

FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO



FEUP

Correspondência entre alterações da voz e do eletrocardiograma

Ricardo Jorge Macedo Pereira de Sousa Dias

PARA APRECIÇÃO POR JÚRI

Mestrado Integrado em Engenharia Electrotécnica e de Computadores

Orientador: Prof. Dr. Aníbal João de Sousa Ferreira

Co-orientador: Prof. Manuel dos Santos Barbosa (FPCEUP)

26 de Junho de 2012

Resumo

A voz é um dos principais meios de interação entre os seres humanos. Este órgão de comunicação tem sido alvo dos mais variados estudos, de que são exemplo a identificação de patologias ou a identificação de oradores em contextos forenses. Este estudo visa estabelecer um conjunto de correspondências entre variações existentes nos sinais de voz e de eletrocardiografia. Esta comparação foi feita em dois contextos distintos: em situação de repouso e mediante a apresentação de um conjunto de estímulos visuais de várias naturezas.

Tendo como ponto de partida comum os mecanismos neurológicos que desencadeiam os movimentos nas pregas vogais e no coração, pretende-se chegar a um conjunto de conclusões inovadoras sobre a forma como a voz nos diz algo sobre o estado do nosso coração. Cruzando os domínios do processamento de sinal e da medicina, procura-se compreender melhor o modo como as pessoas interagem e sentem. Porque expressar emoções a partir da fala é, em larga medida, o que faz de nós humanos.

Foi feita uma recolha de amostras em vinte sujeitos, divididos igualmente entre sexo masculino e feminino. Os estímulos apresentados durante as recolhas são integrantes do *International Affective Picture System*, sendo alvo de uma marcação distintiva individual. Os sinais de voz analisados consistiram nas vogais sustentadas /a/ e /i/. Depois da organização da base de dados de sinais, a implementação teve por objetivo a extração de características de sinal de ECG (*jitter*, batimentos cardíacos, período médio e outros) e de voz (*jitter*, *shimmer*, frequência fundamental, *Harmonic-to-noise ratio* e outros).

Efetuiu-se uma análise estatística dos resultados, recorrendo a um conjunto de descritores e testes apropriados para o espaço amostral em questão. Foram confrontadas diversas situações e comprovaram-se relações associadas a tendências coincidentes nas distribuições destes sinais, nomeadamente no que diz respeito a variações no *jitter* e na frequência fundamental. Também foi possível tirar conclusões sobre a influência do género nestas variações. Uma análise mais detalhada destes sinais, direcionada para uma vertente patológica, poderá contribuir para um estudo de relações ao nível do diagnóstico médico.

Abstract

Voice is one of the most important means of interaction between humans. This organ of communication has been the target of a variety of studies, such as the detection of pathological voices or the identification of speakers in forensic contexts. This study aims to establish a set of correspondences between variations existing in the speech and eletrocardiography signals. This comparison will be made in two different contexts: in the resting condition and upon presentation of a set of visual stimuli with different natures.

Taking as the common starting point the neurological mechanisms that originate movements in the vocal chords and heart, the study intends to draw a set of innovative conclusions about how the voice tells us something about the state of our heart. Interconnecting the fields of signal processing and medicine, this research wants to better understand how people interact and feel. Because the expression of emotions through the voice plays a huge part on the definition of our identity as human beings.

Samples were taken in twenty subjects, divided equally between male and female. The featured stimuli are part of the International Affective Picture System and to all samples was attributed an individual marker. The voice signals consisted in sustained vowels /a/ and /i/. After organizing the signal database, the implementation was made with the purpose of feature extraction on the ECG (jitter, beats per minute, average period and others) and on the voice signals (jitter, shimmer, Harmonic-to-noise ratio and others).

A statistical analysis was performed, using a set of descriptors and suitable statistical tests for the sample used. Several situations were compared and correspondences were detected, specially with coincident trends regarding the signal probability distributions for the values of jitter and fundamental frequency. Conclusions were also drawn about the influence of gender on these variations. A more detailed signal analysis, on a pathological dimension, may contribute to a stydy of relationships on a medical diagnosis level.

Agradecimentos

Agradeço em primeiro lugar aos meus pais e irmã, que sempre estiveram ao meu lado ao longo deste caminho cheio das mais variadas emoções. Agradeço igualmente aos meus avós e familiares próximos pelo entusiasmo com que encararam esta fase da minha vida. Por este sentimento de orgulho e por fazerem de mim aquilo que sou hoje, um muito obrigado.

Quero agradecer ao Professor Doutor Aníbal Ferreira pela orientação nesta dissertação e pela oportunidade que me deu de trabalhar neste tema. Uma palavra de imenso agradecimento vai também para o Doutor Ricardo Sousa pela sua disponibilidade, apoio e paciência no esclarecimento das dúvidas ao longo destes meses. A todas as pessoas com quem partilhei o local de trabalho, pelo bom ambiente de criado. Em especial, ao Mário Barbosa pela ajuda e pelo excelente espírito colaborativo desenvolvido na recolha e processamento da informação para este estudo. Ao Daniel Gomes pela ajuda na revisão da escrita da dissertação.

Não poderei deixar de agradecer igualmente ao co-orientador, Professor Fernando Barbosa, da Faculdade de Psicologia e Ciências da Educação da Universidade do Porto, quer pela cedência do local para recolhas no Laboratório de Neuropsicofisiologia (com o devido consentimento da *Lab Manager* Cassilda Reis), como também pela ajuda prestada noutros temas relacionados com a área da psicofisiologia. A todos os voluntários que deram o seu contributo para o estudo.

Por último, mas não menos importante, a todos com quem partilhei esta viagem e a quem, mesmo de fora, me motivou sempre que necessário. A vossa ajuda foi fundamental.

Ricardo

*“Para ser grande, sê inteiro: nada
Teu exagera ou exclui.*

*Sê todo em cada coisa. Põe quanto és
No mínimo que fazes.*

*Assim em cada lago a lua toda
Brilha, porque alta vive”*

Ricardo Reis

Conteúdo

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Metodologia	2
1.4	Estrutura	3
2	Revisão bibliográfica	5
2.1	Caracterização do eletrocardiograma	5
2.1.1	Enquadramento fisiológico	6
2.1.2	Aplicabilidade	6
2.1.3	Eventos	7
2.1.4	Erros no eletrocardiograma	9
2.1.5	Dispositivos utilizados	9
2.1.6	Exame típico	9
2.1.7	Exemplos de anomalias	10
2.2	Caracterização de parâmetros do som	11
2.2.1	Intensidade sonora	11
2.2.2	Frequência fundamental (F0) e formantes (F1, F2... Fn)	11
2.2.3	<i>Pitch</i> , timbre e intensidade subjetiva	11
2.2.4	Representação e análise de Fourier	12
2.3	Produção e percepção da voz	13
2.3.1	Percepção auditiva	13
2.3.2	Produção de voz	13
2.3.3	Modos e pontos de articulação	14
2.3.4	Mecanismos de fonação	14
2.3.5	Modelo Fonte-Filtro	14
2.4	Medição de parâmetros de voz	14
2.4.1	Critérios de avaliação	15
2.4.2	<i>Jitter</i>	15
2.4.3	<i>Shimmer</i>	15
2.4.4	<i>Harmonic-to-noise ratio</i> (HNR)	16
2.4.5	Tempo máximo de fonação	16
2.4.6	Estimação de impulsos glotais	16
2.5	Influência do sistema nervoso na voz e ritmo cardíaco	17
2.5.1	Nervo laríngeo recorrente	17
2.5.2	Sistema nervoso simpático e parassimpático	18
2.6	Avanços na avaliação de distorções na voz e no ECG	19
2.7	Considerações finais	19

