conta a agradabilidade: se a altura tonal da voz¹ está adaptada ao género e à idade; se a intensidade vocal subjetiva² está apropriada ao contexto, permitindo, ao indivíduo, ser ouvido pelos interlocutores e não chamar atenção indesejada por ser demasiado elevada; e se é flexível ao conteúdo da informação que está a ser transmitida (Behlau, 2004; Guimarães, 2007), questões estas relacionadas, também, com a adequação da prosódia. Em termos acústicos é possível avaliar a voz a partir do sinal sonoro de forma não invasiva e determinar e quantificar a qualidade vocal de cada indivíduo, através dos diferentes parâmetros acústicos que compõem o sinal: periodicidade, amplitude, duração e composição espectral (Guimarães, 2007; Teixeira et al., 2011).

De uma forma mais objetiva a voz pode ser definida como um som audível resultante da atividade laríngea, através de uma relação complexa entre pressão e velocidade do fluxo de ar expiratório, os diferentes padrões de adução e de abdução das pregas vocais e as propriedades de reflexão e configuração do trato vocal (Linville, 2002).

¹ A altura tonal da voz ou *pitch*, é o “correlato audioperceptivo da frequência fundamental. É um parâmetro psicoacústico de avaliação da voz, cuja medida é determinada numa escala de grave a agudo. Para além da frequência fundamental do som vocal ... a altura tonal é influenciada por outros aspetos do espetro sonoro, nomeadamente pelos fenómenos de ressonância.” (SPTF, 2020, p. 34)

² A intensidade vocal subjetiva ou *loudness* é a “percepção auditiva de amplitude, altura de uma onda sonora, determinada pela quantidade e velocidade da pressão do ar subglótico ... é medida numa escala de fraco a forte.” (SPTF, 2020, p. 120)

Do ponto de vista fisiológico, as pregas vocais ao vibrarem, por ação da pressão subglótica exercida pelo ar proveniente dos pulmões, dão origem a uma onda sonora. Esta onda irá ter determinada frequência e intensidade de vibração consoante as características fisiológicas de cada indivíduo (Mateus, Falé, & Freitas, 2005).

A onda sonora produzida na glote, antes de chegar ao exterior passa, também, por um conjunto de espaços do trato vocal que se encontram acima das pregas vocais, designadas de cavidades supraglóticas. Estas cavidades (laringe, faringe, nariz e boca) produzem efeitos sobre a onda sonora que modificam a ressonância da voz, conferindo qualidades únicas que permitem identificar a voz dos diferentes indivíduos (Mateus et al., 2005).

Embora seja pertinente obter resultados objetivos acerca da voz, é necessário contemplar, também, o ponto de vista do paciente sobre a sua voz e as suas implicações profissionais e sociais (Putnoki, Hara, Oliveira, & Behlau, 2010).

A voz pode ser percecionada de diversas formas, consoante a cultura em que os indivíduos se inserem. Quando a voz foge ao paradigma de uma determinada sociedade pode criar efeitos negativos na qualidade de vida, na relação psicossocial, pessoal e ou profissional do indivíduo. Quando procuramos objetivar a avaliação da voz, pretendemos obter auxílio na perceção da alteração registada, no entanto estas avaliações não refletem a dimensão do impacto na vida do indivíduo. O Voice Handicap Index (VHI) (Jacobson et al., 1997), o Voice-Related Quality of Life (V-RQOL) (Hogikyan & Sethuraman, 1999); e o Voice Symptom Scale (VoiSS) (Deary, Wilson, Carding, & MacKenzie, 2003), traduzido, adaptado e validado para PE (Santos, 2012), são exemplos de instrumentos que permitem medir o impacto psicossocial da voz.

A adequação da voz ao género do falante pode ser considerada um dos fatores com mais impacto ao nível psicossocial, uma vez que gera sentimentos de inadequação, em comparação com os restantes indivíduos do mesmo género (Hancock, 2015; Azul, 2015, citado por Barros, 2017).

O Tpf é o profissional responsável por intervir nas perturbações da voz tanto de origem orgânica como funcional (Behlau, 2004). O edema de Reinke, a muda vocal incompleta e os indivíduos transgéneros são exemplos de perturbações da voz que causam impacto psicossocial relacionado com o género (Barros, 2017). Os objetivos de intervenção, nestes casos, devem sempre ser dirigidos às expectativas do indivíduo, de forma a que este se sinta confortável com a sua expressão de género nas interações e meios sociais, e não do ponto de vista de um modelo de masculinidade ou feminilidade previamente concebido (Barros, 2017).

Uma avaliação vocal completa na prática clínica deve incluir uma vasta recolha de informação através da entrevista clínica, uma avaliação por parte de um médico otorrinolaringologista ao nível da fisiologia

laríngea, uma avaliação musculoesquelética, uma avaliação perceptiva que englobe todas as componentes da voz, assim como a componente aerodinâmica, a avaliação acústica e uma autoavaliação do impacto psicossocial da voz no indivíduo (Belo, 2015).

No mesmo sentido, Behlau (2004) defende que na avaliação vocal devem ser contempladas questões que englobem as dimensões biológica, psicológica e socioeducacional. Segundo a dimensão biológica podemos obter informações acerca do sexo, idade, saúde geral, estrutura física geral e específica dos órgãos do aparelho fonador. Na avaliação das estruturas deve ser englobada não só a estrutura, mas também as questões funcionais e a forma como trabalham entre si e com as restantes estruturas do sistema estomatognático. A dimensão psicológica avalia a personalidade e o estado emocional do indivíduo no momento da emissão vocal. A dimensão socioeducacional engloba o sotaque, os regionalismos, modelos vocais considerados mais positivos ou negativos consoante a cultura e consoante a profissão (Behlau, 2004).

A avaliação áudio-perceptiva é realizada através da audição da voz do paciente. Está dependente da experiência do avaliador, dos critérios utilizados e das suas competências de percepção e discriminação auditiva (Behlau, 2004). Embora existam escalas para tornar a avaliação perceptiva mais objetiva, continua a ser preponderante a utilização da análise acústica para complementar a informação recolhida auditivamente (Behlau, 2004; Felipe, Grillo, & Grechi, 2006).

A análise acústica é uma forma de avaliação não invasiva, com o objetivo de determinar e quantificar a voz, através dos parâmetros acústicos que compõem o sinal de fala tais como a periodicidade, a amplitude, a duração e a composição espectral (Guimarães, 2007). É uma forma de obter documentação mais precisa para realizar a linha de base da voz de um indivíduo; verificar a eficácia de um determinado tratamento, durante e após a sua aplicação; acompanhar o desenvolvimento de uma voz profissional ao longo do tempo; e pode ser útil, também, como instrumento de deteção de alterações vocais (Behlau, 2004; Felipe et al., 2006).

«A análise acústica é o estudo de caracterização das propriedades acústicas dos sons da fala envolvendo aspetos como a frequência, amplitude e tempo. Permite determinar o número, a frequência e a amplitude (intensidade) das vibrações que constituem um som complexo. Os parâmetros vocais acústicos de maior importância para o uso clínico são as medidas de ruído, perfil de extensão vocal, espectrografia acústica e a frequência fundamental e os seus índices de perturbação – *jitter* e *shimmer*. Existem ainda outras características do sinal sonoro capazes de fornecer informações importantes, tais como os formantes e o tempo máximo de fonação.»
(SPTF, 2020, p. 35)

A análise acústica permite extrair, através de um equipamento de recolha de som, a informação proveniente das ondas sonoras que têm origem na fonte glótica com a vibração das pregas vocais. As

ondas sonoras, antes de se propagarem no meio exterior, passam pelo trato vocal, que as modifica. Quando as ondas são captadas pelo equipamento de recolha de som apresentam certas configurações que permitem obter um conjunto de dados que refletem a qualidade vocal e a ressonância da voz do indivíduo (Mateus et al., 2005).

Uma vez que é uma avaliação que está dependente de equipamentos tecnológicos e *software* específicos, a fiabilidade da informação torna-se dependente das condições e procedimentos de gravação, armazenamento e análise do sinal, razão pela qual estas questões devem ser levadas a cabo com rigor sempre que é implementada uma análise acústica (Guimarães, 2007).

Embora seja uma avaliação que permite, quantificar a voz do indivíduo e obter dados mais objetivos, a utilização da avaliação acústica em meio clínico é restrita, pois exige a aquisição de equipamentos tecnológicos específicos para gravação e análise dos dados, alguns dos procedimentos de recolha e análise são complexos e o tempo e a tipologia das amostras podem não corresponder à fala espontânea do indivíduo (Guimarães, 2007).

Nos capítulos que se seguem serão aprofundados os métodos de avaliação da voz, tanto perceptivos como acústicos, bem como o estado da arte para a caracterização do género, com base nesses mesmos métodos. Será também apresentada uma breve descrição de alguns quadros clínicos que estão diretamente relacionados com a caracterização de género, seguindo-se a descrição dos objetivos para o presente projeto.

2.1. Avaliação da voz

2.1.1. Avaliação Perceptiva da Voz

A avaliação perceptiva, para ser o mais completa possível, deve incluir uma avaliação detalhada de cada sistema, para tal, devem ser considerados vários aspetos do comportamento vocal, tais como qualidade vocal, sistema de ressonância, coordenação pneumofonoarticulatória, assim como a avaliação corporal e a psicodinâmica vocal (Behlau, 2004; Boone, 2010 citado por Belo, 2015).

De forma a avaliar apenas a componente auditiva da produção vocal é utilizada a avaliação áudio-perceptiva da voz, que consiste no “processo de recolha de informação audioperceptiva realizada por um examinador sobre o estado corrente da voz do cliente, na qual se procura descrever o comportamento de parâmetros audioperceptivos diversos, nomeadamente: respiração, fonação, altura tonal da voz, intensidade subjetiva, ressonância e débito.” (SPTF, 2020, p. 46)

Na respiração deve ser avaliado o tipo e o modo respiratório através da observação do padrão respiratório, bem como a coordenação pneumofonoarticulatória do falante (Behlau, 2004).

Ao avaliar a qualidade vocal o avaliador deve ter em conta as características vocais do indivíduo do ponto de vista de qualidade do som ouvido. Devido à subjetividade desta avaliação existe uma grande variedade de conceitos que permitem classificar a voz. Algumas qualidades estão mais relacionadas com a voz patológica, tais como: rouca, áspera, soprosa, sussurrada, tensa; enquanto que outros aspetos estão mais relacionados com o desenvolvimento normal da voz: infantilizada, presbifónica, feminilizada ou virilizada (Behlau, 2004).

A altura tonal da voz, tal como já foi referido, refere-se à perceção auditiva da frequência (SPTF, 2020). Os registos vocais estão diretamente relacionados com a altura tonal, pois, através de diferentes configurações laríngeas, permitem ao indivíduo produzir uma voz mais grave ou mais aguda. Os registos vocais são mais conhecidos pelas suas descrições baseadas na perceção auditiva ou configuração laríngea, no entanto tendo em conta o pressuposto deste trabalho é importante abordar esta questão com uma perspetiva baseada em parâmetros acústicos. Os mecanismos vibratórios da laringe, embora não sejam diretamente correspondentes aos registos vocais, representam as diferentes configurações da glote que permitem produzir todas as frequências da extensão vocal dos indivíduos. Existem 4 mecanismos que variam entre M0 (Mecanismo 0) e M3 (Mecanismo 3): o M0, relacionado com o registo basal, ou *vocal fry*, é o mecanismo que permite ao ser humano produzir as frequências mais graves, marcado fisiologicamente por cordas vocais curtas, grossas e relaxadas, com uma fase de fechamento mais longa; no M1, relacionado com o registo modal de peito, há maior massa de vibração com intervenção das camadas mais profundas das pregas vocais, mais amplitude de vibração e uma fase de fechamento abrupta e de curta duração; no M2, relacionado com o registo modal de cabeça, a massa de vibração é muito menor, vibrando apenas a parte mais superficial das pregas vocais e, por consequência, com menor amplitude; quanto ao M3, relacionado com o registo de falsete, as cordas encontram-se mais estreitas e alongadas, permitindo produzir frequências mais agudas. Durante a fala os indivíduos utilizam predominantemente os mecanismos M1 e M2 (Roubeau, Henrich, & Castellengo, 2009).

O sistema de ressonância é o conjunto de elementos do aparelho fonador, que entre si, são capazes de modelar e projetar a voz no espaço. Estes elementos são designados cavidades de ressonância (laringe, faringe, cavidade vocal, cavidade nasal e seios perinasais) responsáveis por dar ênfase a certas frequências em detrimento de outras. Estas modificações de ampliação/redução de frequências, podem ocorrer de forma intrínseca relacionados com aspetos fisiológicos do aparelho fonador e com as opções de uso do falante, ou de forma extrínseca dependente do ambiente em que o falante está inserido (Behlau, 2004).

A intensidade vocal subjetiva está relacionada com a perceção da amplitude (SPTF, 2020). Esta é alterada consoante as necessidades e o contexto em que o indivíduo se encontra. A adequação da intensidade vocal exige a consciência da dimensão do interlocutor, bem como o controlo da projeção

vocal. Num contexto de exposição perante um grande público, por exemplo, a intensidade vocal usada será superior à utilizada numa conversa intimista entre duas pessoas próximas (Behlau, 2004).

O débito de fala corresponde ao número de palavras usado por unidade de tempo (SPTF, 2020). Se este débito estiver alterado, ao ponto de apresentar uma velocidade de discurso muito rápida ou muito lenta, pode comprometer a eficácia na transmissão da mensagem (Behlau, 2004).

De forma a tornar a avaliação áudio-percetiva mais objetiva foram criadas diferentes escalas que permitem ao avaliador quantificar a voz do indivíduo. A GRBAS, construída por Hirano em 1981, é um instrumento de análise áudio-percetiva da voz falada que avalia: a gravidade global (Grade), rouquidão (Roughness), sopro (Breathiness), astenia (Astheny) e tensão (Strain), traduzido, adaptado e validado para o PE por Freitas (2012), Guimaraes (2002) e Silvestre, Guimarães, & Teixeira (2011). Pinho e Pontes (2002) também apresentaram uma adaptação da escala (RASAT), na qual propuseram a distinção entre rouquidão e aspereza (Pinho & Pontes, 2002) à qual, mais tarde foi associado o parâmetro Instabilidade (Pinho & Pontes, 2008). Em 2006 surgiu, também a CAPE-V (ASHA, 2006), traduzida, adaptada e validada para o PE por Jesus, Tavares e Hall em 2017 e em 2019 por Almeida, Mendes e Kempster, apresentando uma segunda versão da escala, que permite realizar uma avaliação mais generalizada da voz, incluindo, avaliar não só parâmetros de qualidade vocal, mas também, recolher informações acerca de altura tonal, intensidade subjetiva e da contribuição do trato vocal ao nível da ressonância (Almeida, Mendes, & Kempster, 2019; Jesus, Tavares, & Hall, 2017).

Para avaliar a voz são, geralmente, solicitadas diferentes tarefas que permitem analisar os comportamentos vocais. Estas tarefas podem ser mais próximas da utilização normal da voz, tais como: discurso espontâneo, leitura e o canto; ou tarefas mais específicas, como: fricativas e vogais sustentadas e escalas ou glissandos (Mendes & Castro, 2005).

2.1.2. Avaliação Aerodinâmica da Voz

A avaliação aerodinâmica da voz consiste em avaliar “os volumes, o fluxo e a pressão do ar envolvido na produção de voz” (SPTF, 2020, p. 45). As medidas aerodinâmicas são representativas da atividade laríngea, nomeadamente da adução e abdução das pregas vocais. O volume de ar, o fluxo de ar e a pressão de ar são as medidas mais frequentemente extraídas e permitem obter dados indiretos acerca da biomecânica fonatória (Baken & Orlikoff, 2000).

O Tempo Máximo de Fonação (TMF), o coeficiente S/Z e o padrão respiratório são os parâmetros mais utilizados para obter dados acerca do comportamento e coordenação entre respiração e fonação (eficácia fonorrespiratória), uma vez que não requerem a utilização de equipamentos tecnológicos

(Santos, 2009). O TMF e o coeficiente S/Z são medidas que permitem avaliar a eficiência glótica³ através da contabilização do tempo máximo que o indivíduo aguenta a sustentar determinado som (SPTF, 2020).

O TMF é uma medida de tempo, expressa em segundos, que pode ser obtida através da emissão da vogal [a], [i] ou [u], com a máxima duração e intensidade confortável (SPTF, 2020). Uma vez que reflete o controlo das forças aerodinâmicas e mioelásticas da laringe (Guimarães 2007), representa o controlo da função respiratória, eficiência glótica e atividade valvular da laringe. Esta medida deve ser obtida através da média do tempo de emissão máximo das vogais em duas ou três repetições (Sridhara, Ashok et al. 2003, citado por Santos, 2009).

O coeficiente S/Z, proposto por Eckel e Boon em 1981, avalia o suporte aéreo pulmonar possibilitando a avaliação da fonte friccional do som e da fonte glótica através da emissão prolongada dos fonemas [s] e [z] após inspiração profunda (Boone, 1971, citado por Behlau, 2004; SPTF, 2020). A produção da consoante fricativa [s] reflete a capacidade de controlo expiratório e a fricativa [z] acrescenta à tarefa a componente laríngea (vozeamento). Para a obtenção deste coeficiente é necessária a produção sustentada e prolongada das consoantes [s] e [z] no máximo período de tempo em duas ou três tentativas. Os valores mais elevados obtidos na produção destas duas fricativas são utilizados para a obtenção deste coeficiente (Santos, 2009).

2.1.3. Análise Acústica – Frequência Fundamental

A frequência mais baixa de uma onda sonora complexa designa-se frequência fundamental (f_0). Esta corresponde diretamente à frequência vibratória das pregas vocais e é medida em Hertz (Hz) (Beber & Cielo, 2011) ou em semitons (ST) (Guimarães, 2002). Esta frequência pode também ser designada de primeiro harmónico (H1), sendo que o valor da frequência dos restantes harmónicos é sempre múltiplo do primeiro (Mateus et al., 2005). A f_0 reflete as características biodinâmicas das pregas vocais e a sua integração com a pressão subglótica. Está relacionada com o comprimento, a tensão, a rigidez e a massa das pregas vocais na sua interação com a pressão aérea subglótica (Beber & Cielo, 2011). Quanto maior a massa e menos alongada estiver a prega vocal mais lentos serão os ciclos e mais baixa será a frequência produzida, criando sons mais graves. No sentido inverso, quanto maior a tensão nas pregas vocais, mais rápidos os ciclos e mais elevada a frequência gerada, produzindo sons agudos (Behlau, 2004).

Na avaliação da voz o valor da f_0 obtém-se tanto em vogais sustentadas como em fala encadeada. Pode ser obtido o valor da f_0 média, o desvio padrão (DP), a f_0 mínima, f_0 máxima e Gama Tonal. O DP

³ Eficiência glótica – capacidade de controlo das forças aerodinâmicas do fluxo de ar e mioelásticas da laringe (SPTF, 2020)

corresponde ao desvio acima e abaixo do valor da f_0 média e a Gama Tonal corresponde ao número de ST que o indivíduo usa durante a fala calculado entre a f_0 mínima e f_0 máxima, correspondendo em ambos os casos à variabilidade da f_0 . A extensão vocal corresponde ao número de ST (ou notas musicais) que um indivíduo é capaz de emitir, do mais grave até ao mais agudo, não importando a qualidade vocal nem o esforço necessário para a conseguir (Behlau, 2004).

A f_0 é uma medida que permite uma ampla variabilidade de investigações, nomeadamente estudos relacionados com a sua variação na idade, no género, na raça, em diferentes contextos, tarefas, línguas, hábitos de higiene vocal, no uso profissional da voz e em casos de disfonia. Tem como característica o facto de ser consistente que possibilita a utilização de diferentes instrumentos de recolha e de tratamento de dados (Guimarães, 2002).