3	Ferramentas utilizadas	21
3.1	Recolha de amostras	21
3.1.1	Biopac MP35	21
3.1.2	Biopac Student Lab (BSL)	22
3.1.3	Adobe Audition CS5.5	22
3.1.4	Microfone Sennheiser EARSET 1	22
3.1.5	Roland UA 25EX	23
3.1.6	<i>International Affective Picture System (IAPS)</i>	23
3.1.7	Material adicional	23
3.2	Processamento da Informação	24
3.2.1	Algoritmia de Processamento de Sinal	24
3.2.2	Praat	24
3.2.3	Voice Studio	24
3.2.4	Statistic Package for the Social Sciences (SPSS)	25
3.3	Considerações finais	25
4	Procedimento Experimental	27
4.1	Local de Recolha	27
4.2	Recolha de Voluntários e caracterização da amostra	27
4.3	Consentimento Informado	29
4.4	Escolha de imagens do IAPS	29
4.5	Exame de eletrocardiograma	29
4.6	Protocolo experimental	30
4.6.1	Fases do protocolo	31
4.6.2	Tempo de exposição dos estímulos e intervalo interestímulos	32
4.6.3	Marcação e segmentação de estímulos	32
4.7	Considerações finais	32
5	Processamento de Sinais	33
5.1	Sinais de eletrocardiografia	33
5.1.1	Principais funcionalidades	33
5.1.2	Extração de Picos (método automático)	35
5.1.3	Extração de Picos (método semi-automático usando <i>peakfinder</i>)	35
5.1.4	Extração de Picos (método semi-automático dos limiares)	36
5.1.5	Parâmetros da análise de sinais de eletrocardiografia	38
5.2	Sinais de voz	39
5.3	Considerações finais	40
6	Análise estatística de resultados	41
6.1	Testes da qualidade da amostra	41
6.2	Análise descritiva	42
6.2.1	Sinais de eletrocardiograma	42
6.2.2	Sinais de voz	44
6.3	Testes de correlação	45
6.4	Análise de modelos de regressão	46
6.5	Comparações entre distribuições	47
6.6	Considerações finais	50

7 Conclusões	53
7.1 Satisfação dos objetivos	53
7.2 Principais dificuldades	54
7.3 Trabalho futuro	55
7.4 Observações finais	55
A Conjunto de especificações	57
A.1 Biopac MP35	57
A.2 Sennheiser Earset 1	58
A.3 Roland UA 25EX	59
B Descritores estatísticos de parâmetros de sinais de eletrocardiograma	61
C Descritores estatísticos de parâmetros de sinais de voz (vogal /a/)	65
D Descritores estatísticos de parâmetros de sinais de voz (vogal /i/)	71
E Consentimento informado	77
Referências	79

Lista de Figuras

1.1	Especificação da metodologia utilizada	2
2.1	Mecanismo de despolarização e repolarização.	7
2.2	Alguns Eventos de um ECG	8
2.3	Principais dispositivos de um ECG.	10
2.4	Espectrograma de um violino.	12
2.5	Anatomia do ouvido humano.	13
2.6	Representação de impulsos glotais (à esquerda) e fórmula de cálculo do RAP (à direita).	15
2.7	Representação de impulsos glotais (à esquerda) e fórmula de cálculo do APQ (à direita).	16
2.8	Localização dos nervos associados ao trato vocal.	17
2.9	Sistema nervoso autónomo.	18
3.1	Aparelho Biopac MP35.	22
3.2	Microfone Sennheiser Earset 1.	23
3.3	Exemplo de interface no software SPSS.	25
4.1	Local de recolha de amostras.	28
4.2	Distribuição das idades dos participantes.	28
4.3	Ligação dos elétrodos ao sistema Biopac MP35.	30
4.4	Ligação dos elétrodos para cada sujeito.	30
4.5	Sequência do protocolo, com indicação de marcadores.	31
5.1	Sinais de ECG para a parte 3 do protocolo experimental (vozeamento de vogais).	34
5.2	Funcionamento do algoritmo automático.	35
5.3	Funcionamento do algoritmo semi-automático usando <i>peakfinder</i>	36
5.4	Representação de picos de correlação provenientes do algoritmo semi-automático <i>peakfinder</i> (caso de sucesso).	36
5.5	Funcionamento do algoritmo semi-automático dos limiares.	36
5.6	Deteção de picos pelo método dos limiares.	37
5.7	Complexos QRS após deteção.	37
5.8	Exemplo de deteção indevida de ondas T no algoritmo da reta em sinal de ECG com ruído.	38
5.9	Sinal de voz para análise.	39
5.10	Exemplo de <i>Voice Report</i> do Praat.	40
6.1	Funções de distribuição do PPQ5 para situação de repouso (esquerda) e com apresentação de estímulos (direita).	43

6.2	Exemplo de matriz de coeficientes de correlação (para relação entre parâmetros de ECG e voz).	46
6.3	Exemplo de modelo de regressão para relação entre PPQ5 de eletrocardiograma e voz (no vozeamento da vogal /i/).	48
6.4	Comparação de funções de distribuição para <i>jitter</i> , sem (N) e com (S) apresentação de estímulo.	48
6.5	Comparação das distribuições de valores da frequência fundamental, sem (N) e com (S) apresentação de estímulo.	49
6.6	Comparação das distribuições de valores do <i>shimmer</i> , sem (N) e com (S) apresentação de estímulo	50
A.1	Especificações do aparelho Biopac MP35.	57
A.2	Especificações do microfone Sennheiser Earset 1 (Gerais).	58
A.3	Especificações do microfone Sennheiser Earset 1 (Resposta em frequência e direcionalidade).	58
A.4	Especificações de Roland UA 25EX.	59
E.1	Consentimento informado apresentado a cada voluntário antes do início do teste	78

Lista de Tabelas

6.1	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a totalidade de amostra	43
6.2	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte I)	44
6.3	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte II)	44
B.1	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a totalidade de amostra	61
B.2	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a amostra do sexo masculino	62
B.3	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a amostra do sexo feminino	62
B.4	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a totalidade de amostra	62
B.5	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a amostra do sexo masculino	63
B.6	Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a amostra do sexo feminino	63
C.1	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a totalidade de amostra (Parte I)	65
C.2	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a totalidade de amostra (Parte II)	66
C.3	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo masculino (Parte I)	66
C.4	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo masculino (Parte II)	66
C.5	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo feminino (Parte I)	67
C.6	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo feminino (Parte II)	67
C.7	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte I)	67
C.8	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte II)	68
C.9	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo masculino (Parte I)	68

C.10	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo masculino (Parte II)	68
C.11	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo feminino (Parte I)	69
C.12	Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo feminino (Parte II)	69
D.1	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a totalidade da amostra (Parte I)	71
D.2	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a totalidade da amostra (Parte II)	72
D.3	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo masculino (Parte I)	72
D.4	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo masculino (Parte II)	72
D.5	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo feminino (Parte I)	73
D.6	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo feminino (Parte II)	73
D.7	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a totalidade da amostra (Parte I)	73
D.8	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a totalidade da amostra (Parte II)	74
D.9	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo masculino (Parte I)	74
D.10	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo masculino (Parte II)	74
D.11	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo feminino (Parte I)	75
D.12	Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo feminino (Parte II)	75

Abreviaturas e Símbolos

ANOVA	<i>Analysis of Variance</i>
APQ	<i>Amplitude Perturbation Quotient</i>
BPM	Batimentos por minuto
CAPE-V	Consensus Auditory-perceptual Evaluation of Voice
DFT	<i>Discrete Fourier Transform</i>
ECG	Eletrocardiograma
EGG	Estimação de impulsos glotais
ET	Estatística de teste
FFT	<i>Fast Fourier Transform</i>
GRBAS	Grade, roughness, breathiness, asteny, strain
HNR	<i>Harmonic-to-noise ratio</i>
IAPS	<i>International Affective Picture System</i>
NHR	<i>Noise-to-harmonic ratio</i>
PPQ	<i>Period Perturbation Quotient</i>
RAP	<i>Relative Average Perturbation</i>
SPSS	<i>Statistic Package for the Social Sciences</i>
SPT	<i>Speech Processing Toolbox</i>
WAV	<i>WAVEform audio format</i>
XLS	<i>Microsoft Excel Format</i>

Capítulo 1

Introdução

Ao longo da primeira secção deste documento será feita uma descrição do tema escolhido e dos principais objetivos a serem alcançados. Numa fase seguinte, será indicado um conjunto de abordagens a seguir para os atingir. Será também apresentada a estrutura de todo o documento.

1.1 Motivação

A voz é um dos principais meios de interação entre os seres humanos. Este órgão de comunicação tem sido alvo dos mais variados estudos, de que são exemplo a identificação de patologias ou a identificação de oradores em contextos forenses. A produção de voz é o resultado final de uma articulação sincronizada entre os órgãos constituintes do aparelho fonatório. O estudo de distorções associadas a sinais de voz implica o conhecimento de todo este mecanismo.

No âmbito desta dissertação, pretende-se estabelecer um conjunto de correspondências entre variações existentes nos sinais de voz e de eletrocardiografia. Esta comparação será feita em dois contextos distintos: em situação de repouso e mediante a apresentação de um conjunto de estímulos de natureza agradável, desagradável e neutra. Este estudo pretende conduzir a um conjunto de conclusões inovadoras sobre a forma como a voz nos diz algo sobre o funcionamento cardíaco.

Ao movimento dos músculos responsáveis pela fonação está associado o envio de impulsos ao longo do nervo laríngeo recorrente. Alterações neurológicas involuntárias desencadeadas pela exposição a determinados estímulos podem originar mudanças no envio desta informação e, por consequência, na natureza destes movimentos, alterando a frequência de vibração das pregas vocais. Esta vibração - a frequência fundamental da voz - é de mais simples identificação no caso da fonação ser feita com vogais sustentadas.

Outras alterações associadas aos sistemas nervosos simpático e parassimpático poderão ter como consequência a variação da frequência cardíaca ou da pressão arterial. Tendo como ponto de partida comum o nosso sistema nervoso, pretende-se encontrar uma relação entre as reações fisiológicas involuntárias que ocorrem simultaneamente nas pregas vocais e no coração.

Esta avaliação será feita tanto para o mesmo indivíduo, como para um conjunto de indivíduos, possibilitando tirar conclusões singulares e abrangentes relativamente a este conjunto de correspondências. Este estudo visa cruzar os domínios do processamento de sinal de voz e da medicina, levando a conclusões relativamente ao modo como as pessoas interagem e sentem. Porque expressar emoções a partir da fala é, em larga medida, o que faz de nós humanos.

1.2 Objetivos

Como foi indicado anteriormente, o principal objetivo do trabalho traduzido nesta dissertação é estabelecer um conjunto de ligações entre parâmetros de caracterização ou perturbação de voz e do eletrocardiograma. Para atingir este objetivo será feita, numa fase inicial, uma revisão bibliográfica onde constarão todos os conceitos relacionados com os sinais a estudar. Esta abordagem tem por objetivo conferir um conjunto de conhecimento essencial à compreensão do estudo, permitindo também um melhor enquadramento no tema para os leitores desta dissertação.

Numa fase seguinte, foi recolhido um conjunto de amostras de sinais de voz e ECG num grupo de voluntários. Esta recolha foi feita em dois contextos: em repouso e com apresentação de estímulos visuais (imagens). A metodologia destas tarefas experimentais será descrita detalhadamente mais à frente. Depois de recolhido o conjunto de amostras, proceder-se-á ao estudo destas distorções recorrendo a software específico, nomeadamente o Seegnal Voice Studio ou Praat.

Depois desta análise, será estabelecido um conjunto de correspondências entre estes parâmetros de sinal de voz e de eletrocardiograma. O conjunto concreto de parâmetros a analisar será especificado mais adiante. Será também feita uma comparação entre parâmetros dos mesmos sinais, confrontando a situação de repouso com a de apresentação de estímulos. Todas as relações terão um suporte estatístico para as comprovar, cuja especificação será feita mais à frente.

1.3 Metodologia

A metodologia utilizada para este estudo pode dividir-se em três fases fundamentais. A primeira consiste na recolha de amostras, para a qual é necessária a definição prévia de um procedimento experimental. Para que as recolhas sejam feitas com sucesso, este processo terá de ser bem pensado no que diz respeito à duração de cada segmento e ao despiste de outro tipo de problemas. Numa fase seguinte, será feito o processamento de sinais de voz e eletrocardiografia. Para o efeito, as ferramentas a utilizar serão o Praat e o Matlab, respetivamente. Depois das fases de recolha e processamento de sinais estarem concluídas, proceder-se-á a uma análise estatística dos resultados, utilizando o software *Statistic Package for the Social Sciences* (SPSS). Um esquema da metodologia utilizada pode ser visto na figura 1.1.

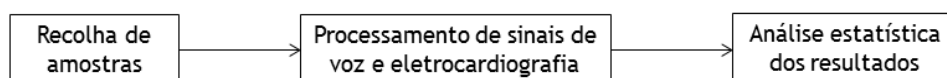


Figura 1.1: Especificação da metodologia utilizada

1.4 Estrutura

Esta dissertação divide-se em sete capítulos. O propósito da primeira foi já descrito no início deste capítulo. No capítulo 2 será feita a revisão bibliográfica para o tema escolhido, abordando todas as temáticas que ajudem à melhor compreensão do estudo de uma forma muito breve, para além dos avanços relativos ao tema propriamente dito. No capítulo 3, serão descritas as ferramentas utilizadas nas duas principais fases do estudo: a recolha e o processamento destes sinais.

O capítulo 4 especificará todo o procedimento experimental, justificando as escolhas feitas relativamente às tarefas experimentais efetuadas, à sua duração e ao local onde foi feita a recolha das amostras. Toda a explicação relativa ao processamento dos sinais de eletrocardiografia e voz encontra-se presente no capítulo 5, onde se confronta um conjunto de algoritmos para extração de parâmetros destes sinais.

No capítulo 6 utiliza-se a informação obtida anteriormente para fazer um tratamento estatístico, procurando justificar as relações encontradas utilizando os testes mais apropriados. Por último, o capítulo 7 é um balanço do estudo, verificando se os objetivos definidos foram cumpridos e identificando um conjunto de dificuldades encontradas.

Capítulo 2

Revisão bibliográfica

Tendo em conta que este estudo envolve o processamento de sinais de voz e de eletrocardiografia, importa primeiro fornecer um conjunto de conhecimentos fundamentais para a sua compreensão. Será feita uma breve caracterização do sinal eletrocardiográfico, através de um enquadramento que explique os processos fisiológicos que o fundamentam. Será também identificado um conjunto de situações comuns nas quais o exame é feito, bem como os seus principais eventos. Apesar da identificação destes eventos não ter, no contexto do processamento de sinal, relevância para o estudo, compreender a sua sequência é importante para o despiste de outro tipo de variações que apresenta (de que é exemplo o ruído), bem como para auxiliar à compreensão do fundamento teórico subjacente ao procedimento.

Os riscos associados ao eletrocardiograma são uma temática que não podia ser ignorada, bem como os dispositivos comumente utilizados. Por último, e apesar das diferenças relativamente ao tipo de recolha e montagem utilizados (cuja justificação será dada mais tarde), é descrito o modo como, tipicamente, a eletrocardiografia é levada a cabo e que tipo de anomalias se conseguem detetar através da sua análise.

Compreender a estrutura de um sinal de som e quais são seus principais parâmetros caracterizadores é um requisito importante para quem pretenda lidar com estes sinais. Depois desta breve explicação, é feito um enquadramento geral sobre como a voz é ouvida e produzida, bem como uma enumeração do conjunto de mecanismos que o permitem.

A secção que aborda a medição de parâmetros de voz complementa a anterior e é um dos suportes teóricos mais importantes desta dissertação. De seguida, é explicado como é que o nosso sistema nervoso reage a situações de maior ou menor stress e de que forma é que isso poderá refletir-se em alterações nos sinais de voz e eletrocardiográficos. Por último, são enumerados os estudos que, mesmo parcialmente, se relacionem com o estudo que foi feito.

2.1 Caracterização do eletrocardiograma

A eletrocardiografia foi inventada por Willen Einthoven, da Universidade de Leiden, na Holanda. Tal invenção valeu-lhe o prémio Nobel da Medicina em 1924. Trata-se de uma análise

transtorácica que mede a atividade eletromecânica no coração, durante um período de tempo, através de um conjunto de elétrodos em contacto com a pele. Esta técnica é utilizada para o diagnóstico e tratamento de praticamente todos os tipos de doenças cardíacas. Trata-se de um exame que é fruto de uma série complexa de processos fisiológicos e tecnológicos [1].