2.1.4. Análise acústica – Medidas de Curto Termo

As medidas de perturbação da frequência e intensidade, que são frequentemente extraídas na avaliação acústica, são as medidas de curta duração *jitter* e *shimmer*. Estas permitem, através da análise dos ciclos vibratórios de vogais sustentadas, obter informação acerca da frequência e amplitude da voz e perceber o quanto um período da onda sonora se diferencia de outro que o sucede (Araújo, Grellet, Pereira, & Rosa, 2002; Beber & Cielo, 2011; Behlau, 2004; Felipe et al., 2006; Freitas, 2012; Rachel, Snehalatha, Vedhasorubini, & Balakrishnan, 2017).

O *jitter* é uma medida que considera a variabilidade da f_0 média ciclo a ciclo, enquanto que o *shimmer* indica a variabilidade da amplitude da onda sonora ciclo a ciclo (Araújo et al., 2002; Baken & Orlikoff, 2000; Behlau, 2004; Titze, 1995, citado por Freitas, 2012; Pinho, 2006; Rachel et al., 2017). As duas medidas só podem ser obtidas durante a emissão de uma vogal sustentada, sem variações voluntárias de frequência ou de intensidade (Behlau, 2004; Dejonckere et al., 2012, citado por Freitas, 2012). Estas medidas de perturbação podem variar com a escolha da vogal, a idade e o género do sujeito (Beber & Cielo, 2011; Behlau, 2004; Freitas, 2012).

Mesmo as vozes normais podem apresentar uma pequena variação na massa ou na tensão das pregas vocais, na distribuição do muco sobre as mesmas, na simetria das estruturas ou ainda na atividade muscular ou neural envolvida. Estas alterações refletem-se no valor do *jitter*, que se correlaciona mais com o parâmetro percetivo Aspereza e/ou Rouquidão (Lieberman, 1961; Koike, 1967; Yumoto et al., 1984; Wolfe & Steinfatt, 1987; Dejonckere et al., 1993; Buder, 2000; Deal e Emanuel, 1978 in Yiu et al., 2000; Titze, 1995; Ma & Yiu, 2006, citados por Freitas, 2012), com o Grau geral de disfonia (Ma & Yiu, 2006; Martens et al., 2007, citados por Freitas, 2012) e a Soprosidade (Eskenazi et al., 1990; Dejonckere, 1995, citados por Freitas, 2012).

Por mais estável que seja a voz do indivíduo ocorrem sempre pequenas variações no fluxo do ar e na intensidade que se refletem no *shimmer*. Esta medida, normalmente, apresenta valores maiores em sons com frequência grave e intensidade fraca. Oferece uma percepção indireta do ruído na produção vocal e relaciona-se com o Grau geral de disфонia (*grade*) (Martens et al., 2007, citado por Freitas, 2012) e com a Soprosidade (Takahashi & Koike, 1976; Feijoo & Hernández, 1990; Dejonckere et al., 1993, citados por Freitas, 2012), ou com a Rugosidade (Hillenbrand, 1988; Feijoo & Hernández, 1990; Tittze, 1995; Eadie & Baylor, 2006; Martens et al., 2007, citados por Freitas, 2012);

Alguns autores assumem que as medidas de variação de frequência e amplitude, respetivamente o *jitter* e o *shimmer*, na emissão de vogais sustentadas são úteis na descrição das características vocais de falantes normais e disfónicos (Felippe et al., 2006). Contudo, outros autores referem que, para vozes disfónicas, o *jitter* e o *shimmer* não são as melhores opções de análise, pois são medidas que nos conferem os limites de cada ciclo vibratório, o que se torna difícil de mensurar em vozes muito alteradas (Watts & Awan, 2011).

O *Harmonic to Noise Ratio* (HNR) é uma medida acústica de curto termo que contrasta o sinal regular (componente harmónico) com o sinal irregular das pregas vocais e do trato vocal (componente ruído) e é quantificada em dB. As medidas de ruído, quantificam o ruído originado na turbulência do ar ao nível da glote (Rachel et al., 2017). Quanto mais harmónica for a voz e menos ruído possuir maior o valor de HNR (Beber & Cielo, 2011). Esta medida é um preditor significativo de amostras de voz que são classificadas perceptivamente como roucas (Martin, Fitch & Wolfe, 1995, Eskenazi, Childers & Hicks, 1990, Krom G, 1995, citado por Ferrand, 2002). O HNR é mais sensível à função vocal em comparação com *jitter* (Ferrand, 2002).

A análise acústica de curto termo requer a recolha precisa dos ciclos de f_0 , que é praticamente impossível de obter em amostras de discurso espontâneo, razão pela qual são usadas as vogais sustentadas. Apesar de estas apresentarem como ponto positivo a exclusão das variações provocadas pela articulação e variações consoante o dialeto, elas não são representativas do discurso espontâneo dos indivíduos. Por este motivo, têm vindo a ser exploradas medidas que permitem obter informação acústica de tarefas de fala encadeada (Delgado-Hernández, León-Gómez, Izquierdo-Arteaga, & Llanos-Fumero, 2018).

A partir de tarefas de curto termo também é possível realizar análises do ponto de vista espectral que serão referidas no próximo capítulo.

2.1.5. Análise acústica – Medidas de Longo Termo

Um dos instrumentos mais usados na avaliação acústica é a espectrografia que demonstra visualmente as características acústicas da emissão sendo, também, um meio complementar da

análise perceptivo-auditiva (Pontes, Vieira, Gonçalves, & Pontes, 2002). A espectrografia pode ser definida como a transformação de padrões sonoros em padrões visuais gráficos utilizando a Transformada de Fourier. Este cálculo matemático converte a forma de onda (amplitude por tempo) no espectro (frequência por amplitude), oferecendo a amplitude das diversas frequências que compõem o som. Como as ondas de fala não são periódicas e sim quase periódicas, é necessário fazer uma adaptação da análise de Fourier através da Transformada Rápida de Fourier (FFT), produzindo um espectro que mostra a amplitude de cada harmónico da f_0 (a diferença entre dois harmónicos é a f_0) (Behlau, 2004). Esta técnica reflete as características da voz, a f_0 e os harmónicos correspondentes, a amplitude de cada harmónico e a duração do registo vocal (Guimarães, 2007).

As restantes partes do trato vocal, como a faringe, a cavidade oral e a nasal, também têm a sua representação nos espectros, alterando certas frequências que determinam a origem de formantes, isto é, grupos de frequências que recebem reforço de intensidade por ação do trato vocal (Beber & Cielo, 2011).

A espectrografia acústica é o método ideal para a análise da fala encadeada, pois evidencia aspetos temporais da emissão, características de coarticulação dos sons da fala, a contribuição da fonte glótica e a ressonância do trato vocal (Behlau, 2004).

De entre várias formas possíveis de representação da onda sonora, existe uma que se baseia na análise espectral, designada por LTAS (*Long-term Average Spectrum*), em português, espectro médio de longa duração. O LTAS tem vindo a ser utilizado em diferentes estudos pois permite avaliar a qualidade da voz, sendo possível observar diferenças entre género (Mendoza, Valencia, Muñoz, & Trujillo, 1996), idade (Linville, 2002), qualidade da voz profissional (Cleveland, Sundberg, & Stone, 2001; Master, De Biase, Chiari, & Laukkanen, 2008) e aspetos relacionados com vozes disfónicas, contribuindo para a avaliação e para posterior intervenção (Da Silva, Master, Andreoni, Pontes, & Ramos, 2011; Master, Biase, Pedrosa, & Chiari, 2006).

Segundo Nordenberg e Sundberg (2003) o LTAS “reflete a contribuição tanto da fonte glótica como do trato vocal na qualidade de uma voz” e, num só espectro, é possível observar a média dos vários espectros de um segmento de fala de longa duração (Nordenberg & Sundberg, 2003).

Ao contrário das avaliações que envolvem análise acústica de curto termo, o LTAS avalia a voz ao longo do discurso, tornando-se irrelevante a periodicidade do sinal. Este fator é importante do ponto de vista clínico uma vez que a maior parte das vozes encontradas no meio clínico são normalmente aperiódicas (Tanner, Roy, Ash, & Buder, 2005). O LTAS anula a variabilidade acústica proveniente da instabilidade articulatória do trato vocal, o que é essencial para se conseguir estudar a fonte sonora e obter mais destaque na região representativa das variações de ressonância (Ng, Liu, Zhao, & Lam, 2009).

Para além da análise visual do próprio traçado, que, por si só, admite retirar algumas conclusões, o LTAS permite obter valores exatos de medidas representativas do comportamento da voz, tais como: α , $(H1-H2)_{LTAS}$ e Leq .

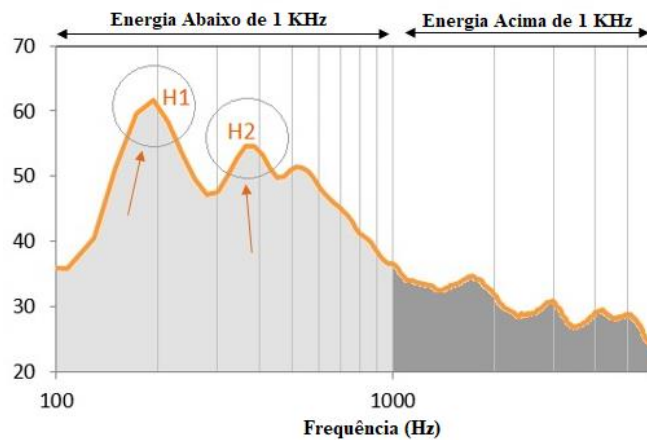


Figura 1- Representação visual dos componentes usados no cálculo do α e do $(H1-H2)_{LTAS}$. Adaptado de Araújo (2015)

O α , ou α ratio, compara a energia entre os 0 e os 1000 Hz com a energia entre os 1000 e os 5000 Hz (Araújo, 2015; Da Silva et al., 2011; Ng et al., 2009; Xie et al., 2016). Esta medida reflete a inclinação do espectro estando relacionada com a velocidade de encerramento glótico. A forma do fluxo de ar durante a passagem na glote influencia a amplitude dos primeiros harmônicos. As tarefas em intensidade normal apresentam uma energia alta inferior a 1000 Hz, produzindo um α elevado. O contrário pode ser observado em tarefas de intensidade forte, que, por norma, são realizadas com maior nível de adução (e conseqüente tensão), provocando a criação de uma maior energia acima dos 1000 Hz e, portanto, um α com valores mais reduzidos. Esta inclinação pode variar, também, com diferenças de intensidade e os diferentes tipos de qualidade vocal (Da Silva et al., 2011; Ng et al., 2009; Xie et al., 2016).

O $(H1-H2)_{LTAS}$ compara a energia no primeiro harmônico (ou f_0) com a energia no segundo harmônico (H2) (Araújo, 2015; Sundberg, Salomão, & Scherer, 2019). Um $(H1-H2)_{LTAS}$ elevado apresenta um H1 superior a um H2, representativo de uma voz mais hipofuncional, soprosa e fraca (como representado na figura 1). Pelo contrário um valor de $(H1-H2)_{LTAS}$ baixo apresenta um H2 com energia superior, estando relacionado com uma voz hiperfuncional, tensa e forte, podendo inclusivamente apresentar valores negativos ($H2 > H1$) (Araújo, 2015; Da Silva et al., 2011).

O Equivalent Sound Level (Leq), usualmente medido em dB a 30 cm da boca, corresponde ao nível constante de pressão sonora ao longo do tempo, permitindo obter um valor médio de intensidade para tarefas de fala (Araújo, 2015; Da Silva et al., 2011).

O Cepstro, é o resultado da inversão do logaritmo de um determinado espectro, ou seja, o mesmo que aplicar duas transformadas de Fourier a uma amostra (Delgado-Hernández et al., 2018). O Cepstro filtra os componentes agudos do espectro, convertendo as ondas em valores numéricos, permitindo identificar os componentes de frequências e as suas amplitudes. A escala horizontal do cepstro representa a quefrência (Behlau, 2004).

A quefrência diz respeito à frequência da frequência do espectro, sendo a unidade de quefrência correspondente a ciclos por frequência. A quefrência uniformemente espaçada com picos de baixa amplitude são chamados de *rahmónicos*, derivados dos harmónicos. O pico com a maior amplitude é designado de pico cepstral. Quando se traça uma linha de regressão, no espectro, representativa da energia sonora média, que é obtido a partir do sinal de voz, a distância medida do pico do cepstro a esta linha de regressão linear é denominada como proeminência do pico cepstral (CPP – *Cepstral Peak Proeminence*) (Sujitha, 2020). Na prática o CPP reflete o quão sinusoidal é o espectro logarítmico no pico da f_0 , ou seja, quão acentuada é a curva correspondente ao pico da f_0 . A amplitude da CPP mede a periodicidade do espectro e não do sinal (Riesgo & Nöth, 2019) e é expressa em dB (Brockmann-Bauser et al., 2019).

Awan, Giovinco & Owens (2012) concluíram que a CPP é influenciada pelas alterações de intensidade e pelo tipo de vogais usadas. A alteração consoante a intensidade utilizada acontece pois, quanto mais intensidade tem a voz, menor é a perturbação encontrada e maior e mais estável é a frequência, o que reflete, também, maior tensão nas pregas vocais (Awan, Giovinco, & Owens, 2012).

Alguns estudos indicam que a CPP é eficaz na identificação de vozes disfónicas, nomeadamente por ser sensível à soproside (Brinca, Batista, Tavares, Gonçalves, & Moreno, 2014; Watts, Awan, & Maryn, 2017). Estando o CPP relacionado com a periodicidade seria de esperar alguma relação com o parâmetro de rouquidão, no entanto, autores que já estudaram este aspeto não o conseguiram comprovar (Heman-Ackah, Michael, & Goding, 2002).

A Proeminência do Pico Cepstral Suavizada (CPPS – *Smoothed Cepstral Peak Proeminence*) é uma atualização à CPP que foi desenvolvida por Hillenbrand et al. (1994) e Hillenbrand and Houde (1996) preparada para medir a severidade da disфонia, sendo possível relacionar com as medidas percetivas. A CPPS apresenta um passo adicional que suaviza o cepstro nos domínios temporal e espectral antes de calcular o pico, ou seja, calcula a média dos dados do sinal em determinado número de pontos, reduzindo dados desviantes (Brockmann-Bauser et al., 2019).

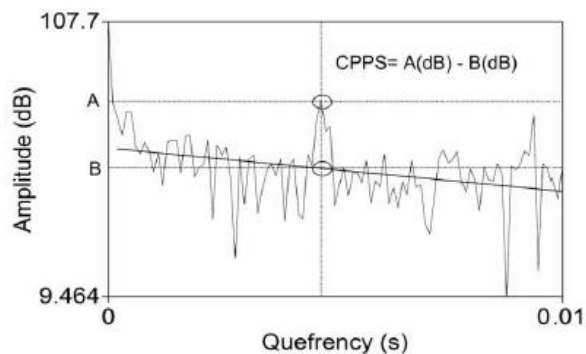


Figura 2 - Representação gráfica do Cepstro e do cálculo da CPPS (Delgado-Hernandez et al., 2018)

2.2. Caracterização do género vocal

Do ponto de vista da caracterização anatómica do género, é possível afirmar que, de forma geral, no género masculino se verifica uma posição vertical da laringe mais baixa. A proeminência da cartilagem laríngea pode até ser evidente no pescoço, ao qual, vulgarmente se designa de “maçã de adão”. Nas mulheres a laringe apresenta uma posição mais elevada, não sendo visível a olho nu. A laringe do género masculino é 40% mais comprida e mais alta do que a laringe feminina. A cartilagem tiroidea é 20% mais larga na dimensão antero-posterior no género feminino. Quanto às pregas vocais dos homens estas são longas e grossas, apresentando maior volume, enquanto que nas mulheres são mais curtas e estreitas (Brasil, Yamasaki, & Leão, 2005; Filho, Melo, Carneiro, Tsuji, & Sennes, 2003; Guimarães, 2002; Titze, 1989). Ao nível hormonal, as características mais associadas ao género masculino são provocadas pela presença de maiores níveis de testosterona, enquanto que características mais associadas ao género feminino estão relacionadas com a presença da hormona estrogénio (Mueller & Mazur, 1998, Andersson, 1994, citado por Feinberg, 2008).

Na voz do género masculino é frequente encontrar uma laringe baixa, pregas vocais grossas, com grande massa em vibração e superfície de contacto extensa, facilitando a emissão de sons graves (Behlau, 2004). Estas características parecem ser próprias do M1 descrito por Roubeau (2009) e resultado do predomínio da ação do músculo tirearitenóideo (Behlau, 2004). No género feminino a laringe está elevada, com pregas vocais alongadas, superfície de contacto reduzida e vibrações mais curtas, havendo uma maior ação do músculo cricotiróideo (Behlau, 2004). Nesta descrição de Behlau (2004) são detetadas mais semelhanças com o M2 (Roubeau et al., 2009). Embora seja possível associar os géneros a determinadas características de emissão vocal é normal existir variação de mecanismos ao longo do discurso (Roubeau et al., 2009), variação esta responsável pela expressividade (Behlau, 2004).

É importante lembrar que, embora se aborde o conceito voz normal, indivíduos com voz normal podem apresentar ligeiros desvios no que diz respeito a assimetrias e aperiodicidade durante a fonação,

provocados por pequenos edemas ou encerramento incompleto das pregas vocais (Bless et al, 1993 citado por Guimarães, 2002). Segundo um estudo realizado em mulheres jovens-adultas por Casper, Brewer and Colton (1987), citados por Guimaraes (2002), é comum encontrar variabilidade no tamanho, forma e simetria, principalmente no complexo aritenóideo durante a fonação (Guimarães, 2002). Nas mulheres também é normal encontrar fendas glóticas triangulares posteriores em vozes perceptivamente normais (Söderstern, Hertegard and Hammargerg, 1995, citado por Guimarães, 2002). Stemple, Sanley and Lee (1995) citados por Guimarães (2002), observaram, também que, após algum tempo em fala encadeada, as mulheres revelam uma pequena fenda na porção anterior da glote.

Nas medidas aerodinâmicas, tendo em conta os valores de TMF, o género masculino apresenta tempos superiores ao género feminino, encontrando-se entre 15 a 25 segundos nas mulheres e 25 a 35 nos homens (Behlau, 2004). Segundo alguns autores, os valores de referência esperados para a produção de vogais sustentadas nos homens adultos são 20 segundos e para as mulheres 15 segundos (Kent, Kent et al. 1987; Zemlin 1998, citado por Santos, 2009). Hirano (1981, citado por Santos, 2009), apresenta como valores de referência para os homens 23,7 segundos e para as mulheres 16,7 segundos.

Em relação ao coeficiente S/Z são considerados normais para os dois géneros, valores de 15 a 25 segundos de emissão prolongada e um valor de coeficiente mais próximo de 1, normalmente dentro do intervalo [0,8; 1,2]. Existe um maior quociente de encerramento das pregas vocais quando o valor da consoante [z] sustentada é maior do que a consoante [s]. Pelo contrário, quando o valor da consoante [s] prolongada é superior à consoante [z] é indicativo de menor coaptação glótica (Behlau, 2004).

Outros autores referem que um coeficiente S/Z de 1,0, com duração da produção das fricativas [s] e [z] de aproximadamente 20-25 segundos nos adultos, sugere uma capacidade respiratória normal e ausência de patologia. Caso seja inferior a 1,0 indica uma possível ineficácia respiratória, no entanto sendo superior a 1,2 pode ser indicador de perturbação relacionada com o funcionamento das pregas vocais, nomeadamente por falta de coaptação glótica (Pindzola, 1987, Eckel e Boone 1991 e Boone e MacFarlane 1994 citado por Santos, 2009).