Para além de medir a frequência cardíaca, o eletrocardiograma permite também analisar o efeito de medicamentos ou regular o funcionamento de dispositivos que auxiliam ao funcionamento do coração, de que é exemplo o *pacemaker*. A palavra "eletrocardiograma" deriva do grego *elektro* (*eletricidade*) + *kardio* (*coração*) + *graph* (*escrever*).

2.1.1 Enquadramento fisiológico

A explicação fisiológica deste exame é feita recorrendo ao conceito de fluxo de iões que circulam entre vários locais do coração, em correntes. Inicialmente, estas correntes de origem iónica formam-se ao longo da membrana celular e das suas células adjacentes. Estas são sincronizadas com a ativação cardíaca e geram um campo elétrico dentro e em torno do coração que variam ao longo de um ciclo cardíaco. Este campo elétrico atravessa outras estruturas do nosso corpo, sendo possível detetar ecos da atividade do coração nestes locais.

As correntes iónicas transmembranares são as principais responsáveis pelos sinais captados e registados num ECG. Considera-se, de um modo arbitrário, que estas correntes se representam por movimentos de cargas positivas. Numa situação de repouso, cada célula do músculo cardíaco apresenta uma carga negativa. Esta é anulada através de um fluxo de iões com carga positiva, como o Sódio (Na^+) ou o Cálcio (Ca^{2+}) – a este fenómeno chama-se despolarização (ver figura 2.1). Este evento é o principal responsável pelo desencadear das contrações cardíacas. No caso convencional, esta despolarização é promovida, a cada ciclo cardíaco, pelas células do nódulo sino-auricular. Esta estrutura é responsável por enviar os impulsos elétricos que percorrem as aurículas e estimulam o nódulo aurículo-ventricular (que posteriormente transmitirá impulsos aos ventrículos), pelo que tem um papel fundamental no controlo da frequência dos batimentos cardíacos [1].

2.1.2 Aplicabilidade

A electrocardiografia é um exame clínico muito utilizado em análises de rotina, quando um paciente é admitido nas urgências ou antes de uma intervenção cirúrgica, de modo a detetar anomalias até então desconhecidas ou para efeitos de monitorização. Em todo o caso, existe alguma flexibilidade nos regulamentos quanto à obrigatoriedade da sua realização nestas situações.

A análise do eletrocardiograma é unanimemente considerada o melhor método de deteção de qualquer arritmia cardíaca [1]. Este é particularmente indicado para encontrar anomalias no tecido condutor onde os sinais elétricos são transmitidos, ou anomalias nas contrações cardíacas causadas por problemas na despolarização. Nos casos de ocorrência de enfartes do miocárdio é possível, na maior parte das vezes, verificar que zonas do músculo cardíaco ficaram afetadas.

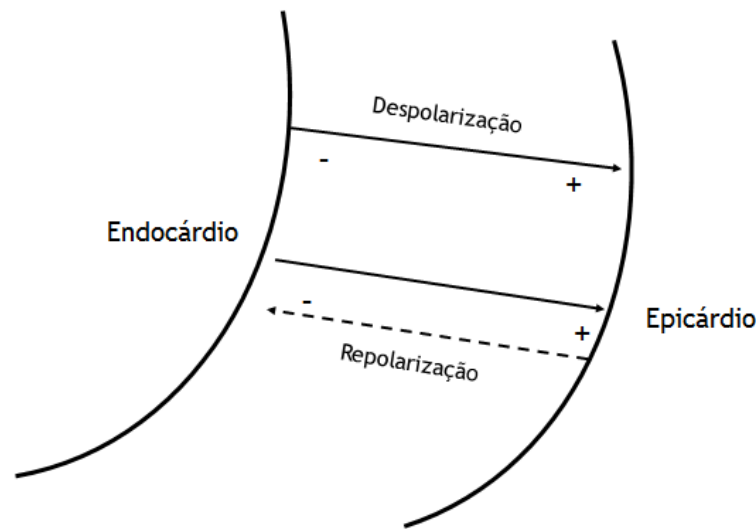


Figura 2.1: Mecanismo de despolarização e repolarização.

A principal limitação deste tipo de técnica é que não consegue medir a capacidade de bombeamento do coração, sendo para isso necessário recorrer a técnicas de ultrassom como o ecocardiograma ou outro tipo de meios na área da medicina nuclear. Estes meios complementares de diagnóstico são de enorme importância, uma vez que é possível um ser humano com um sinal de ECG perfeitamente normal entrar em paragem cardíaca, num fenómeno chamado *pulseless electrical activity*.

2.1.3 Eventos

A análise típica de um sinal de ECG envolve um conjunto de eventos fundamentais: onda P, complexo QRS, onda T e onda U (esta última visível em cerca de metade dos exames). A linha horizontal do ECG é comumente designada de linha de base ou isoeletrica. Para além destes elementos, outros são obtidos por análise indireta: intervalos PR, ST ou QT. Alguns destes exemplos encontram-se ilustrados na figura 2.2.

- **Linha de base ou isoeletrica:** ponto de partida para a atividade elétrica inerente às despolarizações e repolarizações;
- **Onda P:** durante a despolarização auricular normal, o vetor campo elétrico está orientado do nó sino-auricular para o nó aurículo-ventricular, distribuindo-se da aurícula direita para a esquerda, originando a onda P. Apresenta duração típica de 80 ms;
- **Complexo QRS:** reflete as variações rápidas na despolarização ventricular. Como os ventrículos apresentam maior massa muscular comparativamente com as aurículas, a amplitude da onda será muito maior que na onda P. Apresenta uma duração típica em torno dos 80-120 ms;

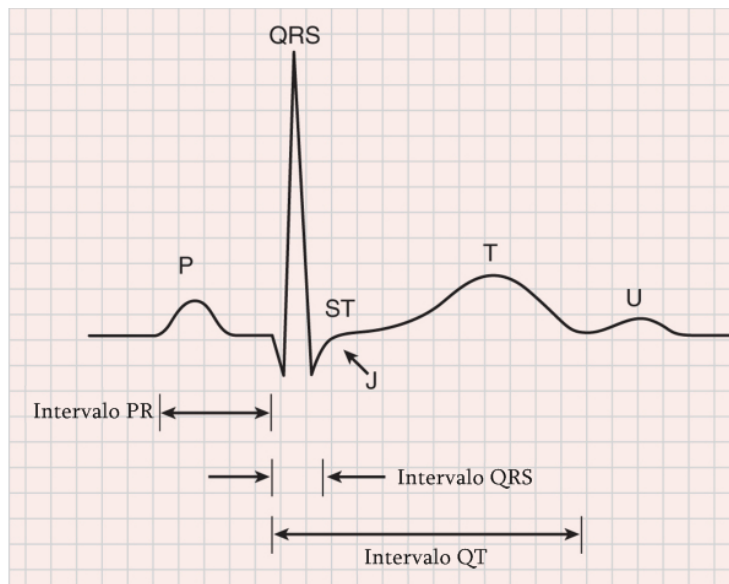


Figura 2.2: Alguns Eventos de um ECG [1].

- **Onda T:** representa a repolarização ventricular. Divide-se pelo pico em duas regiões distintas: período refratário absoluto e relativo (ou vulnerável) (duração típica de 160 ms);
- **Onda U:** repolarização do septo interventricular. Normalmente apresenta amplitude baixa ou quase nula. Caso tal não ocorra, poder-se-á estar perante patologias como hipertiroidismo ou hipercalcemia [2];
- **Intervalo PR:** demarcado pelo início da onda P e início do complexo QRS. Reflete o tempo que o impulso elétrico demora a atravessar desde o nó sinusal até ao nó aurículo-ventricular, antes de entrar nos ventrículos. Tem a duração típica de 120-200 ms;
- **Segmento PR:** semelhante ao anterior, mas começando no final da onda P. Representa o intervalo entre o fim da despolarização auricular e o início da polarização ventricular. Tem a duração típica de 50 a 120 ms;
- **Segmento ST:** intervalo entre o final da onda S e início da onda T, representa o intervalo de excitação ventricular. Tem a duração de 80 a 120 ms;
- **Intervalo QT:** vai desde o início do complexo QRS até ao final da onda T, representando o tempo de sístole elétrica. A sua medição pode ser importante para prevenir taquiarritmias ou morte súbita.