A f_0 é a medida acústica que melhor reflete as diferenças entre género (Beber & Cielo, 2011). A f_0 é influenciada pela quantidade de testosterona presente nas últimas fases da puberdade, que determinam o tamanho da laringe e o comprimento das pregas vocais (Butler et al. 1989; Titze 1994; Harries et al. 1997, 1998). Após atingir a puberdade a f_0 dos homens fica mais grave, num período de tempo curto. No género feminino, a f_0 também fica mais grave, no entanto este procedimento decorre mais devagar em comparação com o género masculino, resultando em valores de f_0 mais elevados em comparação com o género masculino (Huber et al. 1999, citado por Feinberg, Jones, Little, Burt, & Perrett, 2005).

Nos indivíduos do género feminino a f_0 varia entre 150 e 350 Hz, enquanto que nos indivíduos do género masculino esta varia entre 80 e 200Hz (Mateus et al., 2005). Wuyts et al. (1996, citado por Freitas, 2012) definiram para homens gamas de frequência entre 78–166Hz e para as mulheres 167–258Hz.

De notar também que, de acordo com a tarefa de avaliação – vogal sustentada *versus* fala encadeada – os valores de f_0 são condicionados pelo género do falante – os homens baixam a f_0 na produção da vogal sustentada mais do que em conversação; as mulheres tendem ao inverso, isto é, aumentam a f_0 na sustentação de vogais (Freitas, 2012).

Quando medida em vogais sustentadas, a variabilidade (DP) da f_0 não ultrapassa os 2Hz em indivíduos normais. A variabilidade da f_0 na fala encadeada pode ser obtida através da leitura de um texto, conversação espontânea ou fala automática (Freitas, 2012). Nestes casos os indivíduos podem apresentar uma variação em conversação de 2 a 5 ST acima e abaixo da f_0 média, o correspondente entre 4 a 10 ST de Gama Tonal (Behlau, 2004; Mendes & Castro, 2005). Indivíduos com voz normal têm uma extensão vocal mínima superior a duas oitavas (24 ST). Esta extensão vocal desloca-se para os graves no edema de Reinke e para os agudos na muda vocal incompleta (Behlau, 2004).

Tabela 1- Valores de Referência de f_0

Artigos	Língua	Tarefa	Género	f_0 média (Hz) (DP)	f_0 DP (Hz)
(Spazzapan et al., 2018)*	Português (Brasil)	LN	H	128,52 (20,44)	
			M	206,84 (19,24)	
(Delgado, León, Jiménez, & Izquierdo, 2017)*	Espanhol (Espanha)	Va	H	109,99 (15,08)	1,20 (0,48)
			M	194,94 (23,44)	1,65 (0,69)
(Brockmann, Drinnan, Storck, & Carding, 2011)*	Suiço-Alemão e Alemão (Suíça)	Va	H	127 (9)	
			M	214 (9)	
(Ambreen, Bashir, Tarar, & Kausar, 2019)*	Urdu (Paquistão)	Va	H	130,08	
			M	231,8	
(Mazzetto et al, 2014)*	Português (Brasil)	Va	M	215,47	
(Mezzedimi, Di Francesco, Livi, Spinosi, & De Felice, 2017)*	Italiano	Va	H	141,64 (23,38)	
			M	217,54 (28,97)	
(Belo, 2015)*	Português (Portugal)	Va	H	109,5 (16,6)	
			M	198,1 (26,9)	
(Guimarães & Abberton, 2005)**	Português (Portugal)	Va	H	121.0 (6.2)	
			M	212.5 (6.6)	
		Vi	H	135.4 (13.1)	
			M	225.5 (5.5)	
		Vu	H	128.2 (6.7)	
			M	230.1 (6.5)	
		LN	H	116.1 (5.9)	
			M	196.9 (3.8)	
		DE	H	112.7 (4.4)	
			M	189.9 (4.2)	

Legenda: M – Mulheres; H – Homens; f_0 – frequência fundamental; Va – Vogal [a] sustentada; Vi – Vogal [i] sustentada; Vu – Vogal [u] sustentada; LN – Leitura Normal; DE- Discurso Espontâneo; *software: Praat; **software: Speech Studio

Tendo em conta as medidas de perturbação da f_0 , *jitter* e *shimmer*, a maioria das referências encontradas, que podem ser observadas na Tabela 2, apresentam valores maiores nos homens, no entanto esta informação não é consistente na literatura, além de que as diferenças encontradas nem sempre são estatisticamente significativas (Beber & Cielo, 2011; Freitas, 2012). Beber e Cielo (2011) defendem que o *jitter* pode ser mais elevado nos homens, por apresentarem pregas mais longas, que por sua vez podem gerar mais instabilidade na vibração (Beber & Cielo, 2011). Já Freitas (2012) faz referência a estudos que defendem que o *jitter* é maior no género masculino e o *shimmer* maior no género feminino (Albertini et al., 2009; Moon et al., 2012; citados por Freitas, 2012)

Como é possível observar na Tabela 2, o HNR apresenta valores mais baixos nos homens em comparação com as mulheres. O facto de o *jitter* e o *shimmer* apresentarem valores mais elevados nos homens, pode representar um maior nível de ruído na voz. No entanto, como na maioria dos artigos esta diferença não é estatisticamente significativa não foi possível justificar esta relação.

Tabela 2- Valores de Referência para as Medidas de Curto Termo obtidos através do software Praat® para a vogal [a]

Artigos	Língua	Género	<i>Jitter</i> (%)	<i>Shimmer</i> (%)	HNR (dB)
(Delgado et al., 2017)	Espanhol (Espanha)	H	0,49 (0,19)	2,38 (1,10)	22,94(3,23)
		M	0,41 (0,17)	1,48 (0,54)	25,63 (3,76)
(Brockmann et al., 2011)	Suiço-Alemão e Alemão (Suiça)	H	0,30 (0,05)		
		M	0,37 (0,07)		
(Ambreen et al., 2019)	Urdu (Paquistão)	H	0,23 (0,08)	2,40 (0,77)	23,39 (3,18)
		M	0,20 (0,07)	2,20 (1,06)	26,47 (3,19)
(Mazzetto de Menezes et al., 2014)	Português (Brasil)	M	0,4	3,8	19,54
(Mezzedimi et al., 2017)	Italiano	H	0,28 (0,10)	3,25 (1,08)	23,48
		M	0,28 (0,08)	3,12 (0,85)	22,79
(Belo, 2015)	Português (Portugal)	H			16,58 (3,33)
		M			18,16 (3,96)
(Teixeira & Fernandes, 2014)	Alemão	H	0,36 (0,18)	2,29 (1,26)	24,01 (4,37)
		M	0,33 (0,17)	2,75 (2,31)	24,95(4,48)

Legenda: H- Homens; M - Mulheres; *Jitter* – jitter local (%); *Shimmer* – shimmer local (%); HNR – Harmonic to Noise Ratio (dB)

Existem diversos estudos realizados no âmbito da aplicação do LTAS na voz normal, sendo uma delas a distinção entre géneros (Master et al., 2006). Nos estudos que envolvem a identificação do género verifica-se uma relação entre a qualidade vocal do indivíduo e as características fisiológicas laríngeas e supralaríngeas (Mendoza et al., 1996).

A nível supralaríngeo depreende-se que o efeito do trato vocal poderá contribuir para a identificação do género através do enfraquecimento ou amplificação de determinadas frequências de formantes verificando-se, desta forma, diferenças nos espectros gerados por homens e mulheres. No género feminino, por exemplo, há uma menor inclinação espectral e amplitude mais elevada nos 3000Hz

(Linville, 2002). Os valores elevados de f_0 e uma maior energia espectral na região correspondente ao terceiro formante (por volta dos 3 kHz) estão diretamente relacionados com a soproidade e astenia, características da voz feminina que se devem ao encerramento incompleto das pregas vocais (Mendoza et al., 1996).

Alguns autores sugerem uma maior amplitude no H1 comparativamente ao H2, numa voz feminina em oposição ao que acontece na voz masculina (Mendoza et al., 1996). A diferença entre os géneros deriva da formação de uma soproidade na laringe, natural das mulheres em comparação com a dos homens. (Klatt & Klatt, 1990). Desta forma é de esperar um maior valor de $(H1H2)_{LTAS}$ no género feminino.

Tabela 3 – Valores de referência para as medidas obtidas a partir do LTAS

Artigos	Língua	Amostra	Tarefa	α (dB)	$(H1-H2)_{LTAS}$ (dB)	Leq (dB)
(Da Silva et al., 2011)	Português (Brasil)	M	LN	-14,28		71,96
			Lfo	-9,98		80,13
(Leino, 2009)	Finlandês	H	LN	-10,5		70,9
(Araújo, 2015)	Português (cantoras tradicionais)	M (cantadeiras)	DE	-10,63	-0,267	73,67
			LN	-11,34	0,46	74,33
			C (Forte)	-0,809	-1,909	88,09
(Guzman, Correa, Muñoz, & Mayerhoff, 2013)	Espanhol (Atores)	H (atores)	DE	-18,44 ± 3,25		
		M (atrizes)		-17,46 ± 3,29		
(Ng, Chen, & Chan, 2012)	Cantonês	H	LN	-15,75 (1,58)		
		M		-16,44 (2,30)		
	Inglês	H		-16,47 (1,89)		
		M		-17,44 (2,30)		

Legenda: H – Homens; M – Mulheres; LN – Leitura Normal; Lfo – Leitura em intensidade forte; DE – Discurso Espontâneo; C – Canto

A CPP é usada na análise de vogais sustentadas e amostras de discurso (Brockmann-Bauser et al., 2019). Alguns autores, que compararam os valores obtidos nas vogais sustentadas e na leitura, concluíram que os valores são semelhantes nas duas tarefas pelo que, as propriedades mais instáveis da fala não afetam os valores de CPP. No entanto, também foram demonstradas melhores correlações entre disфонia e soproidade com CPP e CPPS em amostras de discurso espontâneo. Hasavand et al, demonstrou que os valores de CPP e CPPS em leitura são significativamente diferentes entre pessoas saudáveis e disfónicas tanto nos homens como mulheres, no entanto, na tarefa de vogais sustentadas este facto apenas foi verificado nas mulheres. Em todo o caso os valores em pacientes disfónicos nas duas tarefas são sempre inferiores em comparação a indivíduos com voz normal (Brockmann-Bauser et al., 2019).

Awan et al demonstraram que o CPPS aumenta com o aumento da intensidade em vogais (valor de SPL) tanto em homens como em mulheres. Outros autores demonstraram que o CPPS aumenta com o aumento da intensidade tanto em vogais sustentadas como no discurso espontâneo em professores

com voz normal. Num estudo de Brockman–bauser, et al (2019) o CPPS aumentou com a intensidade em mulheres com e sem patologia vocal. Os valores de SPL e f_0 mostram efeitos significativos no CPPS, embora a f_0 não seja tão evidente. Isto mostra que as alterações ao nível da intensidade da tarefa podem influenciar a interpretação dos valores de CPPS. Desta forma confirma-se que nas recolhas do sinal acústico deve existir um maior cuidado com a calibração da amplitude do som (Brockmann–Bauser et al., 2019).

Embora a CPP e a CPPS sejam medidas descritas para avaliar a qualidade vocal, e não tenham sido encontradas referências às diferenças entre género, são medidas que têm vindo a mostrar-se relevante na avaliação acústica. Uma vez que já foi provado que estão particularmente associadas ao parâmetro perceptivo de sopro, que, como já foi mencionado, está relacionado com a voz normal do género feminino, e não foram encontrados estudos que explorassem este aspeto, considerou-se pertinente a sua abordagem neste trabalho.

Tabela 4 – Valores de Referência CPP e CPPS obtidos através do software Praat®

Artigos	Língua	Amostra	CPPS (dB)	
			Va	LN
(Núñez-Batalla et al., 2019)	Espanhol (Espanha)	M		9,87 (1,28)
		H	16,4 (2,80)	9,65 (1,28)
(Delgado et al., 2017)	Espanhol (Espanha)	M	14,84 (2,13)	8,40 (0,77)
		H	15,54 (2,29)	7,86 (0,76)
(Delgado–Hernández et al., 2018)	Espanhol (Espanha)	H+M	14,90 (1,81)	7,98 (0,79)

Legenda: H – Homens; M – Mulheres; Va – Vogal [a] sustentada; LN – Leitura em intensidade normal

2.3. Relevância do género vocal para a Terapia da Fala

Segundo o Decreto Lei nº564/99 de 21 de dezembro de 1999 e a *American Speech–Language–Hearing Association* (ASHA) o Tpf é o profissional responsável pela prevenção, avaliação, diagnóstico, intervenção e estudo científico das perturbações da fala, linguagem, comunicação e deglutição, abrangendo não só todas as funções associadas à compreensão e expressão da linguagem oral e escrita, mas também outras formas de comunicação não verbal, em todas as faixas etárias (ASHA, n.d.) Dentro das perturbações da comunicação encontra-se a perturbação vocal, que pode, ou não, ter origem patológica. A intervenção consiste em definir objetivos com base nas expectativas do indivíduo para melhorar a sua forma de comunicar e considerando não só as características do ponto de vista biológico, mas também psicossocial (SPTF, 2020). As questões de identidade de género podem aparecer associadas à procura por uma readaptação à forma de comunicação, razão pela qual indivíduos nestas condições procuram o Tpf.

O termo gênero passou a ser utilizado no intuito de complementar a noção de sexo com o aparecimento do feminismo (Meyer, 2013). O sexo constitui a natureza biológica do ser vivo, na qual cada indivíduo nasce com características que o tornam membro de um determinado sexo. O gênero é uma construção social do sexo, ou seja, refere-se aos comportamentos e atitudes que a cultura passa para o indivíduo. Tal como Silveira (2006) afirma, natureza e cultura compõem o processo de construção social da identidade de gênero.

A voz assume um papel importante na identificação dos gêneros principalmente após a puberdade. A partir da muda vocal, na adolescência, a ação hormonal ocasiona alterações no aparelho fonador que resultam em qualidades vocais distintas entre os sexos (Behlau, 2004).

Quando esta muda vocal não ocorre o indivíduo desenvolve uma alteração vocal designada puberfonia. Esta patologia está relacionada com problemas no desenvolvimento vocal na adolescência, com um impacto mais acentuado no gênero masculino. No homem espera-se que aquando da muda vocal normal a voz se torne temporariamente mais rouca, fraca e instável. Quando a muda vocal finaliza a voz estabiliza e a altura tonal torna-se mais grave (diminuição da f_0 em 12 ST), o ângulo tiroideo aumenta, assim como a proeminência da cartilagem tiroidea, em simultâneo com o crescimento das pregas vocais. No entanto, na muda vocal incompleta permanece o padrão vocal infantil, na qual apenas ocorre a descida de 8 ST na extensão vocal. Perceptivamente a voz é aguda, revela esforço fonatório, fadiga vocal e pouca eficiência glótica. Estas questões, nomeadamente a alteração da altura tonal, fazem com que a voz possa causar confusão ao nível do gênero do falante. Quando ocorre no gênero feminino, traduz-se numa voz feminina infantil ou imatura que por apresentar uma voz excessivamente aguda pode tornar-se um problema a nível psicossocial. No entanto é no gênero masculino que esta patologia ocorre mais frequentemente, estando maioritariamente relacionada com questões psicoemocionais (Cielo et al., 2009).

Considerando a estreita relação entre gênero e sexo, importa explorar problemáticas relacionadas com os mesmos, nomeadamente orientação sexual e identidade de gênero. “A orientação sexual é definida como a atração sexual e/ou emocional de uma pessoa por outra pessoa” (APA, 2015). A identidade de gênero é o sentido que cada um tem sobre si como pertencendo ao gênero masculino ou feminino, relacionando-se com os papéis sociais do gênero com que se identifica, tanto ao nível de comportamento como de aparência (Appleby e Anatas, 1998, citado por Nogueira et al., 2010).

Até este ponto, compreende-se que, no caso da orientação sexual, é abordada apenas a atração do indivíduo, não estando direcionada com a sua conceção de si próprio e não existindo necessariamente procura por adquirir padrões do outro gênero. Por outro lado, na identidade de gênero o indivíduo pode não se identificar com o sexo e os papéis sociais a ele inerentes com que nasceu, procurando alterar o seu comportamento, nomeadamente a forma de comunicação para ir de encontro ao gênero com que se identifica.

O termo transgénero engloba os indivíduos que se identificam como sendo de outro género que não o correspondente ao seu sexo biológico, estando aqui presente a identidade de género. É a partir da muda vocal, que grande parte dos indivíduos transgéneros iniciam as suas transformações (Arán, Zaidhaft, & Murta, 2008). Quando os indivíduos recorrem a tratamentos e/ou cirurgias para alterar a sua componente biológica – o sexo – entramos dentro do campo da transexualidade. Basta que o indivíduo recorra a tratamento hormonal para ser classificado como transexual. A transexualidade implica, contudo, um desejo de viver ou uma experiência de vida de acordo com que é convencionalmente atribuído a outro género. O transgenderismo surge conceptualmente no final do século XX (Virgínia Prince, 1979, citada por Nogueira et al., 2010) e de acordo com a APA (2009, citada por Nogueira et al., 2010), considera-se que inclui todas as identidades ou expressões de género fora das normas de género convencionalmente aceites. Assim, a transexualidade é lida como uma forma de transgenderismo e que inclui outras formas como o/a travesti, o/a drag, entre outras (Nogueira et al., 2010).

Na comunidade transexual é comum que os indivíduos procurem alterar a sua voz de forma a que esta se aproxime ao máximo com o género com o qual se identificam. Para isto os indivíduos transexuais recorrem a tratamentos hormonais e cirurgias específicas que permitem alterar a fisiologia e função vocal. No entanto nem sempre é o suficiente para atingirem os objetivos definidos para a sua própria voz a diferentes níveis: altura tonal por a voz não estar tão aguda ou tão grave como esperavam; qualidade vocal, por inadaptação às alterações do aparelho fonador; outras razões mais específicas relacionadas com outros parâmetros, tais como a ressonância ou respiração. Estes indivíduos podem sentir dificuldade em participar em certas atividades, medo de falar ao telemóvel, desemprego e insucesso nas relações (Byrne, 2007; King, Brown & McCrea, 2011; Neumann & Welzel, 2004, citado por Leung, 2018). Estas restrições vêm muitas vezes acompanhadas de estigma, preconceito e discriminação (Byrne, 2007, citado por Leung, 2018). Na sua generalidade, é uma comunidade que procura encontrar a normalidade, privacidade, pertença e segurança junto da sociedade (Byrne, 2007, citado por Leung, 2018).

Embora não seja muito comum a procura de Terapia da Fala por parte destes indivíduos, é sabido que o Tpf é o profissional especializado para intervir neste tipo de alterações vocais (Leung, Oates, & Chan, 2018). O seu papel passa por ajudar estes indivíduos a desenvolver os padrões de comunicação respeitantes ao género que eles procuram (Coleman et al., 2012, citado por Leung, 2018). Para Bergel, 1999 no caso de transexuais homens, que procuram tornar-se mulheres, para a voz se tornar mais feminina é necessário modificar os parâmetros de entoação, de ressonância, de elevação da f_0 , o vocabulário, a gramática, o enunciado e os gestos durante a fala.