São detetáveis outros eventos na medição do ECG. No entanto, estes são considerados os mais importantes para a compreensão genérica de um gráfico para a análise clínica do electrocardiograma resultante da electrocardiografia.

2.1.4 Erros no eletrocardiograma

Tal como noutros exames médicos, o ECG apresenta uma margem de erro. A eficácia deste exame como ferramenta de diagnóstico está dependente de fatores como as indicações existentes para o procedimento, o contexto clínico no qual é feito, o uso correto dos dispositivos e a capacidade do examinador. O ruído no sinal devido às forças que atuam nos elétrodos, especialmente às altas frequências, pode potenciar ainda mais a ocorrência de erros.

A análise do ECG feita sem conhecimento do historial clínico do paciente pode ser limitativa. Torna-se relevante, por exemplo, saber que medicação está a ser consumida, uma vez que esta pode perturbar os resultados. A existência de registos de ECG anteriores para efeitos comparativos pode também ter utilidade.

Outro problema comum que pode levar a problemas de diagnóstico está relacionado com erros técnicos (por exemplo, troca de elétrodos). Situações destas podem conduzir a posteriores testes e que implicam o desperdício de recursos, bem como à realização de tratamentos que podem colocar em perigo o paciente. Conhecer os padrões de leitura deste tipo de erros ajuda no seu despiste.

Os erros na interpretação de sinais de ECG têm aumentado nos últimos anos. Estes culminam em falhas na deteção e triagem de pacientes com doenças que os podem colocar em situações perigosas. Estudos concluem que erros de grande proporção podem ser identificados em 4 a 32 % das interpretações [1]. Além deste tipo de erros, devem-se ter igualmente em conta os erros computacionais. Apesar de os computadores fornecerem uma ajuda significativa em termos de processamento e algoritmia, os instrumentos de medida ainda apresentam limitações que impedem a deteção de problemas mais complexos. Novas técnicas baseadas em inteligência artificial, aliadas a avanços no hardware, tentarão colmatar esta falha da melhor maneira possível.

2.1.5 Dispositivos utilizados

Para realizar uma eletrocardiografia, é necessário um conjunto de elétrodos para captar as diferenças de potencial. É também necessário um sistema de visualização (ver figura 2.3). A qualidade desta visualização está dependente do tipo de sistemas eletrónicos usado para amplificação, filtragem e digitalização da informação obtida. Depois deste tratamento de informação, o registo é normalmente impresso num papel quadriculado com dimensões próprias, num gráfico de tensão elétrica *versus* tempo. Antes da medição, é prática comum o uso de um sinal de calibração.

A portabilidade e conseqüente aceitação de novos dispositivos de medição de ECG, especialmente em pacientes de idade mais avançada, deve ser um fator a ter em conta no seu projeto, contribuindo para uma monitorização mais permanente [3].

2.1.6 Exame típico

O exame mais comum é o ECG de 12 derivações. Cada derivação corresponde a uma diferença de potencial entre dois elétrodos colocados em zonas diferentes do corpo. Antes de iniciar o procedimento, o examinador deverá explicar as etapas do processo. A sala terá de apresentar uma

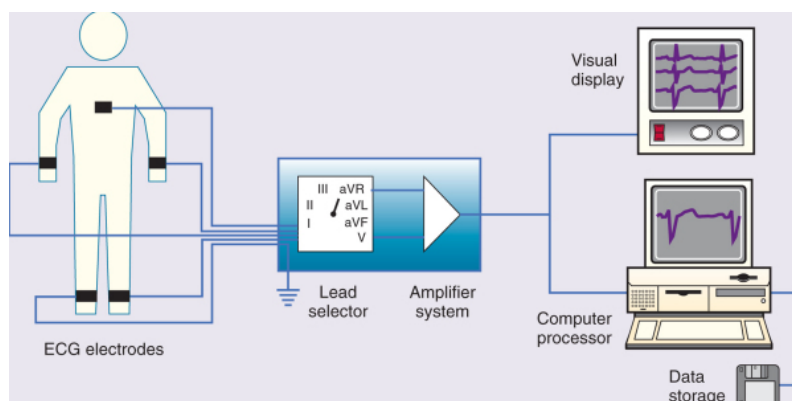


Figura 2.3: Principais dispositivos de um ECG [1].

temperatura agradável e o paciente deverá estar relaxado antes do procedimento. Deve-se também conhecer a medicação consumida no curto período de tempo anterior ao exame.

Depois de colocar o paciente em decúbito dorsal, deve determinar-se a posição das derivações. Normalmente, quatro irão para os membros e as seis restantes para posições do peito. De seguida, é colocado o gel de condução nos respetivos locais, ligando-se os elétrodos ao aparelho de medição. Pode ser necessária uma tricotomia, especialmente no caso masculino. Após captação e amplificação, os sinais são registados para visualização on-line ou análise especializada posterior [1].

O tipo de eletrocardiograma a ser feito no âmbito deste estudo terá menos derivações que aquele que, tipicamente, é realizado num contexto médico. O princípio de funcionamento é o mesmo, mas aplica-se apenas uma derivação, com um elétrodo colocado no pulso e outro no tornozelo, acrescidos de um elétrodo de ligação à terra. Este procedimento será explicado em maior detalhe no capítulo 4, que incide sobre o procedimento adotado para este estudo em concreto.

2.1.7 Exemplos de anomalias

Por análise dos eventos anteriormente descritos torna-se possível, para um profissional treinado, fazer um diagnóstico diretamente a partir da leitura do ECG. Por exemplo, um intervalo QT prolongado/encurtado pode indicar hipercalcemia ou uma anomalia genética. Uma onda T invertida ou achatada, isquemia coronária ou hipertrofia no ventrículo esquerdo. Uma onda T hiperaguda pode ser sugestiva de um futuro enfarte no miocárdio. Uma extrassístole ventricular é diagnosticada graças à presença de um complexo QRS prematuro e de morfologia larga ou deformada com duração superior a 120 ms, seguido de uma pausa compensadora.

Como estes, existem muitos outros exemplos de variações nos eventos que apontam para um diagnóstico, bem como mecanismos automáticos para classificação de vários tipos de arritmias [4]. No entanto, importa salientar que este exame não é uma ferramenta isolada na deteção de anomalias, funcionando apenas como um auxiliar na sua deteção numa primeira instância. Caso

se pretenda uma confirmação, terá de se proceder a um conjunto de exames complementares mais detalhados.

Tendo em conta que este estudo não visa aplicações diagnósticas nem clínicas, não se aprofundará mais esta abordagem.

2.2 Caracterização de parâmetros do som

O som é produzido por uma onda mecânica que representa uma oscilação de pressão transmitida por um meio físico (sólido, líquido ou gasoso), composto por uma ou mais frequências. Estas frequências podem ou não estimular os mecanismos de audição, isto é, serem ou não audíveis.

Um som, no seu estado natural, resulta de combinações complexas entre várias ondas sinusoidais. Cada um destes sinais é caracterizado através de um conjunto de parâmetros como a frequência ou amplitude. O ser humano é capaz de ouvir numa gama de frequências compreendida entre os 20 e os 20 kHz, sendo que os sons cujas frequências são menores ou maiores que esta gama se denominam por infrassom e ultrassom, respetivamente. Neste sub-capítulo apresenta-se um conjunto de parâmetros caracterizadores do som que auxiliam na sua análise acústica.

2.2.1 Intensidade sonora

É a perceção da amplitude da onda sonora. A perceção desta grandeza pelo ouvido humano apresenta uma característica logarítmica. Isto implica que variações apenas são percecionadas no caso das amplitudes variarem exponencialmente. A medida de pressão sonora em relação à perceção auditiva é feita recorrendo a uma unidade logarítmica, o deciBel (dB) [5].

2.2.2 Frequência fundamental (F0) e formantes (F1, F2... Fn)

A frequência fundamental corresponde ao primeiro harmónico de uma onda sonora, sendo a frequência mais influente na perceção de um dado som. No caso da voz, existe variação deste valor consoante a idade e o sexo. Num homem adulto está entre os 85 e os 180 Hz, e na mulher entre os 165 e os 255 Hz. É um dos principais elementos caracterizadores da voz.

Os formantes são concentrações de energia numa região do espectro sonoro. Desse modo, os harmónicos que se encontram nessa região serão realçados. Trata-se de uma informação distintiva importante na voz falada e cantada. O primeiro formante apresenta a menor frequência, e assim sucessivamente. Os dois primeiros formantes (F1 e F2) costumam ser suficientes para identificar uma vogal sustentada [5].

2.2.3 Pitch, timbre e intensidade subjetiva

Na acústica, o *pitch* (ou altura tonal) é reconhecido como a perceção da frequência de vibração do som, ajudando à distinção entre graves e agudos. Um tom simples é composto por apenas uma frequência, e um complexo por várias, sendo que a mais baixa indica o tom fundamental. As

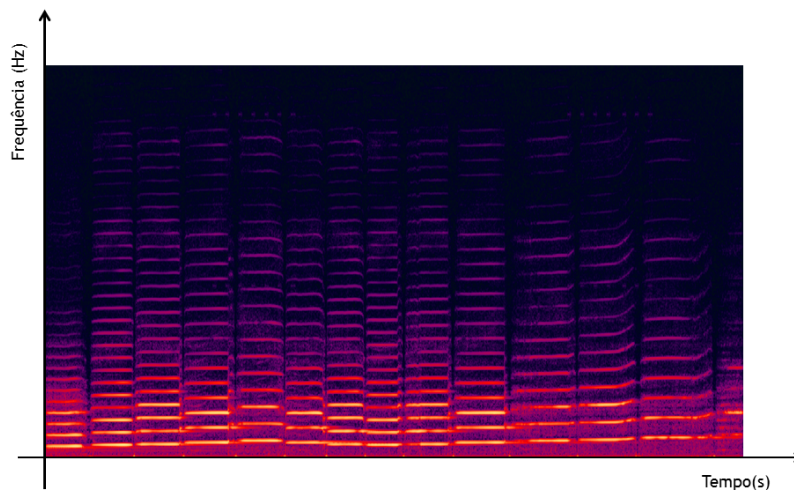


Figura 2.4: Espetrograma de um violino [7].

frequências dos tons seguintes são múltiplos da sua frequência fundamental, constituindo uma estrutura harmônica. Esta combinação agradável de tons compõe um tom musical [6].

O timbre é a característica de um tom que o distingue dos restantes num contexto de produção sonora. Dois instrumentos que produzam a mesma nota musical ao mesmo volume apresentam características diferentes no domínio das frequências, logo diferentes timbres, o mesmo se aplicando à voz.

A intensidade subjetiva (ou *loudness*) relaciona alguns dos parâmetros anteriormente descritos com a percepção qualitativa do ouvido humano. Pode variar de pessoa para pessoa, pelo que não é possível atribuir-lhe uma métrica fixa.

2.2.4 Representação e análise de Fourier

Para fazer a análise dos sons é necessário convertê-los primeiro para o domínio das frequências. Neste domínio é comum utilizar representações gráficas do módulo e da fase do sinal. Para uso computacional é frequentemente utilizada a FFT (*Fast Fourier Transform*), que é um algoritmo eficiente para processar a Transformada de Fourier Discreta (DFT – *Discrete Fourier Transform*).

Os sinais aos quais é aplicada a transformada são representados em espectros. Todos os sinais com variação temporal são representáveis desta forma. Esta representação torna os harmônicos mais visíveis, ajudando a uma análise mais eficaz do sinal. Decompor um sinal complexo em partes mais simples constitui uma prática comum neste tipo de análise, especialmente em sinais periódicos.

O espectrograma representa a variação temporal de um espectro de frequências, através de um gráfico com um eixo horizontal que representa o tempo, um vertical que representa a frequência e variações na cor que indicam a magnitude de uma dada frequência. Um exemplo desta representação pode ser visto na figura 2.4.

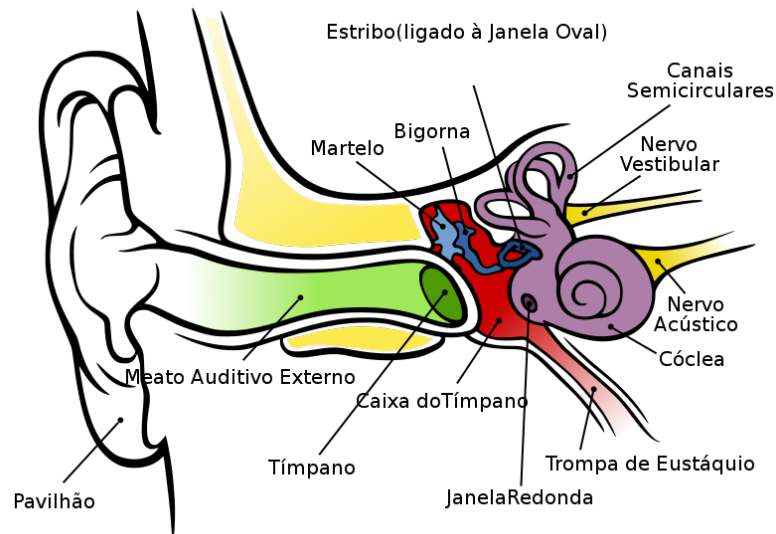


Figura 2.5: Anatomia do ouvido humano [5].

2.3 Produção e percepção da voz

Após caracterização dos parâmetros fundamentais do som, importa perceber quais os mecanismos desencadeados pelo ser humano para os produzir/ouvir.

2.3.1 Percepção auditiva

Compreender o modo como são captados os sons implica algum conhecimento da anatomia do ouvido. De um modo simplificado, este é composto pelo ouvido externo, onde as ondas sonoras são refletidas e atenuadas, fornecendo a informação necessária ao cérebro sobre a fonte do som. No canal auditivo, os sons entre 3 e 12 kHz são amplificados. Depois de atingida a membrana timpânica, o som atinge o ouvido médio. Aqui, atravessa um conjunto de ossículos (martelo, bigorna e estribo) que aumentam a pressão das vibrações para preparar a etapa seguinte.

O ouvido interno é fundamentalmente constituído pela cóclea, que contém os terminais nervosos responsáveis pela audição. Trata-se de um tubo ósseo espiral composto por três compartimentos membranares e um central que contém as células que ajudam a conferir a sensação de audição através dos movimentos dos fluídos no seu interior [5].

2.3.2 Produção de voz

Define-se produção de fala como o processo a partir do qual palavras faladas são escolhidas e articuladas recorrendo ao aparelho vocal humano. Normalmente, a fala é criada através da pressão promovida pelos pulmões, gerando som por fonação na glote. A característica da vibração das pregas vocais varia conforme a palavra a articular. Consideram-se três fases de processamento

na fonação: conceptualização (no qual se define a intenção de criar o fonema), criação do *lemma* (palavra sem som associado) e articulação (que envolve todas as mudanças físicas inerentes à fala) [8].

2.3.3 Modos e pontos de articulação

O modo como a fala é produzida depende do modo como o ar chega ao exterior. No caso de se tratar de um fluxo periódico de ar por vibração das cordas vocais, estamos perante fala vozeada. No caso do fluxo de ar ser forçado através de uma constrição, por exemplo, na glote, gera-se uma turbulência que dá lugar à produção de sons não vozeados. A existência dos modos vozeado e não vozeado permite a distinção entre estes fonemas, de que são exemplo o /ff/ ou o /v/.

No caso das vogais, o ponto de articulação caracteriza-se pela posição da língua, pela abertura do maxilar e pelo arredondamento da boca. No caso das consoantes, a única informação disponível é o local da constrição, que normalmente é supra-glótica [5].

2.3.4 Mecanismos de fonação

O tipo de som que produzimos pode ser classificado de acordo com a disposição das pregas vocais. Se estiverem estáticas, considera-se a situação de fonação não vozeada. Se a glote for comprimida, existe um ponto intermédio onde a vibração das pregas vocais é máxima – este ponto define a voz modal presente nas vogais [5].

Existe também a voz murmurada, quando existem porções sem vozeamento (por exemplo, na palavra “japonês”). Outros exemplos poderiam ser considerados, como a voz nasalada ou de falsete.