Explorando as problemáticas relacionadas com a patologia vocal, vertente mais associada ao trabalho do TpF, o Edema de Reinke é uma lesão laríngea que provoca alterações na voz e que frequentemente tem influência na perceção auditiva do género do falante, especialmente no caso das mulheres. Nesta patologia ocorre acumulação de fluído na camada superficial da lâmina própria das pregas vocais. É uma patologia que está fortemente relacionada com o tabagismo, embora também possa estar relacionada com outros fatores, tais como hormonais, envelhecimento e maus usos vocais. Audio-percetivamente a voz associada ao Edema de Reinke apresenta rugosidade, altura tonal grave e com pouca variação, voz fluida, crepitação e TMF reduzido. Acusticamente há redução da f_0 e alterações ao nível *jitter*, *shimmer* e HNR (Pereira et al., 2018). A voz da mulher com Edema de Reinke pode tornar-se tão grave, devido ao aumento da massa das pregas vocais, que pode chegar a ser confundida com a voz masculina (Behlau, 2004), razão pela qual, a maior parte dos utentes que chegam à Terapia da Fala são do género feminino (Pereira et al., 2018). No caso dos homens esta alteração não é na maior parte das vezes sentida como um problema, sendo inclusivamente percecionada como positiva, já que acentua através do agravamento da altura tonal o marcador de masculinidade (Behlau, 2004).

Embora seja um tema menos abordado e pouco associado ao TpF, a procura por uma forma diferente de comunicar, ou de melhorar as competências de comunicação em determinados contextos, é também um motivo de procura de intervenção junto do TpF. Estas questões podem estar relacionadas apenas com questões de gosto pessoal em relação à sua forma de comunicar, mas também podem estar relacionadas com parâmetros que englobam o papel de cada género na sociedade.

Na sociedade atual, apesar de as diferenças entre homens e mulheres estarem menos evidentes, em comparação com o século anterior, o homem continua a ter um papel de liderança na sociedade. Esta conceção sobre o homem e a opção pelo género masculino para ocupação de cargos de liderança em determinados contextos, pode ser explicada por, na sua generalidade, terem a capacidade de ser mais assertivos, controlados e seguros (Eagly & Karau, 2002). Este aspeto tem sido estudado por diversos autores que procuram entender as perceções da voz de indivíduos do mesmo género e indivíduos do género oposto. É possível perceber através destes estudos que, homens com vozes mais graves são considerados mais dominantes e mais atraentes. A mulheres com vozes mais graves também é atribuída a característica de dominante, no entanto nas mulheres as vozes mais agudas (até 260Hz) são mais atraentes, embora esta ligação não seja linear (Borkowska & Pawlowski, 2011). Na mesma linha de pensamento para o género masculino, as mulheres com uma altura tonal mais baixa são vistas como mais competentes, fortes e confiáveis (Klofstad, Anderson, & Peters, 2012)

2.4. Objetivos da investigação

Tendo em conta a introdução teórica apresentada e o facto de serem ainda escassos valores de referência de alguns parâmetros vocais para o PE, o objetivo geral proposto para a presente investigação é «analisar as propriedades acústicas da voz associadas à caracterização do género em falantes de PE», do qual surgem os seguintes objetivos específicos:

- i. Comparar parâmetros vocais acústicos de acordo com o género dos falantes;
- ii. Identificar combinações de medidas robustas na caracterização vocal do género feminino e masculino;
- iii. Identificar os marcadores de género mais predominantes numa amostra combinada de género masculino e feminino.

3. Métodos

3.1 Tipo e Desenho de estudo

O presente estudo está integrado numa linha de investigação denominada “Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação”, da responsabilidade do Professor Doutor André Araújo, aprovada pela Comissão de Ética da Escola Superior de Saúde do Politécnico do Porto (ESS), segundo confirma o parecer com o número de registo 1296/2017 (**Anexo 1**), à qual foi solicitada uma adenda para continuação dos estudos iniciados em 2017.

Este projeto de investigação pressupõe a observação e descrição de relações entre variáveis, descrição com base na estatística e verificação de teorias previamente concebidas com base na pesquisa teórica prévia. Uma vez que houve apenas um momento de recolha de dados, podemos afirmar que estamos perante um estudo do tipo observacional analítico de carácter transversal. Ao realizar este tipo de estudo, a equipa de investigação não intervém na recolha dos dados, apenas desenvolve procedimentos para observar o comportamento desses mesmos dados (Ribeiro, 2010). Neste caso, através de gravações, foi possível observar o comportamento de diferentes parâmetros de medidas acústicas em indivíduos do género feminino e masculino. Este comportamento é descrito através de valores estatísticos, seguindo, portanto, uma metodologia quantitativa (Carmo & Ferreira, 2008)

3.2. Amostra

A amostra utilizada neste estudo está incluída numa base de dados extraída pela investigadora num projeto prévio realizado em 2018 (Pereira, 2018), também ele integrado na mesma linha de investigação.

A base de dados é composta por uma amostra não probabilística de conveniência (Carmo & Ferreira, 2008), constituída por um grupo de 30 participantes voluntários, 15 do género feminino (média de idade - 21,3) e 15 do género masculino (média de idade - 19,6). Apenas foram selecionados os indivíduos com idade compreendida entre os 18 e os 25 anos, que apresentassem o PE como língua materna e uma avaliação perceptiva, realizada pela equipa de investigação no momento da recolha, através da GRBAS (Hirano, 1981, citado por Guimarães, 2007; Freitas, 2012), sem alterações. Como critérios de exclusão, os participantes não podiam apresentar queixas ou historial de problemas respiratórios, ou alguma perturbação da fala diagnosticada, nomeadamente ao nível da voz e articulação (Ambreen et al., 2019; Guimarães & Abberton, 2005; Spazzapan et al., 2018).

3.3. Instrumentos

O instrumento de recolha de dados, disponível no **Apêndice 1**, foi construído tendo em conta os procedimentos e tarefas definidas para outros protocolos de avaliação (Pinho & Pontes, 2002; Tavares, 2014) e em investigações prévias, com base no *LTAS*, realizadas pelo investigador responsável (Araújo, 2015).

Este instrumento foi construído para ser utilizado na íntegra aquando das gravações de cada indivíduo, sendo composto por duas secções. A primeira secção é constituída por um protocolo de avaliação, onde é possível registar os dados de cada participante, tendo em atenção o código atribuído a cada um deles. Este código foi usado posteriormente para a identificação dos participantes de forma a garantir a confidencialidade dos dados. Nesta mesma secção é possível registar aspetos que possam influenciar o desempenho vocal do indivíduo, nomeadamente, ao nível da postura, respiração e alterações no trato vocal. Consta, também, uma análise vocal perceptiva tendo por base os parâmetros da GRBAS que permite perceber previamente se a voz do indivíduo cumpre os critérios de inclusão.

Na segunda parte constam todos os procedimentos a adotar antes da chegada do participante, ao nível da montagem dos componentes eletrónicos e as tarefas a desempenhar pelo indivíduo durante a gravação, com indicações acerca do código atribuído a cada tarefa, da ordem e da distância a que cada tarefa deve ser realizada.

As tarefas selecionadas para constar no instrumento de recolha de dados são as mais utilizadas na avaliação vocal:

- discurso espontâneo (DE): através de um pequeno questionário de respostas abertas que permite obter uma amostra de fala sem ser “conscientemente planeada, preparada ou lida” (SPTF, 2020). Esta tarefa tem a vantagem de obter resultados mais realistas e representar melhor a dinâmica vocal, no entanto, por apresentar maior variabilidade, implica técnicas de análise mais sofisticadas (Laver, Hiller and Beck, 1992, citado por Guimarães, 2002);
- amostras de discurso semi-espontâneo (DSE), através da descrição de uma, ou mais, imagens complexas (Santos, 2013; SPTF, 2020);
- leitura em diferentes intensidades – normal (LN), fraca (LFr), forte (LFo) – do texto “O Sol”. O uso de tarefas de leitura tem a vantagem de apresentar uma boa consistência em diferentes momentos e permite obter valores de f_0 mais próximos do discurso espontâneo (Fitch, 1990; Baken and Orlikoff, 2000; citado por Guimarães, 2002). «O Sol» é uma narrativa construída especificamente para PE por Mendes e Castro (2005). É composto por todos os fonemas e formatos silábicos existentes e uma média de frequência relativa de ocorrência de fonemas semelhantes ao discurso espontâneo (Mendes et al., 2011; SPTF, 2020);
- canto (C) da música popular “Parabéns a Você” por ser de conhecimento geral.

- vogais [a] (Va), [i] (Vi) e [u] (Vu) sustentadas. Estas vogais representam os extremos articatórios do espaço vocálico, permitindo observar as duas principais dimensões: alta/baixa e anterior/posterior (Horii,1979, citado por Guimarães & Abberton, 2005). São tarefas que refletem uma condição relativamente estável do sistema fonatório (Horii, 1979, Titze, 1987, citado por Guimarães & Abberton, 2005), para além de que são mais utilizadas para obter medidas relacionadas com a f_0 , nomeadamente por permitirem avaliar a estabilidade da laringe (Horii, 1979, Titze, 1987, citado por Guimarães & Abberton, 2005), não estarem dependentes de fenómenos de coarticulação ou da entoação dos indivíduos e por serem mais simples de recolher e analisar (Laver, Hiller and Beck, 1992; citado por Guimarães, 2002).
- glissandos ascendentes (GA) e descendentes (GD) em extensão vocal máxima (ExtV).
- fricativas [s] e [z] em tempo máximo de emissão.

Para testar a aplicabilidade do instrumento utilizado foi realizado um teste piloto, onde foi solicitado, a dois voluntários participantes, que dessem a sua opinião acerca dos procedimentos realizados. Deste teste surgiu a necessidade de reestruturar a ordem das tarefas solicitadas, bem como as estratégias usadas para orientar os participantes, de forma a facilitar a compreensão e a execução.

Para a recolha das amostras foi cedido, pela área técnico científica de Terapia da Fala da ESS, um microfone (*Audio-Technica Midnight Blues®*), uma coluna comum e uma placa de som (iO2 Express da *Alesis®*). O registo do sinal de fala foi calibrado através de um decibelímetro (*CESVA® SC-3010*), disponibilizado pela área técnico científica de Saúde Ambiental da ESS. De forma a registar o sinal, o grupo de investigação utilizou um computador com o *software* Audacity®.

Para proceder à análise acústica foram utilizados dois *software*, tendo em consideração as referências utilizadas para este trabalho. Para obter os valores relacionados com a f_0 (f_0 media, f_0 DP e Gama Tonal) (Tabela 1) e medidas de curto-termo (*jitter*, *shimmer* e HNR) (Tabela 2) foi utilizado o *software* Praat®, assim como para obter os valores de CPP e CPPS (Tabela 4).

Para proceder à extração de medidas do LTAS, foi utilizado o *software* Sopran® por estar disponível gratuitamente e por permitir extrair os valores do espectro de longo termo para tratamento externo. Este apresenta várias funções, tais como: calibrar o sinal, realizar alterações na imagem, retirar valores exatos num certo ponto e obter a lista dos principais valores de todo o espectro, essencial para os cálculos das medidas necessárias. Este *software* foi utilizado por Sundberg, Salomão, & Scherer (2019) para estudar a expressividade emocional em cantores através do LTAS. Uma vez extraídos, os valores do LTAS foram inseridos num ficheiro Excel com todas as fórmulas necessárias para o cálculo das respetivas medidas.

As análises estatísticas posteriormente mencionadas foram efetuadas com o Software SPSS Statistics® (v.26).

3.4. Procedimentos

3.4.1. Recrutamento e seleção dos participantes

Os participantes foram recrutados através da rede de contactos pessoais da equipa de investigação. Foram selecionados tendo em conta os critérios de inclusão e de exclusão e mediante a sua disponibilidade foram agendadas datas e horas, na tentativa de proceder ao maior número de gravações no mesmo período, de forma a garantir as mesmas condições de gravação. Todos os participantes autorizaram a gravação e tratamento de dados através de um consentimento informado, disponível no Anexo 2.

3.4.2. Recolha do Sinal

As recolhas foram realizadas no laboratório de Terapia da Fala da ESS. Embora tenham sido efetuadas sempre no mesmo local, foram realizadas em dias diferentes, pelo que foi necessário assegurar as mesmas condições de gravação através da calibração do sinal sonoro. Este procedimento é essencial para garantir a fiabilidade das recolhas, que permite recolher a verdadeira intensidade das ondas sonoras num certo momento com determinadas condições de ambiente. Para tal, através de uma coluna, foi emitido o sinal de um tom puro (440Hz) que foi medido às três distâncias – 7,5 cm, 15 cm e 30 cm – foram definidas consoante a bibliografia e testes realizados tendo por base as três intensidades que foram utilizadas durante as recolhas – fraco, médio e forte – respetivamente (Master et al., 2006).

Para a gravação o microfone foi posicionado num canto da sala, evitando ao máximo, a reflexão do som em paredes ou outros objetos. O mesmo foi ajustado de forma a ficar com uma inclinação de aproximadamente 45º da boca do participante, diminuindo assim, a captação de sinais aperiódicos originados pelas consoantes oclusivas e sons fricativos presentes no discurso (A. . Mendes & Castro, 2005). Junto ao microfone foi colocada uma marcação de referência para os participantes, para manterem o seu posicionamento ao longo da gravação. Foi instalada uma estante para segurar as folhas com as imagens e textos utilizados. Esta estante foi posicionada de frente para o indivíduo, mas inclinada e ligeiramente mais baixa, de forma a evitar a reflexão do som. O computador ficou em cima de uma mesa perto do microfone virado no sentido oposto ao do participante, para que este não se distraísse com o mesmo ou tivesse *feedback* da gravação. Após estabelecidas as condições de gravação, foi aplicado o instrumento de recolha de dados.

As tarefas de fala encadeada (DE, DSE, LN, LFr, LFo) e canto foram adaptadas de forma a obter 60 segundos de gravação, garantindo um mínimo de 40 segundos da amostra de uso vocal efetivo, o tempo ideal para que as medidas de longo termo não sejam afetadas por alterações de acentuação,

pelo padrão de articulação ou outras particularidades da emissão de cada indivíduo (Master et al., 2006; A. P. Mendes et al., 2015).

3.4.2. Procedimentos de Análise Acústica

Para as medidas aerodinâmicas foram recolhidos os tempos das emissões de todas as vogais e consoantes. Uma vez que cada tarefa foi repetida três vezes, para o TMF foi calculada a média dos valores para cada vogal, enquanto que para o coeficiente S/Z foram selecionados os valores mais elevados de cada consoante (Santos, 2009). No coeficiente S/Z foi ainda realizada a divisão do tempo máximo da emissão prolongada da consoante [s] pelo tempo da consoante [z] (Boone, 1971, citado por Behlau, 2004; SPTF, 2020).

Para obter a f_0 média, f_0 DP e Gama Tonal, em tarefas de longa duração a amostra foi previamente tratada removendo todos os excertos que apresentassem frequências desviantes provocadas por sons sibilantes ou registo basal. De seguida foi selecionada toda a amostra e recolhidos os valores: f_0 median (Hz); f_0 mean (Hz); f_0 std (Hz); f_0 min (Hz); f_0 max (Hz). Nas vogais sustentadas foram selecionados três segundos da segunda repetição de cada vogal e recolhidos os mesmos valores (Ambreen et al., 2019; Brockmann et al., 2011; Mezzedimi et al., 2017). Para obter a Extensão Vocal foi retirado o valor em Hz mais elevado obtido no GA e o valor mais baixo obtido no GD. De seguida procedeu-se à conversão dos valores nas notas musicais respetivas e contabilizados os ST entre a nota mais baixa e a nota mais alta.

À semelhança da obtenção da f_0 , nos parâmetros de curto-termo foram selecionados três segundos da segunda repetição de cada vogal e extraídos os valores: *jitter local*(%)⁴, *shimmer local*(%)⁵ e HNR (dB) (Ambreen et al., 2019; Brockmann et al., 2011; Mezzedimi et al., 2017).

Numa primeira fase de análise do LTAS, foi acrescentado ao sinal de gravação da tarefa, o tom de calibração obtido no dia de gravação à distância da respetiva tarefa, tendo-se procedido à calibração do mesmo através da funcionalidade presente no Sopran®. Ao selecionar a parte do ficheiro correspondente à tarefa, o *software* fornece o valor de Leq médio. Este valor foi retirado e inserido num ficheiro em formato *Excel*, previamente preparado para o efeito. Todos os espectros foram extraídos com os mesmos critérios de visualização: janela *Hamming*, um tempo de resolução de 40 milissegundos e uma banda de 50 Hz. Estes critérios foram selecionados de forma a permitir uma melhor visualização do espectro. Foram retirados todos os valores de frequência e intensidade

⁴ Jitter [%]: representa a diferença da frequência entre ciclos consecutivos e a frequência média do sinal (Boersma & David Weenink, 2021).

⁵ Shimmer [%]: representa a diferença da amplitude entre ciclos consecutivos e a amplitude média do sinal (Boersma & David Weenink, 2021).

correspondente até aos 6000 Hz e inscritos no ficheiro *Excel* referido acima, permitindo, através das fórmulas previamente inseridas, obter os valores finais de Leq , α e $(H1-H2)_{LTAS}$.

As tarefas utilizadas para obter a CPP e a CPPS foram as leituras em diferentes intensidades e a vogal [a] prolongada. Para além de ser a vogal mais utilizada na análise acústica, foi selecionada a vogal sustentada [a], porque mantém constantes as características laríngeas e do trato vocal, sofrendo menor influência de fatores fonéticos e suprasegmentais (Parsa & Jamieson, 2001, citado por Belo, 2015). As tarefas de leitura permitem manter a consistência na amostra entre sujeitos e, em simultâneo, são tarefas mais próximas do discurso espontâneo (Fitch, 1990; Baken and Orlikoff, 2000; citado por Guimarães, 2002). Os procedimentos para extração destas medidas estão descritos na Tabela 5 (Delgado-Hernández et al., 2018; Phadke et al., 2018):

Tabela 5 – Procedimentos para extração dos valores de CPP e CPPS, para versão 6.1.34 do Praat, baseado em Phadke et al. (2018).

1º Selecionar a amostra de vogal sustentada ou de leitura

2º Abrir "Analyze periodicity" e escolher "To Power Cepstrogram" na janela do Praat Objects

3º Usar as seguintes definições para gerar o cepstrograma:

Parâmetros	CPP	CPPS
Pitch floor (Hz)	60	60
Time step (s)	0,002	0,002
Maximum frequency (Hz)	5000	5000
Pre-emphasis form (Hz)	50	50

4º ao selecionar o "powercepstrogram" gerado, escolher "Query", selecionar "Get CPPS" do menu e usar as seguintes definições

Select subtract tilt before smoothing	Yes	No
Time averaging window (s)	0,001	0,01
Quefrequency averaging window (s)	0,00005	0,001
Peak search pitch range (Hz)	60-330	60-330
Peak search tolerance (0-1)	0,05	0,05
Interpolation	Parabolic	Parabolic
Tilt line quefrequency range (s)	0,001-0,0 (=end)	0,001-0,0 (=end)
Line type	Exponential decay	Satright
Fit method	Robust	Robust

3.4.3. Procedimentos de Análise Estatística

Tendo em conta que o objetivo da presente dissertação é realizar uma caracterização do género dos falantes, o primeiro passo foi fazer uma comparação entre o género feminino e o género masculino, de forma a perceber quais as medidas que melhor diferenciam os géneros ao nível vocal. A significância da diferença entre géneros foi avaliada a partir de um teste de hipóteses para amostras independentes, onde: H0: Os valores obtidos nas medidas acústicas são iguais entre os géneros. H1: Os valores obtidos nas medidas acústicas são diferentes entre os géneros. Os pressupostos deste método estatístico, nomeadamente a normalidade das distribuições e a homogeneidade das variâncias nos dois grupos foram avaliados com o teste Shapiro-Wilk e com o teste de Levene, respetivamente. Nas medidas para as quais se verificou o pressuposto da normalidade procedeu-se à aplicação do teste *t*-Student. Após verificar a homogeneidade das variâncias foram observados os valores de *p-value* respetivos, consoante cumpriam ou não a homogeneidade, com um nível de significância de 0,05. Nas medidas em que não se verificou o pressuposto da normalidade foi realizado o teste não paramétrico de Wilcoxon - Mann-Whitney, para o qual, à semelhança do teste paramétrico, apenas se consideram significativas as médias cujo *p-value* for inferior ou igual a 0,05 (Marôco, 2011).