2.3.5 Modelo Fonte-Filtro

Para o estudo dos mecanismos anteriormente descritos é muito utilizado o modelo fonte-filtro, em que a fonte são os impulsos/ruídos glotais e o filtro as influências de ressonância do trato vocal (oral e nasal). Por ser uma aproximação simplificada e precisa, é utilizada nas mais variadas aplicações, nomeadamente em compressão de sinais de fala e áudio.

Como os impulsos glotais têm um comportamento periódico, o seu espectro será um conjunto de harmónicos da frequência fundamental cuja magnitude decresce à razão de 12 dB por oitava. A esta fonte é aplicado um filtro que, no caso mais simples, será composto unicamente por pólos cujos coeficientes são obtidos através de mecanismos de predição linear [5].

2.4 Medição de parâmetros de voz

Depois de compreender os mecanismos que geram e captam a voz, importa introduzir um conjunto de medidas que a avaliam e quantificam objetivamente. Estas medidas poderão ter em conta vários critérios.

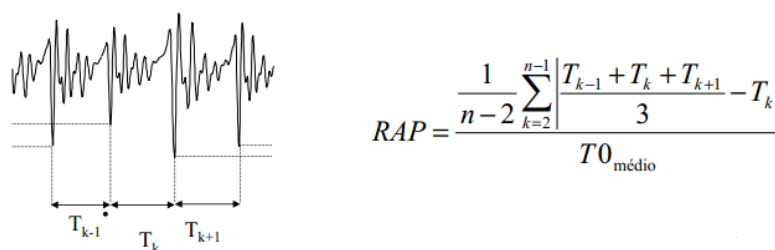


Figura 2.6: Representação de impulsos glotais (à esquerda) e fórmula de cálculo do RAP (à direita) [10].

2.4.1 Critérios de avaliação

Existem três critérios fundamentais para avaliar um sinal de voz: fisiológico, perceptivo e acústico. No caso de se tratar de uma avaliação fisiológica, exames clínicos como a laringoscopia ou a estroboscopia são os mais indicados. Numa avaliação perceptiva, existe um conjunto de protocolos que ajudam a classificar uma voz. Os mais utilizados neste tipo de avaliação são, por exemplo, o GRBAS (*grade, roughness, breathiness, asteny, strain*) ou o CAPE-V (*Consensus Auditory-perceptual Evaluation of Voice*). No entanto, a sua elevada subjetividade conduz a análises que podem ser inconsistentes.

Na avaliação acústica da voz, existe um conjunto de parâmetros objetivos que permitem uma comparação objetiva e quantificada. Nesta avaliação promove-se o uso de fonação sustentada, captando um mínimo de cem impulsos glotais para avaliação. Avalia-se também a frequência fundamental (número de impulsos glóticos por segundo) em termos de média, desvio padrão ou outros descritores. Este critério foi o utilizado ao longo deste estudo [5].

É frequente o uso de avaliação acústica e perceptiva simultaneamente, de modo a encontrar modelos que as correlacionem. Este tipo de abordagem é ainda alvo de intensa investigação.

2.4.2 Jitter

O *jitter* é a variabilidade do período fundamental de um sinal de voz, ciclo a ciclo. Do ponto de vista prático, manifesta-se através de vozes rugosas ou ásperas. Este parâmetro pode ter muitas origens fisiológicas como o movimento involuntário de músculos laríngeos, a pulsação sanguínea, assimetrias nas pregas vocais, vibrações na tiróide, entre outras [9].

Existem vários métodos para o seu cálculo, entre os quais o RAP (*Relative Average Perturbation*) ou o PPQ (*Period Perturbation Quotient*). Para uma gama normal em vogal sustentada (> 100 ciclos), o RAP não deverá ser superior a 1.5% [10].

2.4.3 Shimmer

O *shimmer* é a variabilidade involuntária da amplitude dos ciclos glotais, estando relacionado com a rouquidão da voz. Do ponto de vista prático, uma pessoa rouca apresenta valores elevados

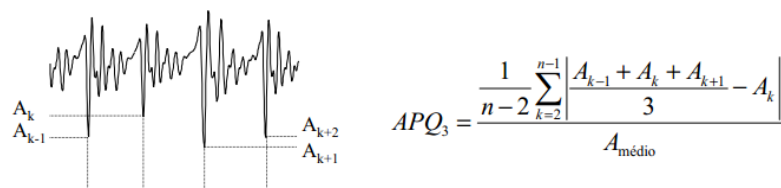


Figura 2.7: Representação de impulsos glotais (à esquerda) e fórmula de cálculo do APQ (à direita) [10].

neste parâmetro, bem como uma modulação em amplitude junto aos picos do sinal de voz. Fisiologicamente, o *shimmer* está associado à rugosidade da mucosa, à redução da resistência glótica ou à presença de lesões nessa região [9].

Uma das fórmulas de cálculo é o APQ (*Amplitude Perturbation Quotient*). Para as condições referidas anteriormente este valor, para os ciclos na sua vizinhança, não deverá ultrapassar os 3% [10].

2.4.4 Harmonic-to-noise ratio (HNR)

O ruído é o principal responsável pela característica não periódica dos sinais de voz, daí a importância da extração de parâmetros com ele relacionados. Uma voz patológica ou com muito ruído pode apresentar uma pior inteligibilidade.

Separando as componentes periódicas das não periódicas no sinal de voz, é possível relacionar as suas energias, obtendo o *Harmonic-to-noise ratio*. Esta grandeza exprime-se em dB para se aproximar melhor da apreciação perceptiva, por humanos, da qualidade da voz. Para as condições de voz normal referidas anteriormente, o HNR deve ser superior a 12 dB. Está maioritariamente relacionado com sopro (escape de ar) e astenia (ou fraqueza) presentes na voz [5]. O cálculo desta grandeza é feito recorrendo à equação 2.1.

$$HNR = 10 \log_{10} \frac{\text{Energia}_{\text{harmonica}}}{\text{Energia}_{\text{ru\u00eddo}}} = 10 \log_{10} \frac{\int |H(f)|^2}{\int |N(f)|^2} \quad (2.1)$$

2.4.5 Tempo máximo de fonação

Este parâmetro ajuda a medir a eficiência glótica. Mede o tempo máximo durante o qual se dá a emissão de uma vogal, com intensidade e *pitch* constantes. Esta medição é feita por duas ou três vezes, obtendo-se o melhor resultado. Caso este valor seja inferior a 14 segundos, pode-se estar perante um quadro patológico [5].

2.4.6 Estimação de impulsos glotais

Este mecanismo é de enorme importância tendo em conta a modelização destes impulsos para síntese artificial da voz. A estimação é feita recorrendo, por exemplo, a uma eletroglotografia.

Nesta técnica, mede-se a resistência elétrica do contacto das pregas vocais, colocando-lhes, simetricamente, dois eléttodos. As formas de onda obtidas indicam as suas fases de abertura e fecho.

Este método é imune a ruído, uma vez que não é influenciado pelo trato vocal – a medição é feita diretamente na fonte. A deteção de desvios na posição e no tempo de abertura/fecho nas pregas vocais pode relacionar-se com certas patologias.

Entre as limitações do teste está a dificuldade de captação devido a uma deficiente colocação dos eléttodos, especialmente devido a variações na estrutura da laringe dos examinados.

2.5 Influência do sistema nervoso na voz e ritmo cardíaco

Como foi explicado na secção 1.1, a natureza dos movimentos associados à fonação e as alterações no nosso ritmo cardíaco são eventos que estão diretamente associados a respostas provenientes do sistema nervoso humano. Ao longo deste sub-capítulo pretende-se explicar, de um modo simples, como é que este processo se desenrola nos dois casos.

2.5.1 Nervo laríngeo recorrente

O cérebro é responsável pela coordenação da voz através de sinais enviados por terminações nervosas específicas. Para a movimentação dos músculos do trato vocal, os responsáveis são o nervo laríngeo recorrente e o nervo laríngeo superior. O primeiro é assim chamado devido ao seu trajeto, proveniente do nervo vago até ao peito e voltando para o pescoço para terminar na laringe (ver figura 2.8).