A Análise de Componentes Principais (*Categorical Principal Components Analysis-CATPCA*) é um método de análise multivariada que permite resumir o número de variáveis em componentes principais, na qual o primeiro componente corresponde ao conjunto de variáveis que representam a maior variação da amostra (Marôco, 2011). Para proceder a esta análise, foram realizados vários testes, de forma a perceber de que forma a mesma poderia contribuir para os objetivos da presente dissertação. Uma vez verificado que a repetição de medidas originava difíceis interpretações e diminuía o peso de cada parâmetro dentro dos componentes, foram selecionadas apenas as tarefas mais representativas da investigação que tem vindo a ser realizada. Assim, em representação das medidas de f_0 (f_0 média, f_0 desvio padrão, gama tonal e extensão vocal) e de LTAS (α , $(H1-H2)_{LTAS}$, Leq) , utilizou-se a tarefa de leitura normal por ser estatisticamente semelhante às restantes tarefas de fala encadeada (Fitch, 1990; Baken and Orlikoff, 2000; citado por Guimarães, 2002; Pereira, 2018). Para as medidas de perturbação da frequência (*jitter*), intensidade (*shimmer*) e HNR foi selecionada a tarefa de vogal [a] sustentada, por ser a mais utilizada em investigação (Tabela 2). Em relação às medidas de CPP e CPPS, foi incluída a leitura em intensidade normal e a vogal [a] sustentada (Tabela 4). Como critério de retenção das componentes foi utilizada a regra de *eigenvalue* superior a 1, assim como a análise do *Scree plot* dos valores de *eigenvalue*. Nesta análise são selecionados todas as componentes até ao ponto de inflexão da curva, ou seja, até que a curva comece a ficar mais horizontal. A consistência interna de cada componente foi medida com o *alpha de Cronbach*. Após a análise realizada através da CATPCA procedeu-se à interpretação dos resultados de forma a perceber a relação entre fatores que estão dentro de cada componente (Marôco, 2011).

4. Resultados

O comportamento das medidas entre os dois géneros foi estudado através do teste t-Student para amostras independentes para as medidas que cumprem o pressuposto da normalidade. Para as medidas que não cumprem o pressuposto foi utilizado o teste de Wilcoxon-Mann Whithney. Considera-se que existem diferenças significativas entre os géneros, quando o p-value do teste é inferior ou igual a 0,05. Como nem todas as medidas seguem a distribuição normal, a estatística descritiva é apresentada por mediana e valores mínimos e máximos (Marôco, 2011).

Tabela 6 – Estatística Descritiva e Testes de Hipóteses para Medidas Aerodinâmicas

Medidas	Total	Homens	Mulheres	Teste t	Teste WM
TMF_Va	n=30	n=15	n=15	p-value	p-value
	Mediana	18,719	18,747	18,690	0,906
(Min, Max)	(9,87; 33,86)	(9,87; 33,86)	(13,98; 23,83)		
TMF_Vi	Mediana	21,194	20,148	23,181	0,369
	(Min, Max)	(11,23; 38,62)	(11,23; 38,62)	(15,44; 35,43)	
TMF_Vu	Mediana	19,983	18,278	21,094	0,202
	(Min, Max)	(11,25; 36,38)	(11,25; 36,38)	(14,19; 26,65)	
Coef_S/Z	Mediana	1,039	1,122	0,965	0,087
	(Min, Max)	(0,67; 1,84)	(0,68; 1,58)	(0,67; 1,84)	

Legenda: Va – Vogal [a] sustentada; Vi – Vogal [i] sustentada; Vu – Vogal [u] sustentada; TMF – Tempo Máximo de Fonação; Coef – Coeficiente; Teste t – Teste t-student para amostras independentes; Teste WM – Teste de Wilcoxon-Mann Whitney

Na Tabela 6 é possível perceber que não se verificam diferenças significativas entre géneros para as Medidas Aerodinâmicas. Embora as diferenças não sejam estatisticamente diferentes é possível verificar que, nesta amostra, os valores de TMF são superiores no género feminino para a vogal [i] e [u]. No entanto é possível observar que os valores são mais discrepantes no género masculino, uma vez que apresenta os valores mínimos mais baixos e os valores máximos mais elevados em comparação com o género feminino.

No coeficiente S/Z é possível observar um valor superior a 1 para os homens e inferior para as mulheres, embora o valor máximo do coeficiente (1,84) esteja na amostra do género feminino.

Tabela 7 – Estatística Descritiva e Testes de Hipóteses para valores de f_0

Medidas		Total	Homens	Mulheres	Testet	Teste WM
		n=30	n=15	n=15	p-value	p-value
f_0 media_DE ¹	Mediana	153,130	110,886	190,810	0,000*	
	(Min, Max)	(94,32; 225,48)	(94,32; 136,82)	(169,44; 225,48)		
f_0 media_DSE	Mediana	154,917	112,269	202,288	0,000*	
	(Min, Max)	(84,94; 232,18)	(84,94; 139,58)	(170,25; 232,18)		
f_0 media_LN	Mediana	162,650	122,145	210,352	0,000*	
	(Min, Max)	(84,20; 228,01)	(84,20; 151,67)	(173,63; 228,01)		
f_0 media_LFr	Mediana	154,097	114,078	189,689	0,000*	
	(Min, Max)	(82,92; 211,39)	(82,92; 142,16)	(166,04; 211,39)		
f_0 media_LFo	Mediana	287,402	221,134	344,201	0,000*	
	(Min, Max)	(167,65; 402,55)	(167,65; 281,05)	(293,76; 402,55)		
f_0 media_C	Mediana	198,633	135,637	235,687	0,000*	
	(Min, Max)	(91,60; 280,27)	(91,60; 197,09)	(200,34; 280,27)		
f_0 media_Va	Mediana	151,361	113,275	198,745	0,000*	
	(Min, Max)	(78,68; 264,62)	(78,68; 134,18)	(168,54; 264,62)		
f_0 media_Vi	Mediana	165,123	116,874	217,892	0,000*	
	(Min, Max)	(78,67; 274,42)	(78,67; 144,09)	(186,16; 274,42)		
f_0 media_Vu	Mediana	164,244	115,042	209,244	0,000*	
	(Min, Max)	(78,86; 275,92)	(78,86; 149,33)	(179,16; 275,92)		
f_0 DP_DE ¹	Mediana	17,132	12,252	24,734	0,000*	
	(Min, Max)	(8,03; 30,17)	(8,03; 24,04)	(15,23; 30,17)		
f_0 DP_DSE	Mediana	19,358	12,967	22,967	0,000*	
	(Min, Max)	(8,26; 46,86)	(8,26; 24,58)	(12,75; 46,86)		
f_0 DP_LN	Mediana	19,737	13,655	23,684	0,000*	
	(Min, Max)	(7,13; 43,65)	(7,13; 24,03)	(15,30; 43,65)		
f_0 DP_LFr	Mediana	14,221	11,153	17,192	0,003*	
	(Min, Max)	(5,23; 28,23)	(5,23; 19,27)	(11,11; 28,23)		
f_0 DP_LFo	Mediana	44,836	31,26	53,745	0,000*	
	(Min, Max)	(22,35; 70,54)	(22,35; 49,30)	(33,35; 70,54)		
f_0 DP_C	Mediana	36,402	25,846	45,485	0,000*	
	(Min, Max)	(14,03; 55,63)	(14,03; 40,98)	(34,34; 55,63)		
f_0 DP_Va	Mediana	1,169	1,022	1,243		0,001
	(Min, Max)	(0,63; 1,92)	(0,63; 1,43)	(1,02; 1,92)		
f_0 DP_Vi	Mediana	1,287	0,987	1,604	0,000*	
	(Min, Max)	(0,58; 2,27)	(0,58; 1,56)	(1,17; 2,27)		
f_0 DP_Vu	Mediana	1,127	0,782	1,504	0,000*	
	(Min, Max)	(0,47; 2,76)	(0,47; 1,50)	(1,04; 2,76)		
Gama_DE ¹	Mediana	15,500	16	15	0,951	
	(Min, Max)	(12; 23)	(12; 23)	(12; 23)		
Gama_DSE	Mediana	15,000	15	17	0,006*	
	(Min, Max)	(11; 22)	(11; 18)	(14; 22)		
Gama_LN	Mediana	15,000	14	18	0,000*	
	(Min, Max)	(11; 22)	(11; 17)	(13; 22)		
Gama_LFr	Mediana	14,000	14	15	0,062	
	(Min, Max)	(10; 22)	(11; 17)	(10; 22)		
Gama_LFo	Mediana	20,000	20	20	0,880	
	(Min, Max)	(15; 24)	(17; 24)	(15; 24)		
Gama_C	Mediana	1,127	0,782	1,504	0,877	
	(Min, Max)	(0,47; 2,76)	(0,47; 1,50)	(1,04; 2,76)		
ExtV_GAGD	Mediana	35,000	38	33	0,001*	
	(Min, Max)	(28; 45)	(33; 45)	(28; 37)		

Legenda: DE – Discurso Espontâneo; DSE – Discurso Semi-Espontâneo; LN – Leitura em intensidade normal; LFr – Leitura em intensidade fraca; LFo – Leitura em intensidade forte; C – Canto; f_0 – frequência fundamental; DP – desvio padrão; Gama – Gama Tonal; ExtV – Extensão Vocal; Teste t – Teste t – student para amostras independentes; Teste WM – Teste de Wilcoxon–Mann Withney; *n=28 (14+14); *p<0,05

Perante a apresentação de resultados para os valores de f_0 na Tabela 7, é possível verificar que existem diferenças significativas entre géneros em todas as tarefas para os valores de f_0 média e f_0DP , visivelmente mais elevados nas mulheres.

A Gama Tonal apresentou valores mais elevados no género feminino para as tarefas DSE e LN, enquanto que nas restantes tarefas não foram estatisticamente verificadas diferenças significativas.

A Extensão Vocal também apresenta diferença significativa entre géneros sendo, neste caso, mais elevada no género masculino.

Tabela 8 – Estatística Descritiva e Testes de Hipóteses para Parâmetros de Curto Termo

Parâmetros		Total	Homens	Mulheres	Teste t	Teste WM
		n=30	n=15	n=15	p-value	p-value
Jitter_Va	Mediana	0,396	0,469	0,331	0,005*	
	(Min, Max)	(0,20; 0,68)	(0,27; 0,68)	(0,20; 0,66)		
Jitter_Vi	Mediana	0,341	0,391	0,300		0,345
	(Min, Max)	(0,13; 0,98)	(0,20; 0,75)	(0,13; 0,98)		
Jitter_Vu	Mediana	0,341	0,366	0,307	0,789	
	(Min, Max)	(0,22; 0,51)	(0,22; 0,51)	(0,24; 0,51)		
Shimmer_Va	Mediana	2,850	3,268	2,680		0,148
	(Min, Max)	(1,87; 5,58)	(2,19; 5,58)	(1,87; 5,12)		
Shimmer_Vi	Mediana	2,001	2,655	1,799		0,063
	(Min, Max)	(0,76; 5,41)	(1,51; 3,51)	(0,76; 5,41)		
Shimmer_Vu	Mediana	2,205	2,406	1,818		0,011*
	(Min, Max)	(1,02; 4,30)	(1,79; 4,16)	(1,02; 4,30)		
HNR_Va	Mediana	17,715	17,140	18,302	0,049	
	(Min, Max)	(11,83; 22,87)	(11,83; 22,61)	(14,69; 22,87)		
HNR_Vi	Mediana	21,447	20,793	22,272	0,154	
	(Min, Max)	(15,03; 25,71)	(15,03; 23,98)	(18,32; 25,71)		
HNR_Vu	Mediana	23,185	22,218	25,323	0,001*	
	(Min, Max)	(17,47; 27,59)	(17,47; 25,49)	(20,90; 27,59)		

Legenda: Va – Vogal [a] sustentada; Vi – Vogal [i] sustentada; Vu – vogal [u] sustentada; *Jitter* – *jitter local*(%); *Shimmer* – *shimmer local*(%); HNR – Harmonic to Noise Ratio; Teste t – Teste t-student para amostras independentes; Teste WM – Teste de Wilcoxon-Mann Withney; *p<0,05

Consoante os valores estatísticos descritos na Tabela 8, existem diferenças significativas entre os géneros para *jitter_Va*, *shimmer_Vu* e *HNR_Vu*, não sendo possível afirmar, no entanto, que existam as mesmas diferenças para *shimmer_Va*, *HNR_Va*, *jitter_Vi*, *shimmer_Vi*, *HNR_Vi* e *jitter_Vu*;

Embora as medidas não sejam, na sua maioria, estatisticamente diferentes entre géneros, é possível observar maiores valores de *jitter* e *shimmer* no género masculino, enquanto que os valores de HNR se encontram mais elevados no género feminino.

Tabela 9 – Estatística Descritiva e Testes de Hipóteses para medidas de LTAS

Medidas		Total	Homens	Mulheres	Teste t	Teste WM
		n=30	n=15	n=15	p-value	p-value
<i>Alpha</i> _DE ^I	Mediana	4,940	4,68	5,04	0,010*	
	(Min, Max)	(3,74; 6,12)	(3,74; 5,33)	(4,35; 6,12)		
<i>Alpha</i> _DSE	Mediana	4,730	4,59	4,93		0,016*
	(Min, Max)	(3,45; 5,62)	(3,45; 5,11)	(4,01; 5,62)		
<i>Alpha</i> _LN	Mediana	5,095	4,96	5,20		0,026*
	(Min, Max)	(4,09; 5,78)	(4,09; 5,31)	(4,67; 5,78)		
<i>Alpha</i> _LFr	Mediana	4,915	4,54	5,07		0,000*
	(Min, Max)	(3,09; 6,15)	(3,09; 5,27)	(4,59; 6,15)		
<i>Alpha</i> _LFo	Mediana	6,205	6,04	6,56	0,000*	
	(Min, Max)	(5,64; 6,89)	(5,66; 6,31)	(5,64; 6,89)		
<i>Alpha</i> _C	Mediana	4,925	4,53	5,06	0,019*	
	(Min, Max)	(3,17; 5,86)	(3,17; 5,50)	(4,25; 5,86)		
H1H2 _{LTAS} _DE ^I	Mediana	2,970	1,71	4,20	0,021*	
	(Min, Max)	(-2,39; 6,76)	(-2,39; 6,76)	(0,17; 6,38)		
H1H2 _{LTAS} _DSE	Mediana	3,780	2,20	4,57	0,001*	
	(Min, Max)	(-3,22; 7,68)	(-3,22; 6,74)	(0,38; 7,68)		
H1H2 _{LTAS} _LN	Mediana	2,425	1,01	4,25	0,012*	
	(Min, Max)	(-3,75; 6,96)	(-3,75; 4,75)	(-2,62; 6,96)		
H1H2 _{LTAS} _LFr	Mediana	8,765	6,76	10,37	0,001*	
	(Min, Max)	(3,80; 14,65)	(3,80; 13,98)	(6,52; 14,65)		
H1H2 _{LTAS} _LFo	Mediana	-0,530	-2,77	1,67		0,870
	(Min, Max)	(-5,16; 4,89)	(-5,16; 0,94)	(-2,27; 4,89)		
H1H2 _{LTAS} _C	Mediana	4,670	4,60	4,70	0,020*	
	(Min, Max)	(-0,54; 9,73)	(-0,54; 7,19)	(2,69; 9,73)		
Leq_DE ^I	Mediana	62,885	63,35	61,86	0,912	
	(Min, Max)	(54,44; 73,44)	(54,44; 69,78)	(54,86; 73,44)		
Leq_DSE	Mediana	61,310	60,95	61,43	0,588	
	(Min, Max)	(53,14; 73,49)	(53,14; 69,88)	(53,56; 73,49)		
Leq_LN	Mediana	62,815	63,45	62,71	0,020*	
	(Min, Max)	(54,82; 73,09)	(54,82; 67,79)	(56,68; 73,09)		
Leq_LFr	Mediana	55,605	54,34	56,73	0,597	
	(Min, Max)	(46,89; 63,46)	(46,89; 63,46)	(50,87; 61,58)		
Leq_LFo	Mediana	83,8000	83,92	83,35		0,870
	(Min, Max)	(75,34; 99,68)	(75,34; 89,30)	(80,73; 99,68)		
Leq_C	Mediana	65,695	65,73	65,66	0,774	
	(Min, Max)	(55,12; 76,16)	(55,12; 76,16)	(63,21; 70,17)		

Legenda: Teste t – DE – Discurso Espontâneo; DSE – Discurso Semi-Espontâneo; LN – Leitura em intensidade normal; LFr – Leitura em intensidade fraca; LFo – Leitura em intensidade forte; C – Canto; fO – frequência fundamental; dp – desvio padrão; Gama – Gama Tonal; ExtV – Extensão Vocal; Teste t-student para amostras independentes; Teste WM – Teste de Wilcoxon-Mann Withney; *n=28 (14+14); *p<0,05

Os valores de *alpha* (inclinação espectral) e (H1H2)_{LTAS} (diferença entre o primeiro e segundo harmónico), à exceção do (H1H2)_{LTAS} na LFo, os valores são estatisticamente diferentes entre géneros, sendo mais elevados nas mulheres, tal como é possível observar na Tabela 9. Relativamente aos valores de Leq, parâmetro de descrição da intensidade vocal, não são visíveis diferenças significativas entre géneros.

Tabela 10 – Estatística Descritiva e Testes de Hipóteses para medidas cepstrais

Medidas		Total	Homens	Mulheres	Teste T	Teste MW
		n=30	n=15	n=15	p-value	p-value
CPP_LN	Mediana	18,793	19,418	18,210		0,002*
	(Min, Max)	(17,38; 20,19)	(17,38; 20,19)	(17,92; 19,62)		
CPP_LFr	Mediana	16,684	16,803	16,117	0,017*	
	(Min, Max)	(14,63; 18,77)	(15,55; 18,77)	(14,63; 18,07)		
CPP_LFo	Mediana	18,359	19,171	16,834		0,000*
	(Min, Max)	(16,20; 20,10)	(18,31; 20,10)	(16,20; 18,49)		
CPP_Va	Mediana	25,861	26,576	25,757		0,838
	(Min, Max)	(18,81; 31,79)	(20,72; 31,79)	(18,81; 27, 85)		
CPPS_LN	Mediana	8,992	8,704	9,207	0,038*	
	(Min, Max)	(7,65; 11,31)	(7,65; 9,92)	(8,37; 11,31)		
CPPS_LFr	Mediana	6,757	6,536	7,262	0,126	
	(Min, Max)	(4,66; 8,68)	(5,07; 8,34)	(4,66; 8,68)		
CPPS_LFo	Mediana	10,923	10,924	10,797	0,239	
	(Min, Max)	(9,14; 12,88)	(9,28; 12,88)	(9,14; 11,85)		
CPPS_Va	Mediana	14,521	14,420	14,583	0,826	
	(Min, Max)	(10,87; 18,20)	(11,03; 18,20)	(10,87; 17,45)		

Legenda: LN – Leitura em intensidade normal; LFr – Leitura em intensidade fraca; LFo – Leitura em intensidade forte; Va – vogal [a] sustentada; Teste t – Teste t-studente para amostras independentes; Teste de Wilcoxon-Mann Whitney; *p<0,05

Segundo a Tabela 10 é possível afirmar que existem diferenças significativas entre os géneros para CPP_LFr, CPP_LN, CPP_LFo e CPPS_LN.

Os valores de CPP são mais elevados nos homens, mesmo na vogal [a], embora estatisticamente não seja possível confirmar que existam diferenças significativas. Os valores de CPPS, embora só sejam estatisticamente diferentes na tarefa de LN, são mais elevados nas mulheres, à exceção da tarefa LFo.

A CATPCA que é apresentada nas tabelas que se seguem tem a finalidade de encontrar medidas robustas que permitam caracterizar vocalmente a voz feminina e a voz masculina.

As tabelas 12, 14 e 16 representam o *alpha de cronbach* que reflete a fiabilidade de cada dimensão, que é tanto melhor quanto maior o valor. Os gráficos das figuras 3, 4 e 5 representam os Scree Plots dos valores próprios (Eigenvalues) que permitem definir o número de componentes que devem ser selecionadas para cada amostra. As tabelas 11, 13 e 15 representam os pesos das diferentes medidas em cada dimensão. Uma vez que não é viável utilizar um grande número de variáveis, foram selecionadas as que são mais utilizadas na prática clínica e que aparecem mais frequentemente na literatura, tal como descrito no capítulo dos procedimentos de análise estatística.

Tabela 11- CATPCA – Homens

	Dimensões	
	1	2
alpha_LN	0,655	-0,411
H1H2 _{LTAS} _LN	0,369	0,839
Leq_LN	0,551	-0,561
f ₀ media_LN	0,860	-0,264
f ₀ DP_LN	0,868	-0,128
gama_LN	0,788	0,074
CPPS_Va	0,298	0,027
CPPS_LN	0,177	-0,722
jitter_Va	-0,892	-0,304
shimmer_Va	-0,598	-0,488
HNR_Va	0,731	0,144

■ 0,5-0,6 ■ 0,6-07 ■ 0,7-0,8 ■ 0,8-0,9

Tabela 12 – Alpha de Cronbach Homens

Dimensões	Alpha de Cronbach
1	0,870
2	0,589
Total	,942 ^a

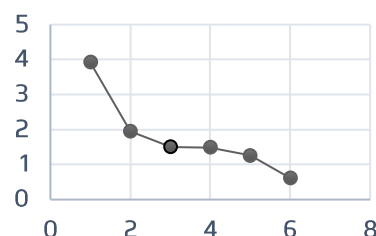


Figura 3- Scree Plot Eigenvalues Homens

De acordo com a regra do *eigenvalue* superior a 1 e com a análise do traçado presente no *Scree plot* foi possível resumir a informação em duas componentes. Na tabela 11 estão representadas as componentes que explicam 94,2% da variância total das variáveis originais do género masculino, assim como os “pesos” que cada variável tem na respetiva componente. É possível perceber que as variáveis que têm mais peso na primeira componente são a *f₀ media*, a *f₀ DP*, a *gama* e o *HNR*, enquanto que o *jitter* e o *shimmer* apresentam valores elevados nesta componente, mas no sentido oposto. Na segunda componente, a variável que apresenta mais peso é o *(H1H2)_{LTAS}* em oposição à *CPPS_LN*.

Tabela 13 – CATPCA – Mulheres

	Dimensões		
	1	2	3
alpha_LN	0,370	0,703	0,491
H1H2 _{LTAS} _LN	0,230	0,583	-0,152
Leq_LN	0,365	0,575	0,640
f ₀ media_LN	0,728	0,436	-0,414
f ₀ DP_LN	-0,416	0,791	-0,310
gama_LN	-0,354	0,624	-0,493
CPPS_Va	0,749	-0,534	0,248
CPPS_LN	0,523	0,548	0,435
jitter_Va	-0,792	0,335	0,339
shimmer_Va	-0,790	0,162	0,219
HNR_Va	0,571	0,237	-0,618

■ 0,5-0,6 ■ 0,6-07 ■ 0,7-0,8 ■ 0,8-0,9

Tabela 14 – Alpha de Cronbach Mulheres

Dimensões	Alpha de Chronbach
1	0,791
2	0,751
3	0,544
Total	,973 ^a

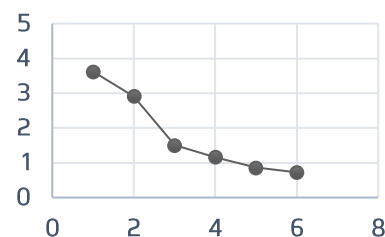


Figura 4- Scree Plot Eigenvalues Mulheres

Na tabela 13, correspondente às variáveis do género feminino, foi possível considerar 3 componentes, que explicam 97,3% da variação total. A primeira componente apresenta como variáveis com mais peso a *f₀*

média e o CPPS na LN em oposição ao *jitter* e ao *shimmer*. Na segunda componente as variáveis com mais peso são a f_0DP e o *alpha*. Por último, na terceira componente as variáveis que apresentam mais peso são o Leq em oposição ao HNR.

Tabela 15 – CATPCA – Homens+Mulheres

	Dimensões	
	1	2
<i>alpha</i> _LN	0,585	0,717
H1H2 _L TAS_LN	0,583	0,375
Leq_LN	0,348	0,787
f_0 media_LN	0,880	-0,081
f_0DP _LN	0,870	-0,106
gama_LN	0,735	-0,134
CPPS_Va	0,380	-0,250
CPPS_LN	0,524	0,584
<i>jitter</i> _Va	-0,793	0,381
<i>shimmer</i> _Va	-0,618	0,482
HNR_Va	0,677	-0,384

■ 0,5-0,6 ■ 0,6-0,7 ■ 0,7-0,8 ■ 0,8-0,9

Tabela 16- Alpha de Cronbach Homens+Mulheres

Dimensões	Alpha de Cronbach
1	0,869
2	0,609
Total	,943 ^a

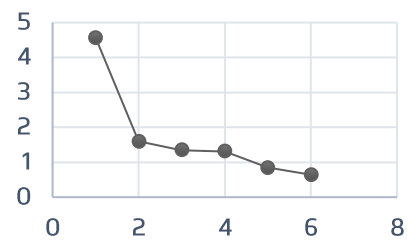


Figura 5- Scree Plot Eigenvalues Homens+Mulheres

Na tabela 15, na qual estão representadas as variáveis dos dois géneros em conjunto, as duas componentes identificadas explicam 94,3% da variação total. À semelhança do que foi possível observar na análise de componentes principais do género masculino, a primeira componente revela mais peso nos parâmetros relacionados com a f_0 e com o HNR em detrimento do *jitter*. Na segunda componente estão destacados os valores de *alpha* e de Leq.

5. Discussão

Quando é abordada a voz normal com o intuito de caracterizar o género, a frequência é o parâmetro que mais se destaca, por ser de fácil avaliação tanto acústica (f_0) como perceptiva (altura tonal da voz). No entanto, tal como a voz é mais do que a altura tonal, as diferenças vocais entre os géneros também vão mais além de uma voz mais grave ou mais aguda.

Relativamente aos valores aerodinâmicos, não foi possível verificar, nesta amostra, que os homens apresentam valores mais elevados do que as mulheres ao nível do TMF, ao contrário do que seria de esperar, tendo em conta os valores de referência consultados (Behlau, 2004; Hirano, 1981, Kent, Kent et al. 1987 e Zemlin 1998, citado por Santos, 2009). Apesar de, nesta amostra, alguns homens poderem ter um TMF mais reduzido, este fator não tem necessariamente de interferir na masculinidade ou feminilidade da voz. É importante ter em conta que o TMF é um parâmetro que é influenciado, para além da componente fonorrespiratória, pelas características fisiológicas próprias de cada indivíduo.

O coeficiente S/Z embora não tenha apresentado diferenças significativas apresenta um valor superior a 1 para o género masculino e inferior a 1 para o género feminino. Estes resultados, embora estejam dentro do intervalo definido para não serem considerados patológicos, são inferiores nas mulheres em comparação com os homens. Tendo em conta a literatura seria de esperar que o género feminino apresentasse menor coaptação glótica, pela presença mais frequente de fendas triangulares posteriores e por apresentar uma qualidade vocal com mais sopro (Guimaraes, 2002). No entanto, na amostra recolhida, ao apresentar uma mediana inferior a 1, nesta amostra, o género feminino revela maior quociente de fechamento das pregas vocais do que o género masculino, ao contrário do que seria de esperar (Behlau, 2004).

Ao proceder à comparação entre géneros para todas as tarefas foi possível perceber que efetivamente a f_0 é mais elevada nas mulheres, nos mesmo intervalos que estão já definidos na literatura (Tabela 1). Embora os valores obtidos não sejam estatisticamente significativos, através da análise dos dados descritivos (Tabela 8), foi possível encontrar nos homens valores mais elevados de *jitter* e *shimmer* e mais baixos de HNR, o que corrobora a literatura encontrada (Tabela 2). Este aspeto pode representar mais perturbação na voz do homem tanto ao nível da amplitude como da frequência do sinal, bem como diminuição da proporção entre harmónicos e ruído. No entanto, como os resultados estatísticos não são significativos, tanto neste trabalho como na literatura consultada, seria relevante explorar esta variabilidade, no sentido de encontrar uma justificação de base acústica e fisiológica para a mesma.

Em relação às medidas de LTAS as mulheres apresentam valores de $(H1H2)_{LTAS}$ superiores aos homens, o que é de esperar, uma vez que a voz da mulher é marcada por uma ligeira sopro e astenia que fazem aumentar a diferença entre o H1 e o H2. Os valores de *alpha*, por outro lado, não estão de acordo com o que a literatura indica para as diferenças entre os géneros. No género feminino são esperados

valores mais elevados acima dos 1000Hz, justificados pela maior presença de soproidade neste género, que em princípio origina uma inclinação espectral menos pronunciada. No entanto os valores encontrados sugerem uma menor inclinação espectral para o género masculino. Este facto também pode ser justificado pela presença mais acentuada de tensão nas vozes masculinas utilizadas para esta amostra (Mendoza et al., 1996).

Na comparação com base nos valores de CPP que, ao contrário do CPPS, deu resultados estatisticamente significativos, é possível observar que os valores de CPP são menores para o género feminino. Uma vez que, de acordo com a evidência encontrada, existe correlação entre os valores de CPP e o parâmetro áudio-percetivo “soproidade” (Brockmann-Bauser et al., 2019), este resultado valida a possível presença de soproidade na voz feminina, que pode ser justificada pela presença habitual de fendas triangulares posteriores neste género (Guimaraes, 2002).

Na perspetiva de uma avaliação por parte do TpF no género vocal, será relevante dar destaque aos parâmetros que são mais diferentes entre géneros, nomeadamente f_0 , f_0DP , $(H1H2)_{LTAS}$, $alpha$ e CPP. Uma vez que estas tarefas apresentam valores mais discrepantes entre géneros, ao usá-las na avaliação será mais fácil encontrar semelhanças ou desvios de acordo com o que é esperado para um determinado género.

A CATPCA realizada teve como pressuposto a identificação de medidas robustas para caracterizar os géneros, de forma a cumprir o segundo objetivo específico proposto para esta dissertação. Nos homens, a primeira componente, reflete uma relação de oposição entre f_0 média, as suas variações (DP e gama) e a harmonia (HNR) com a irregularidade fonatória que é representada pela medida de perturbação *jitter* (e também *shimmer*, embora com menor expressão). Esta relação sugere que as medidas relacionadas com a f_0 e com o HNR são maiores se os valores de perturbação da frequência da voz forem menores, ou seja, vozes mais periódicas serão também mais harmónicas e têm maior flexibilidade tonal. Nesta mesma componente, também aparece identificada a energia espectral ($alpha$) e a intensidade (Leq) como beneficiando do mesmo fenómeno, ou seja, quanto mais periódica a fonação, mais intensidade e fluidez (energia abaixo dos 1000 Hz) o sinal vocal adquire.

A segunda componente apresenta pesos inversos para $(H1H2)_{LTAS}$ e CPPS_LN. Não foram encontrados achados na literatura que justifiquem diretamente esta observação. Contudo, em oposição ao que está descrito na literatura, no qual a voz é tendencialmente mais soproosa para o género feminino, pode equacionar-se que, havendo no género masculino uma maior coaptação glótica (Araújo, 2015; Da Silva et al., 2011), surjam valores mais baixos de $(H1H2)_{LTAS}$ e, por consequência, por revelarem menor soproidade, valores de CPPS mais elevados. Ao compararmos estes resultados com o que foi possível encontrar na estatística descritiva é possível verificar que, embora se esteja a fazer esta relação com a

voz mais soprosa na mulher, é possível verificar que os valores de CPPS, ainda assim, são mais elevados nas mulheres. No entanto, como foi apresentado no teste de hipóteses o CPPS não apresenta diferença significativa entre géneros. Já o CPP apresenta diferenças e os valores são mais elevados nos homens. Desta forma, embora a literatura aponte o CPPS como preferencial para se correlacionar com vozes de qualidade alterada, o CPP pode ser mais relevante para distinguir o género, já que não estamos à espera de graus de alteração muito elevados. Tendo em conta este aspeto seria importante utilizar o CPP na CATPCA e perceber quais as relações que se conseguem obter utilizando esta medida.

Com base na metodologia utilizada para este estudo, para estudar a voz masculina poderão ser mais importantes em primeiro lugar aspetos relacionados com as medidas de f_0 , HNR e *Jitter*, ou seja frequência e as suas variações, harmonia e regularidade vibratória. Em segundo lugar é importante ter em consideração a coaptação glótica e a qualidade vocal, nomeadamente o parâmetro de soprosidade.

A primeira componente da CATPCA para o género feminino revela valores discrepantes com bastante peso para f_0 e CPPS_Va em contraste com *shimmer* e *jitter*. Isto reflete uma relação inversa entre a voz aguda e menos soprosa (ou melhor qualidade vocal) com a regularidade de vibração da fonte glótica.

A segunda componente do género feminino sugere relação positiva entre a variabilidade da f_0 (f_0DP e *gamma*), a fluidez (*alpha*, $(H1H2)_{LTAS}$), intensidade (Leq) e qualidade vocal (CPPS na LN). A existência de valores em oposição de CPPS na tarefa de vogal sustentada não é algo expectável e torna difícil a interpretação e justificação desta segunda componente.

A análise das medidas que têm mais peso no terceiro componente sugere que na mulher há uma relação inversa entre intensidade (Leq) e harmonia na voz (HNR). Ou seja, uma voz com intensidade mais fraca pode apresentar maiores níveis de harmonia, enquanto que uma voz com intensidade mais forte pode apresentar uma voz com mais ruído. Neste caso, além de não se relacionar com o que foi encontrado na segunda componente, não é possível encontrar relação baseada na literatura. Neste ponto é importante destacar que se está a relacionar medidas que são obtidas em tarefas distintas (LN e Va) e que, por essa, razão podem afetar os resultados obtidos.

Embora as componentes associadas ao género feminino não sejam facilmente justificáveis, é possível resumir que numa voz feminina as medidas mais relevantes são f_0 , CPPS_LN, *jitter* e *shimmer* ou seja, frequência, qualidade vocal (soprosidade) e regularidade vibratória.

Quanto aos marcadores mais predominantes em toda a amostra, a Análise de Componentes Principais permite identificar na primeira componente, à semelhança do género masculino, uma relação entre as medidas de f_0 e HNR com a medida de perturbação *jitter* e *shimmer*, embora a primeira tenha mais peso.

Na segunda componente destacam-se pesos significativos positivos para Leq e α . Segundo a literatura, em tarefas em intensidade normal o α é mais elevado, no entanto se a intensidade (Leq) aumentar é esperado que o α diminua, por apresentar maior força de coaptação glótica e por sua vez maior energia abaixo dos 1000 Hz. Desta forma esta variação positiva encontrada não está de acordo com a literatura (Da Silva et al., 2011; Ng et al., 2009; Xie et al., 2016).

Uma vez que a primeira componente da amostra total está fortemente relacionada com a primeira componente do género masculino, é possível verificar, que na amostra utilizada a maior variabilidade é explicada por marcadores mais semelhantes ao género masculino. Não é possível, contudo, garantir que estes marcadores sejam de facto ancorados à caracterização do género, ou simplesmente indicadores de qualidade vocal normal, já que este era um pressuposto associado à seleção da amostra. Hipotetizando sobre a possibilidade de estarmos perante um padrão vocal predominantemente masculino, esta tendência, embora restrita à amostra e metodologia utilizada neste estudo, permite expor várias questões relacionadas com o comportamento vocal das mulheres:

- Será que as mulheres utilizam um padrão mais masculino? De forma consciente ou inconsciente? Por que motivo?;
- Será que este uso pode estar relacionado com questões de papéis sociais e profissionais?;
- Será que este padrão é mais utilizado em diferentes faixas etárias ou em determinados contextos sociodemográficos?

Estas questões podem ser importantes quando se a fala da procura da Terapia da Fala para melhorar as competências comunicativas, principalmente em profissionais da voz, onde pode ser necessária uma adaptação vocal aos requisitos que a mulher profissional apresenta, nomeadamente em questões relacionadas com o papel que exerce na sua profissão, tais como: cargos políticos, ensino, chefia de empresas, entre outros que implicitamente exijam um cargo de poder. Apesar de já terem sido realizados estudos que utilizam a perceção da altura tonal que comprovam que uma mulher com voz mais grave confere uma perceção de maior competência e liderança (Klofstad et al., 2012), após analisar os resultados deste estudo seria importante utilizar outros parâmetros, nomeadamente que englobem a variabilidade da f_0 (DP e Gama), a relação entre harmonia e ruído na voz, bem como o grau de perturbação. Neste sentido, podia ser estudado até que ponto estes parâmetros influenciam a perceção dos ouvintes, em relação à voz.

Entrando no tema da identidade de género tanto a mulher como o homem podem procurar padrões de fala distintos daqueles que lhe são inculcados geneticamente. Seria importante perceber se a procura espontânea por estes padrões pode ter algum efeito ao nível da saúde e conforto vocal. O TpF é o profissional mais indicado para intervir a este nível, estando evidente que existem mais marcadores para além da f_0 importantes para a caracterização de cada género, não sendo suficiente intervir ao nível da altura tonal do indivíduo.

Tal como foi referido, para este estudo foi utilizada uma base de dados já anteriormente recolhida, inicialmente suficiente para os objetivos propostos. Neste momento, seria importante utilizar uma amostra mais significativa da população de forma a evitar que possíveis variáveis confundidoras possam interferir com os resultados. Embora tenha sido realizada previamente uma avaliação perceptiva da voz e todos os indivíduos estarem dentro dos parâmetros de classificação da voz normal, surge a importância de caracterizar perceptivamente as vozes dos indivíduos do ponto de vista da feminilidade e masculinidade, ou seja, o quanto a voz dos indivíduos se aproxima de uma voz mais feminina ou mais masculina. Esta questão podia ser recolhida através de um questionário aos próprios indivíduos e através da avaliação áudio-perceptiva de um painel de ouvintes. Tal como é referido na literatura, a avaliação acústica deve ser sempre complementar à avaliação perceptiva (Behlau, 2004; Felipe et al., 2006). Neste caso, estando a explorar as medidas do ponto de vista da caracterização do género era importante ter um termo de comparação que ajudasse a validar os resultados observados.

Importante referir que, para a Análise de Componentes Principais, compararam-se medidas que são obtidas em tarefas distintas. Uma vez que a estatística foi realizada com base na análise multivariada de dados, permitiu-se utilizar este tipo de comparação para estudar possíveis relações. No entanto, não se verificando concordância entre tarefas, considera-se pertinente, num estudo futuro, utilizar apenas medidas que sejam obtidas na mesma tarefa.

Tendo em conta que os procedimentos utilizados foram bastante abrangentes a nível de medidas acústicas e tarefas, é importante destacar que, de forma a simplificar os procedimentos e basear a discussão numa perspetiva mais relacionada com a caracterização do género em específico, seria benéfico fazer uma seleção prévia das medidas e tarefas a serem utilizadas. Em relação às tarefas de longa duração, justifica-se a utilização apenas da Leitura em diferentes intensidades, enquanto que ao nível das tarefas de curto termo, apenas se justifica a utilização de uma das vogais, que por ser mais consistente, a mais usada é a vogal [a] sustentada. Em relação aos parâmetros a utilizar, com base nos resultados obtidos para esta amostra, podiam ser excluídos o Leq, o CPPS e as medidas aerodinâmicas.

De forma a enriquecer a informação obtida ao nível da caracterização do género vocal seria relevante, obter amostras de diferentes faixas etárias e diferentes características sociodemográficas, onde fossem explorados estes marcadores de género, como por exemplo: perceber se mulheres que ocupam cargos de liderança apresentam mais semelhanças com a voz dos homens e em que parâmetros é mais relevante essa diferença; perceber se a idade da mulher que ocupa o cargo de liderança tem efeito nestes parâmetros; ou perceber se existem parâmetros mais relevantes na caracterização do género em pessoas de diferentes regiões (meios urbanos *versus* meios rurais).

Seria relevante, também, a realização de estudos, tanto dentro da voz normal como patológica, que utilizassem indivíduos para os quais são importantes as diferenças entre géneros, como por exemplo:

perceber quais os parâmetros que sofrem mais alterações na voz dos indivíduos antes e depois da mudança de género, da intervenção num caso de puberfonia e/ou de edema de Reinke. Dentro destas opções seria importante perceber se as alterações acústicas observadas se refletem da mesma forma na avaliação audio-percetiva realizada por profissionais com experiência neste tipo de avaliação e se se relacionam com a auto-perceção do género por parte dos próprios indivíduos em estudo, não só a nível da fala, mas também ao nível psicoemocional e social.

6. Conclusão

Concluída a discussão dos resultados é possível dar algumas respostas aos objetivos definidos para este projeto, de forma a perceber de que forma este contribuiu para uma melhor caracterização do género vocal.

Desta forma, considerando o primeiro objetivo específico, que consiste em comparar as medidas acústicas utilizadas com base no género do falante, concluiu-se que, para a amostra utilizada, as medidas de f_0 média, f_0 DP, α e $(H1-H2)_{LTAS}$ são superiores no género feminino, enquanto que a extensão vocal e os valores de CPP são superiores no género masculino. Relativamente às medidas de perturbação da frequência (*jitter*) e intensidade (*shimmer*) os valores não são claros, no entanto são mais elevados nos homens, enquanto que o HNR é mais baixo. Esta relação aparece também na bibliografia encontrada, no entanto como os valores não são sinificativamente diferentes é necessário explorar melhor este aspeto. Ao nível das medidas aerodinâmicas, não foram encontradas diferenças entre géneros. Apesar de o valor de CPPS ser uma atualização ao CPP e mostrar melhores evidências ao nível da perturbação vocal, os seus valores não são estatisticamente significativos nas diferenças entre género, não demonstrando ser uma boa medida para avaliar esta componente. Já o CPP demonstra valores significativos e que corroboram a teoria encontrada como é o caso da presença de soprosidade na voz feminina, pelo que poderia ser explorada a sua utilização na voz normal.

No que respeita a encontrar medidas robustas para a caracterização do género quando se procura um padrão de fala mais masculino, além da f_0 , será importante ter em conta também a variabilidade da mesma (f_0 DP e gama), a relação harmonia e ruído (HNR), bem como a perturbação da voz (*jitter* e *shimmer*). Para um padrão de fala mais feminino serão importantes questões mais relacionadas com a f_0 , qualidade vocal (soprosidade) (CPPS) e índices de perturbação da voz (*jitter* e *shimmer*). Interpretando os resultados do ponto de vista de relação entre medidas é possível concluir que, nos homens, vozes mais periódicas demonstram ser vozes com f_0 mais elevada, maior variabilidade e maior harmonia. No género feminino vozes com menor perturbação apresentam maior f_0 e menor soprosidade.

Na tentativa de encontrar marcadores de género mais predominantes na amostra total foi possível perceber que as medidas que apresentam maior peso na componente estatística são semelhantes às medidas encontradas para o género masculino, podendo contudo ser meramente marcadores de qualidade vocal normal, sem especificação de género. Esta evidência requer uma melhor exploração e relação com a componente psicossocial de forma a perceber se existe alguma relação com o papel do género na sociedade

Apesar de, devido à dimensão amostral deste estudo, não ser possível dar indicações claras sobre quais as combinações de parâmetros mais robustas na caracterização do género, fica demonstrada a possibilidade de a investigação caminhar nesta direção. Com esta, identificam-se parâmetros pouco utilizados e qualidades por vezes negligenciadas. Futuramente, a identificação mais robusta de marcadores vocais de género poderá dar indicações mais claras sobre os parâmetros a considerar em terapia vocal direta, levando à seleção de técnicas e estratégias mais adequadas e a melhores resultados. Por exemplo, os níveis de adução (com efeito na soproidade e nos registos vocais), as inflexões vocais, a regularidade vibratória e a ressonância podem ser aspetos úteis, complementos à f_0 . Estes marcadores poderão ainda ser úteis na medição de resultados de intervenções, não apenas de terapia vocal, mas também em intervenções cirúrgicas na área da otorrinolaringologia, bem como de terapias combinadas.

Para além das características de análise audio-percetiva e acústica da voz é relevante incluir aspetos não vocais associados à comunicação que não foram considerados neste estudo, nomeadamente características articulatórias, linguísticas e não-verbais. No futuro, estes aspetos poderão merecer atenção, estudando-se a sua interação com os fenómenos vocais.

A presente dissertação permitiu identificar parâmetros acústicos importantes e mostrar que o estudo de combinações de parâmetros é útil na caracterização vocal do género. Permite explorar questões pertinentes para que mais estudos sejam realizados nesta área e demonstra a importância da contribuição da análise acústica para a prática clínica.

Embora não sejam muito abordadas na área da Terapia da Fala questões relacionadas com os indivíduos que procuram alterar o seu comportamento vocal a evidência científica pode contribuir para um maior conhecimento nesta área, despoletar a procura por parte dos indivíduos que procuram alterar o seu comportamento vocal e transmitir ao TpF segurança na intervenção, quer com esta população, como com as patologias que interferem na perceção vocal do género já conhecidas.

Referências Bibliográficas

- Almeida, S. C. de, Mendes, A. P., & Kempster, G. B. (2019). The Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) Psychometric Characteristics: II European Portuguese Version (II EP CAPE-V). *Journal of Voice*, 33(4), 582.e5-582.e13. <https://doi.org/10.1016/j.jvoice.2018.02.013>
- Ambreen, S., Bashir, N., Tarar, S. A., & Kausar, R. (2019). Acoustic analysis of normal voice patterns in Pakistani adults. *Journal of Voice*, 33(1), 124.e49-124.e58. <https://doi.org/10.1016/j.jvoice.2017.09.003>
- APA. (2015). Guidelines for psychological practice with transgender and gender nonconforming people. *The American Psychologist*, 70(9), 832-864. <https://doi.org/10.1037/a0039906>
- Arán, M., Zaidhaft, S., & Murta, D. (2008). Transexualidade: corpo, subjetividade e saúde coletiva. *Psicologia & Sociedade*, 20(1), 70-79. <https://doi.org/10.1590/s0102-71822008000100008>
- Araújo, P. A. G. M. (2015). *Describing Vocal Performance of Portuguese Cantadeiras of Alto Minho*. Tese de Doutorado. Departamento de Comunicação e Arte: Universidade de Aveiro.
- Araújo, S. A., Grellet, M., Pereira, J. C., & Rosa, M. O. (2002). Normatização de medidas acústicas da voz normal. *Revista Brasileira de Otorrinolaringologia*, 68(4), 540-544. <https://doi.org/10.1590/S0034-72992002000400014>
- ASHA. (n.d.). Speech-Language Pathologists. Retrieved January 23, 2021, from <https://www.asha.org/Students/Speech-Language-Pathologists/#careers>
- ASHA. (2006). Consensus Auditory-Perceptual Evaluation of Voice. *Asha*, 1-3.
- Awan, S. N., Giovinco, A., & Owens, J. (2012). Effects of vocal intensity and vowel type on cepstral analysis of voice. *Journal of Voice*, 26(5), 670.e15-670.e20. <https://doi.org/10.1016/j.jvoice.2011.12.001>
- Baken, R. J., & Orlikoff, R. F. (2000). *Clinical Measurement of Speech and Voice*. San Diego: Singular Publishing Group.
- Barros, A. D. (2017). *A Relação Entre a Voz e Expressão de Gênero : a percepção de pessoas transexuais*. Dissertação de Mestrado. Universidade de Brasília. Retrieved from https://repositorio.unb.br/bitstream/10482/31164/1/2017_AlanaDantasBarros.pdf
- Beber, B. C., & Cielo, C. A. (2011). Características vocais acústicas de homens com voz e laringe normal, 13(1), 340-351. <https://doi.org/10.1590/S1516-18462010005000040>
- Behlau, M. (2004). *Voz: o livro do especialista* (Vol. 1). Rio de Janeiro: Revinter
- Belo, I. (2015). *Valores de Referência de Parâmetros Acústicos para a voz normal no Português Europeu*. Dissertação de Mestrado. Secção Autónoma de Ciências da Saúde: Universidade de Aveiro
- Bergel, S. (1999). *Voz do Transexual Masculino*. Monografia. Centro de Especialização em Fonoaudiologia Clínica.
- Boersma, P., & David Weenink. (2021). Praat. Amsterdam - The Netherlands: Univrsity of Amsterdam.

Retrieved from www.praat.org

- Borkowska, B., & Pawlowski, B. (2011). Female voice frequency in the context of dominance and attractiveness perception. *Animal Behaviour*, *82*(1), 55–59. <https://doi.org/10.1016/j.anbehav.2011.03.024>
- Brasil, O. de O. C. do, Yamasaki, R., & Leão, S. H. de S. (2005). Proposta de medição da posição vertical da laringe em repouso. *Revista Brasileira de Otorrinolaringologia*, *71*(3), 313–317. <https://doi.org/10.1590/S0034-72992005000300008>
- Brinca, L. F., Batista, A. P. F., Tavares, A. I., Gonçalves, I. C., & Moreno, M. L. (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. <https://doi.org/10.1016/j.jvoice.2013.10.001>
- Brockmann-Bauser, M., Stan, J. H. V., Sampaio, M. C., 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. <https://doi.org/10.1016/j.jvoice.2019.11.015>
- Brockmann, M., Drinnan, M. J., Storck, C., & Carding, P. N. (2011). Reliable jitter and shimmer measurements in voice clinics: The relevance of vowel, gender, vocal intensity, and fundamental frequency effects in a typical clinical task. *Journal of Voice*, *25*(1), 44–53. <https://doi.org/10.1016/j.jvoice.2009.07.002>
- Carmo, H., & Ferreira, M. M. (2008). *Metodologia da investigação – Guia para Auto-aprendizagem* (2ª edição). Lisboa: Universidade Aberta
- Cielo, C. A., Beber, B. C., Maggi, C. R., Körbes, D., Oliveira, C. F., Weber, D. E., & Tusi, A. R. (2009). Disfonia funcional psicogênica por puberfonia do tipo muda vocal incompleta: aspectos fisiológicos e psicológicos. *Estudos de Psicologia (Campinas)*, *26*(2), 227–236. <https://doi.org/10.1590/s0103-166x2009000200010>
- Cleveland, T. F., Sundberg, J., & Stone, R. E. (2001). Long-term-average spectrum characteristics of country singers during speaking and singing. *Journal of Voice*, *15*(1), 54–60. [https://doi.org/10.1016/S0892-1997\(01\)00006-6](https://doi.org/10.1016/S0892-1997(01)00006-6)
- Da Silva, P. T., Master, S., Andreoni, S., Pontes, P., & Ramos, L. R. (2011). Acoustic and long-term average spectrum measures to detect vocal aging in women. *Journal of Voice*, *25*(4), 411–419. <https://doi.org/10.1016/j.jvoice.2010.04.002>
- Deary, I. J., Wilson, J. A., Carding, P. N., & MacKenzie, K. (2003). VoiSS: A patient-derived Voice Symptom Scale. *Journal of Psychosomatic Research*, *54*(5), 483–489. [https://doi.org/10.1016/S0022-3999\(02\)00469-5](https://doi.org/10.1016/S0022-3999(02)00469-5)
- Decreto Lei nº564/99 de 21 de dezembro de 1999 do Ministério da Saúde. Diário da República: Série I-A

- de 1999-12-21. Retrieved from <https://data.dre.pt/eli/dec-lei/564/1999/12/21/p/dre/pt/html>
- Delgado-Hernández, J., León-Gómez, N. M., Izquierdo-Arteaga, L. M., & Llanos-Fumero, Y. (2018). Cepstral analysis of normal and pathological voice in Spanish adults. Smoothed Cepstral Peak Prominence in sustained vowels versus connected speech. *Acta Otorrinolaringologica (English Edition)*, 69(3), 134–140. <https://doi.org/10.1016/j.otoeng.2017.05.002>
- Delgado, J., León, N., Jiménez, A., & Izquierdo, L. (2017). Análisis acústico de la voz: medidas temporales, espectrales y cepstrales en la voz normal con el Praat en una muestra de hablantes de español. *Revista de Investigación En Logopedia*, 2(2), 108–127. <https://doi.org/10.5209/RLOG.58191>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573–598. <https://doi.org/10.1037/0033-295X.109.3.573>
- Feinberg, D. R., Jones, B. C., Little, A. C., Burt, D. M., & Perrett, D. I. (2005). Manipulations of fundamental and formant frequencies influence the attractiveness of human male voices. *Animal Behaviour*, 69(3), 561–568. <https://doi.org/10.1016/j.anbehav.2004.06.012>
- Felippe, A. C. N. de, Grillo, M. H. M. M., & Grechi, T. H. (2006). Normatização de medidas acústicas para vozes normais. *Revista Brasileira de Otorrinolaringologia*, 72(5), 659–664. <https://doi.org/10.1590/S0034-72992006000500013>
- Ferrand, C. T. (2002). Harmonics-to-noise ratio: An index of vocal aging. *Journal of Voice*, 16(4), 480–487. [https://doi.org/10.1016/S0892-1997\(02\)00123-6](https://doi.org/10.1016/S0892-1997(02)00123-6)
- Filho, J. A. X., Melo, E. C. M. de, Carneiro, C. de G., Tsuji, D. H., & Sennes, L. U. (2003). Correlação entre a altura e as dimensões das pregas vocais. *Revista Brasileira de Otorrinolaringologia*, 69(36), 371–374. <https://doi.org/10.1590/S0034-72992003000300012>
- Freitas, S. A. V. dos S. (2012). *Avaliação Acústica e Áudio Percetiva na Caracterização da Voz Humana*. Tese de Doutoramento. Faculdade de Engenharia da Universidade do Porto.
- Guimarães, I. (2007). *A Ciência e a Arte da Voz Humana*. Escola Superior de Saúde de Alcoitão: Santa Casa da Misericórdia de Lisboa.
- Guimarães, I., & Abberton, E. (2005). Fundamental frequency in speakers of Portuguese for different voice samples. *Journal of Voice*, 19(4), 592–606. <https://doi.org/10.1016/j.jvoice.2004.11.004>
- Guimaraes, I. C. R. P. (2002). *An electrolaryngographic study of dysphonic Portuguese speakers*. Tese de Doutoramento. University College London
- 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. <https://doi.org/10.1016/j.jvoice.2012.08.008>
- 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. [https://doi.org/10.1016/S0892-1997\(02\)00067-X](https://doi.org/10.1016/S0892-1997(02)00067-X)

- Hogikyan, N. D., & Sethuraman, G. (1999). Validation of an instrument to measure voice-related quality of life (V-RQOL). *Journal of Voice*, 13(4), 557–569. [https://doi.org/10.1016/S0892-1997\(99\)80010-1](https://doi.org/10.1016/S0892-1997(99)80010-1)
- Jacobson, B. H., Johnson, A., Grywalski, C., Silbergleit, A., Jacobson, G., Benninger, M. S., & Newman, C. W. (1997). The Voice Handicap Index (VHI): Development and Validation. *American Journal of Speech-Language Pathology*, 6(3), 66–69. <https://doi.org/10.1044/1058-0360.0603.66>
- Jesus, L. M. T., Tavares, A. I., & Hall, A. (2017). Cross-Cultural adaption of the GRBAS and CAPE-V scales for Portugal and a new training programme for perceptual voice evaluation. *Advances in Speech-Language Pathology*. <https://doi.org/10.5772/intechopen.69644>
- Klatt, D. H., & Klatt, L. C. (1990). Analysis, synthesis, and perception of voice quality variations among female and male talkers. *The Journal of the Acoustical Society of America*, 87(2), 820–857. <https://doi.org/10.1121/1.398894>
- Klofstad, C. A., Anderson, R. C., & Peters, S. (2012). Sounds like a winner: Voice pitch influences perception of leadership capacity in both men and women. *Proceedings of the Royal Society B: Biological Sciences*, 279(1738), 2698–2704. <https://doi.org/10.1098/rspb.2012.0311>
- 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. <https://doi.org/10.1016/j.jvoice.2008.03.008>
- Leung, Y., Oates, J., & Chan, S. P. (2018). Voice, articulation, and prosody contribute to listener perceptions of speaker gender: A systematic review and meta-analysis. *Journal of Speech, Language, and Hearing Research*, 61(2), 266–297. https://doi.org/10.1044/2017_JSLHR-S-17-0067
- Linville, S. E. (2002). Source characteristics of aged voice assessed from long-term average spectra. *Journal of Voice*, 16(4), 472–479. [https://doi.org/10.1016/S0892-1997\(02\)00122-4](https://doi.org/10.1016/S0892-1997(02)00122-4)
- Marôco, J. (2011). *Análise Estatística com o SPSS Statistics* (5ª edição). Pêro Pinheiro.
- 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*, 111–120. <https://doi.org/10.1590/S0104-56872006000100013>
- Master, S., De Biase, N., Chiari, B. M., & Laukkanen, A. M. (2008). Acoustic and perceptual analyses of brazilian male actors' and nonactors' voices: Long-term Average Spectrum and the "Actor's Formant." *Journal of Voice*, 22(2), 146–154. <https://doi.org/10.1016/j.jvoice.2006.09.006>
- Mateus, M. H. M., Falé, I., & Freitas, M. J. (2005). *Fonética e Fonologia do Português* (1ª edição). Lisboa: Universidade Aberta.
- Mazzetto de Menezes, K. S., Master, S., Guzman, M., Bortnem, C., & Ramos, L. R. (2014). Diferencias de los parámetros acústicos y perceptuales de la voz entre sujetos ancianos y jóvenes con intensidad habitual y alta. *Acta Otorrinolaringologica Espanola*, 65(2), 76–84.

<https://doi.org/10.1016/j.otorri.2013.07.009>

- Mendes, A. ., & Castro, E. (2005). Análise Acústica da Avaliação Vocal I: Tarefas Fonatórias e Medidas Acústicas. *Rev. Port. ORL*, *43*, 127–136. <https://doi.org/10.34631/sporl.517>
- Mendes, A. P., Costa, A. N., Martins, A. D., Fernandes, A. F., Vicente, S. M., & Freitas, T. C. (2011). Contributos para a construção de um texto foneticamente equilibrado para o Português-Europeu. *CEFAC*. <https://doi.org/10.1590/S1516-18462012005000030>
- Mendes, A. P., Moreira, M., Costa, A., Murtinheira, A., Jorge, A., & Araújo, M. (2015). Construção de um Instrumento de Triagem de Perturbações Motoras da Fala: standardização do texto foneticamente equilibrado "O Sol." *Revista Ibero-Americana de Saúde e Envelhecimento*, *1*(1), 5–20.
- Mendoza, E., Valencia, N., Muñoz, J., & Trujillo, H. (1996). Differences in voice quality between men and women: Use of the long-term average spectrum (LTAS). *Journal of Voice*, *10*(1), 59–66. [https://doi.org/10.1016/S0892-1997\(96\)80019-1](https://doi.org/10.1016/S0892-1997(96)80019-1)
- Meyer, D. E. (2013). Género e Educação: teoria e política. In Vozes (Ed.), *Corpo, género e sexualidade: um debate contemporâneo na educação* (9ª edição). Rio de Janeiro: Petrópolis.
- Mezzedimi, C., Di Francesco, M., Livi, W., Spinosi, M. C., & De Felice, C. (2017). Objective evaluation of presbyphonia: Spectroacoustic study on 142 patients with Praat. *Journal of Voice*, *31*(2), 257.e25–257.e32. <https://doi.org/10.1016/j.jvoice.2016.05.022>
- 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. <https://doi.org/10.1016/j.jvoice.2011.07.013>
- Ng, M. L., Liu, H., Zhao, Q., & Lam, P. K. Y. (2009). Long-term average spectral characteristics of Cantonese alaryngeal speech. *Auris, Nasus, Larynx*, *36*(5), 571–577. <https://doi.org/10.1016/j.anl.2008.12.005>
- Nogueira, C., Oliveira, J. M. de, Almeida, M. V. de, Costa, C. G., Rodrigues, L., & Pereira, M. (2010). *Estudo Sobre a Discriminação em Função da Orientação Sexual e da Identidade de Género* (Vol. 4). Lisboa: Comissão para a Cidadania e a Igualdade de Género.
- Nordenberg, M., & Sundberg, J. (2003). Effect on LTAS of vocal loudness variation. *TMH-QPSR*, *45*, 093–100. <https://doi.org/10.1080/14015430410004689>
- 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. <https://doi.org/10.1016/j.otoeng.2018.04.005>
- Pereira, A. M., Dassié-Leite, A. P., Pereira, E. C., Cavichiolo, J. B., Rosa, M. de O., & Fugmann, E. A. (2018). Auditory perception of lay judges about gender identification of women with Reinke's edema. *Codas*, *30*(4). <https://doi.org/10.1590/2317-1782/20182017046>

- Pereira, T. D. (2018). *Caracterização de Falantes do Português Europeu através do Espectro Médio de Longa Duração: comparação entre diferentes tarefas de fala*. (Projeto de Licenciatura não publicado) Instituto 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*. <https://doi.org/10.1016/j.jvoice.2018.09.010>
- Pinho, S. M. R. (2006). *Fundamentos em Laringologia e Voz*. Rio de Janeiro: Revinter.
- Pinho, S. M. R., & Pontes, P. (2002). Escala de Avaliação Percetiva da Fonte Glótica: RASAT. *Voxbrasilis*, 3, 11–13.
- Pinho, S. M. R., & Pontes, P. (2008). *Músculos Intrínsecos da Laringe e Dinâmica Vocal. (Série Desvendando os Segredos da Voz)*. (Vol.1). Rio de Janeiro: Revinter.
- Pontes, P. a. L., Vieira, V. P., Gonçalves, M. I. R., & Pontes, A. a. L. (2002). Características das vozes roucas, ásperas e normais: análise acústica espectrográfica comparativa. *Revista Brasileira de Otorrinolaringologia*, 68(2), 182–188. <https://doi.org/10.1590/S0034-72992002000200005>
- Putnoki, D. de S., Hara, F., Oliveira, G., & Behlau, M. (2010). Qualidade de vida em voz: o impacto de uma disfonia de acordo com gênero, idade e uso vocal profissional. *Revista Da Sociedade Brasileira de Fonoaudiologia*, 15(4), 485–490. <https://doi.org/10.1590/s1516-80342010000400003>
- Rachel, S. S., Snekhalatha, U., Vedhasorubini, & Balakrishnan, D. (2017). Spectral analysis and feature extraction of speech signal in dysphonia patients. *International Journal of Pure and Applied Mathematics*, 113(May), 151–160.
- Ribeiro, J. L. P. (2010). *Investigação e Avaliação em Psicologia e Saúde*. (2ª edição). Lisboa: L. Placebo Editora.
- Riesgo, C. A. F., & Nöth, E. (2019). What makes the Cepstral Peak Prominence different to other acoustic correlates of vocal quality? *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2019.01.004>
- Roubeau, B., Henrich, N., & Castellengo, M. (2009). Laryngeal vibratory mechanisms: The notion of vocal register revisited. *Journal of Voice*, 23(4), 425–438. <https://doi.org/10.1016/j.jvoice.2007.10.014>
- Santos, A. R. S. (2012). Contributo para a tradução e adaptação cultural da “VoiSS - Vocal Symptom Scale” para o Português Europeu. *Universidade Fernando Pessoa*. <http://hdl.handle.net/10284/3211>
- Santos, J. D. (2013). A Capacidade de Evocação Verbal em Pessoas com Afasia: estudo comparativo sobre a produção de nomes e verbos em três contextos de produção verbal. Dissertação de mestrado. *Universidade Católica Portuguesa*. Retrieved from <http://hdl.handle.net/10400.14/13977>
- Santos, R. J. (2009). *Avaliação de Pacientes com Paralisia Unilateral das Pregas Vocais*. Dissertação de Mestrado. *Universidade de Aveiro*. Retrieved from <http://hdl.handle.net/10773/3298>
- Silveira, E. M. C. (2006). *De Tudo Fica um Pouco: a construção social da identidade do transexual*. *Pontifícia Universidade Católica do Rio Grande do Sul*. Faculdade de Serviço Social. Retrieved from

<https://hdl.handle.net/10923/5180>

- Silvestre, I., Guimarães, I., & Teixeira, A. (2011). Qualidade vocal em mulheres com diagnóstico de nódulos vocais : Estudo preliminar Voice quality on women with vocal nodules diagnosis : Preliminary study. *Otorrinolaringologia e Cirurgia de Cabeça e Pescoço*, 49(1), 69–77. <https://doi.org/https://doi.org/10.34631/sporl.160>
- Spazzapan, E. A., Cardoso, V. M., Fabron, E. M. G., Berti, L. C., Brasolotto, A. G., & Marino, V. C. de C. (2018). Acoustic characteristics of healthy voices of adults: From young to middle age. *Codas*, 30(5), 1–7. <https://doi.org/10.1590/2317-1782/20182017225>
- SPTF. (2020). *Dicionário Terminológico de Terapia da Fala*. Lisboa: Papa-Letras
- Sundberg, J., Salomão, G. L., & Scherer, K. R. (2019). Analyzing emotion expression in singing via flow glottograms, Long-Term-Average Spectra, and expert listener evaluation. *Journal of Voice*. <https://doi.org/10.1016/j.jvoice.2019.08.007>
- Tanner, K., Roy, N., Ash, A., & Buder, E. H. (2005). Spectral moments of the long-term average spectrum: sensitive indices of voice change after therapy? *Journal of Voice : Official Journal of the Voice Foundation*, 19(2), 211–222. <https://doi.org/10.1016/j.jvoice.2004.02.005>
- Tavares, A. (2014). *Avaliação perceptiva da voz: GRBAS e CAPE-V*. Dissertação de Mestrado. Universidade de Aveiro. Retrieved from <http://hdl.handle.net/10773/13604>
- Teixeira, J. P., & Fernandes, P. O. (2014). Jitter, Shimmer and HNR classification within gender, tones and vowels in healthy voices. *Procedia Technology*, 16, 1228–1237. <https://doi.org/10.1016/j.protcy.2014.10.138>
- Titze, I. R. (1989). On the relation between subglottal pressure and fundamental frequency in phonation. *The Journal of the Acoustical Society of America*, 85(2), 901–906. <https://doi.org/10.1121/1.397562>
- Watts, C. R., & Awan, S. N. (2011). Use of spectral/cepstral analyses for differentiating normal from hypofunctional voices in sustained vowel and continuous speech contexts. *Journal of Speech, Language, and Hearing Research*, 54(December), 1525–1538. [https://doi.org/10.1044/1092-4388\(2011/10-0209\)analyses](https://doi.org/10.1044/1092-4388(2011/10-0209)analyses)
- Watts, C. R., Awan, S. N., & Maryn, Y. (2017). A Comparison of Cepstral Peak Prominence measures from two acoustic analysis programs. *Journal of Voice*, 31(3), 387.e1-387.e10. <https://doi.org/10.1016/j.jvoice.2016.09.012>
- Xie, S., Yan, N., Yu, P., Ng, M. L., Wang, L., & Ji, Z. (2016). Deep neural networks for voice quality assessment based on the GRBAS scale. *Interspeech*, 2656–2660. <https://doi.org/http://dx.doi.org/10.21437/Interspeech.2016-986>

Apêndices e Anexos

Apêndice 1

Instrumento de Recolha de Dados

CheckList_Procedimentos

Materiais		
(1) Ligar computador à corrente		
(2) Montar o tripé para o microfone (Posição e distância)		
(3) Ligar a placa de som ao computador		
(4) Ligar o microfone		
• Selecionar o microfone		
• Testar o microfone		
• Ganho de 0,5		
(5) Calibrar o microfone		
(6) Montar a estante (Posição e distância)		
Tarefas		Dist. Microfone
(1) Discurso espontâneo [DE]: questionário		15cm
a. Qual a sua profissão? Fale-me dela descrevendo-a.		
b. Fale-me dos seus interesses. O que gosta de fazer?		
c. Conte-me como é um dia de semana, desde que se levanta até que se deita, e descreva tudo o que faz.		
d. Diga-me o que pensa sobre a educação que hoje em dia as pessoas dão aos seus filhos comparativamente com a sua infância		
(2) Discurso semi-espontâneo [DSE]: descrição de figura(s)		15cm
a. “Ladrão de biscoitos” b. Carro dos bombeiros c. Pic-nic		
(3) Leitura de Frases [LF]		15 cm
(4) Leitura do texto “O Sol”		
a. Leitura em intensidade normal [LN]		15cm
b. Leitura em intensidade fraca [LFr]		7,5cm
c. Leitura em intensidade forte [LFo]		30cm
(5) Índice s/z		15 cm
a. Produção prolongada de fricativa não vozeada /s/ [ls]		
b. Produção prolongada de fricativa vozeada /z/ [lz]		
(6) Produção de vogais sustentadas, em tempo máximo de emissão		15cm
a. Vogal /a/ [Va]		
b. Vogal /i/ [Vi]		
c. Vogal /u/ [Vu]		
(7) Canto da música Parabéns [C]		15cm
(8) Produção de Glissandos		
a. Glissando ascendente [GA]		15 cm
b. Glissando descendente [GD]		7,5 cm

Código: _____

Nome: _____

Idade: _____ **Género:** _____

Profissão: _____

Nacionalidade: _____

Naturalidade: _____

Caracterização:

Postura:	Normal []	Alt. minor []	Alt. Major []
	Cifose___ Escoliose___ Lordose___ Ombros elevados___ Assimetria___		
	Extensão da cabeça___ Flexão da cabeça___ Flexão do corpo___		
Respiração:	Normal []	Alt. minor []	Alt. Major []
Modo:	Oral___ Nasal___ Misto___		
Tipo:	Costal superior___ Abdominal___ Costodiafragmática___		
Trato vocal:	Normal []	Alt. minor []	Alt. Major []
Fonação:	Normal []	Alt. minor []	Alt. Major []
	G (Gravidade Global) ___ R (Rouquidão) ___ B (Soprosidade) ___ A (Astenia) ___ S (Tensão)___		
Ressonância:	Normal []	Alt. minor []	Alt. Major []
	Equilibrada___ Hipernasal___ Hiponasal___ Faríngea___ Laringofaríngea___ <i>Cul de sac</i> ___		
Altura tonal:	Normal []	Alt. minor []	Alt. Major []
Intensidade:	Normal []	Alt. minor []	Alt. Major []
Articulação:	Normal []	Alt. minor []	Alt. Major []

Imagens

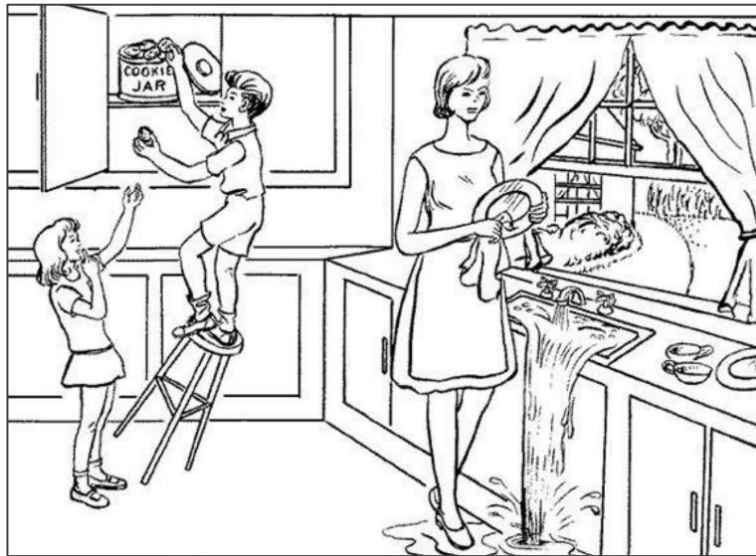


Figura 1 - Ladrão de Biscoitos (Goodglass & Kaplan, 1995)

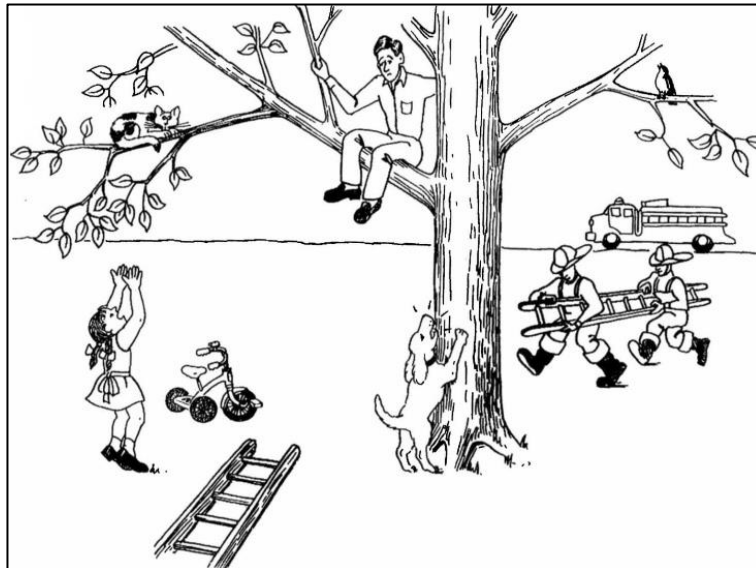


Figura 2 - Carro de Bombeiros

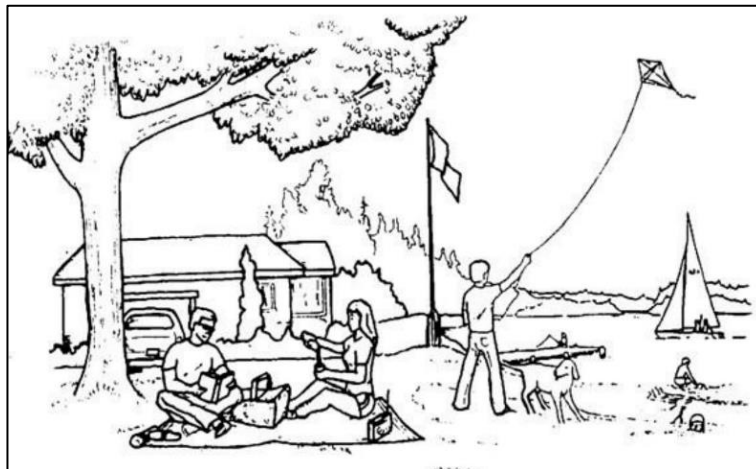


Figura 3 - Piquenique

Leitura de frases⁶

1. A Marta e o avô vivem naquele casarão rosa velho.
2. Sofia saiu cedo da sala.
3. A asa do avião andava avariada.
4. Agora é hora de acabar.
5. A minha mãe mandou-me embora.
6. O Tiago comeu quatro pêras.

Texto "O Sol"⁷

O Sol

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.

Texto "O Sol", 2009

Música "Parabéns a Você"

"Parabéns a você
Nesta data querida
Muitas felicidades,
Muitos anos de vida
Hoje é dia de festa
Cantam as nossas almas
Para o(a) menino(a)...
Uma salva de palmas"

⁶ Frases adaptadas de ASHA. Consensus Auditory-Perceptual Evaluation of Voice. ASHA. 2006;1-3.

⁷ Mendes AP, Costa NA, Martins AD, Fernandes AFP, Vicente SMR, Freitas TCS. Contributos para a construção de um texto foneticamente equilibrado para o Português-Europeu. Revista CEFAC. 2011.

Anexo 1
Parecer Comissão de Ética

ESCOLA SUPERIOR DE
TECNOLOGIA DA SAÚDE
DO PORTO
DATA 19 JUN 2017
N.º 2214
ENTRADA

P. PORTO

2º PARECER

PARECER DA COMISSÃO DE ÉTICA

1296/2017
Número de Registo da Comissão de Ética
12.06.2017 (2137/2017)
Data receção do Documento
Sim
Existência de entradas anteriores
TÍTULO DO TRABALHO
Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação
INVESTIGADOR RESPONSÁVEL
André Araújo
DATA PREVISTA PARA A REALIZAÇÃO DO TRABALHO
Início Out 2017, Fim: Dez 2018
RESUMO DO ESTUDO
OBJETIVOS
Nada a referir.
AMOSTRA
Nada a referir
FORMULÁRIO DE DADOS A RECOLHER
Refere a existência de "formulários em papel" mas não os apresenta
MATERIAL
Nada a referir (gravação áudio)
MÉTODOS
Nada a referir
RISCOS
Não existentes
CONSENTIMENTO INFORMADO
Conforme
AUTORIZAÇÃO PELOS RESPONSÁVEIS LOCAIS
Presentes e assinados, faltando apenas a do Presidente da ESS
APRECIÇÃO DA COMISSÃO DE ÉTICA
Consideram-se esclarecidas as questões suscitadas no 1º parecer
PARECER FINAL DA COMISSÃO DE ÉTICA
De acordo com os dados analisados, o parecer é favorável, mas salienta-se a necessidade de cumprimento de todas as diretrizes submetidas a esta Comissão, com prejuízo de a decisão ser suspensa caso haja algum incumprimento grave.
DATA: 13.06.2017



Anexo 2
Consentimento Informado

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

Desenvolvimento de métodos de avaliação acústica no estudo da voz com finalidade de uso em contexto clínico e em investigação

DESIGNAÇÃO DO ESTUDO

Eu, abaixo-assinado _____ :

Fui informado de que o Estudo de Investigação acima mencionado se destina a explorar e desenvolver a aplicabilidade de métodos inovadores de avaliação acústica na prática clínica da terapia da fala.

Sei que neste estudo está prevista a realização de gravações áudio e questionários tendo-me sido explicado em que consistem e quais os seus possíveis efeitos.

Foi-me garantido que todos os dados relativos à identificação dos Participantes neste estudo são confidenciais e que será mantido o anonimato.

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

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

Aceito participar de livre vontade no estudo acima mencionado.

Também autorizo a divulgação dos resultados obtidos no meio científico, garantindo o anonimato.

Nome do Investigador e Contacto:

PEDRO ANDRÉ GUERREIRO MARTINS ARAÚJO

paa@ess.ipp.pt

DATA

ASSINATURA