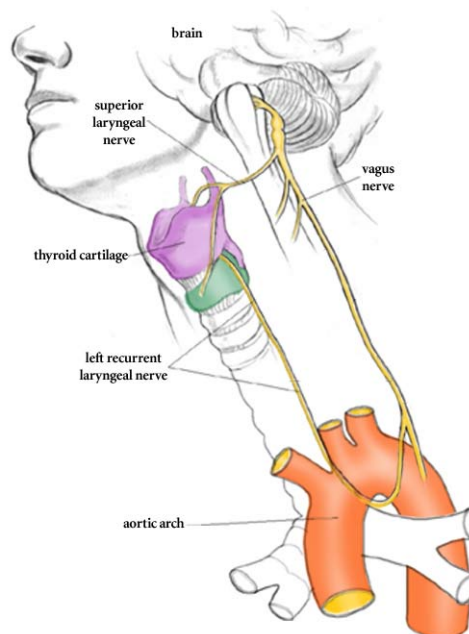


Figura 2.8: Localização dos nervos associados ao trato vocal [11].

Esta localização pode colocar o nervo em perigo em situações de cirurgia de peito aberto, causando danos ou eventual paralisia. Vozes patológicas são frequentemente associadas a problemas no fecho total das pregas vocais, por defeitos no envio da informação ao longo destas terminações nervosas [12]. Neste estudo, o propósito não passa pela análise deste tipo de vozes, mas sim de encontrar pequenas variações inerentes à interpretação feita a estímulos visuais.

2.5.2 Sistema nervoso simpático e parassimpático

Pode-se dizer que o sistema nervoso autónomo se encontra dividido em dois grandes ramos: os sistemas simpático e parassimpático. Ambos operam através de uma série de neurónios interligados que comunicam entre os sistemas nervoso central e periférico através de sinapses.

O sistema simpático é o que permite ao nosso corpo funcionar em condições de maior pressão, desencadeando um conjunto de reações fisiológicas, de que são exemplo o aumento do ritmo cardíaco, o aumento do peristaltismo do esófago ou a dilatação pupilar. Estas alterações processam-se de forma automática e involuntária [13].

O sistema parassimpático tem uma função oposta, sendo responsável por desencadear o conjunto de respostas fisiológicas perante situações de relaxamento, como o descanso ou o adormecer. Deste modo, estes mecanismos traduzem-se numa redução da frequência cardíaca, da adrenalina ou do açúcar no sangue. Uma distinção entre os efeitos destes dois sistemas pode ser vista na figura 2.9 [14].

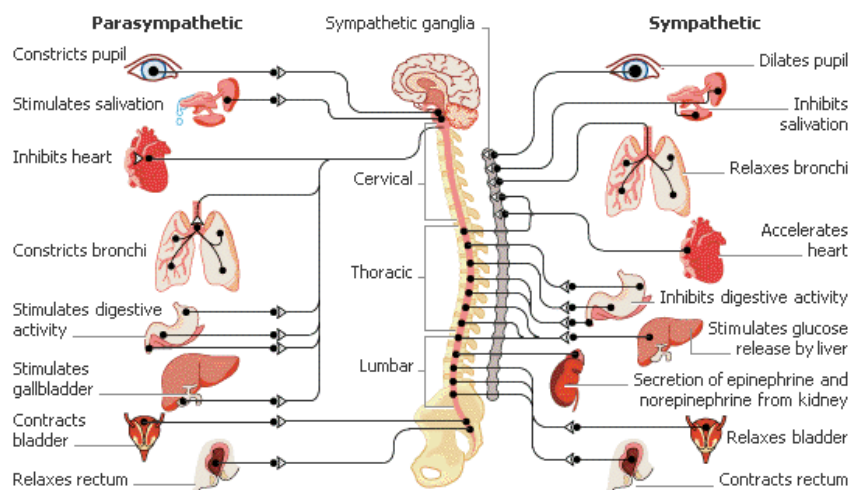


Figura 2.9: Sistema nervoso autónomo [15].

No caso deste estudo, as respostas que irão ser avaliadas consistem apenas em mudanças no ritmo cardíaco. Pretende-se aferir até que ponto a actuação destes dois sistemas (simpático e parassimpático) conduz a variações induzidas pela interpretação dos estímulos a apresentar aos participantes do estudo.

2.6 Avanços na avaliação de distorções na voz e no ECG

Neste sub-capítulo será feito um levantamento de alguns avanços feitos na relação entre parâmetros de distorção de voz e ECG.

Numa investigação feita por Skopin e Baglikov, é provado que a modulação dos impulsos cardíacos influencia a voz humana e que essa informação pode ser extraída no domínio das frequências [16]. Esta extração é feita recorrendo ao espectro proveniente da fonação de uma vogal, considerando o modelo fonte-filtro explicado anteriormente.

Tendo em conta que a laringe contém músculos cobertos por vasos sanguíneos, torna-se igualmente possível relacionar os batimentos cardíacos com as variações dinâmicas de parâmetros da voz como o volume ou tempo de fonação. O sistema de captação consiste num microfone e um sistema portátil de ECG [17].

Existem também avanços no que diz respeito à identificação de emoções a partir da voz e do ECG. Parâmetros de fala, de que são exemplo os formantes F2 e F3, identificam vários estados emocionais. No estudo realizado por Ying-Chien Wei et al. [18], descobriram-se os efeitos que a raiva e o medo têm sobre o ciclo de batimentos cardíacos. Para obter a relação pretendida, os sinais de voz e ECG foram captados simultaneamente.

Existem também avanços no que diz respeito à identificação de emoções na voz cantada. Utilizando parâmetros como afinação, amplitude e timbre, é possível extrair informação dos 5 primeiros formantes provenientes da captação de vogais distintas em peças musicais de conteúdo emocional específico. Com esta informação é possível associar cada um dos formantes numa emoção específica: alegria (F3), tristeza (F4) e raiva (F5) [19].

Num outro contexto, foi levado a cabo um estudo que indicava que a fonação em pacientes que sofrem de apneia do sono é diferente daqueles que não apresentam esta patologia. Para chegar a esta conclusão foram recolhidas amostras de som de um conjunto de voluntários com um perfil específico [20]. Em oposição à análise convencional utilizada neste caso, é também possível seguir uma análise dinâmica não linear dos sinais de voz, associando-os a vozes saudáveis ou patológicas [21].

2.7 Considerações finais

Todos os conceitos teóricos aqui apresentados, quer para os sinais eletrocardiográficos quer para os sinais de voz, se revelarão úteis para a compreensão do estudo descrito nesta dissertação. A pesquisa feita ajudou a uma maior ambientação com um conjunto de conceitos completamente novos, especialmente no que diz respeito ao domínio da eletrocardiografia.

De um modo geral, o estudo de alterações na voz e no ECG e a sua comparação em registos recolhidos na mesma janela temporal é uma temática que, tanto quanto a pesquisa efetuada conseguiu apurar, está relativamente inexplorado. Não quer isso dizer que a informação recolhida seja inválida, uma vez que confronta um conjunto de perspetivas diferentes que conferirão um grupo de conhecimentos e abordagens que poderão revelar-se úteis.

Capítulo 3

Ferramentas utilizadas

Depois de recolhida a principal informação de base para o desenrolar do projeto, importa referir quais as ferramentas que foram utilizadas para levar a cabo a prova do conceito. Nesta identificação, as ferramentas foram divididas em dois grandes grupos, bem demarcados temporalmente: os recursos utilizados aquando da recolha das amostras e os utilizados na fase de tratamento e análise de sinais.

3.1 Recolha de amostras

Na fase de recolha de amostras, foi necessária a obtenção de um conjunto de recursos muito específico que, por vezes, exigiu alguma paciência e persistência para levar a cabo toda a logística inerente.

3.1.1 Biopac MP35

Este foi o dispositivo utilizado para a captação dos eletrocardiogramas. É também capaz de efetuar registos de outros sinais fisiológicos, como eletromiografias ou eletroencefalografias. Possui quatro canais para aquisição de dados e é acompanhado de um software para controlar as aquisições e efetuar os registos em tempo real. Tem diversas opções de visualização e análise, bem como a capacidade de exportação de dados para outros formatos e ferramentas de análise de sinal. A taxa de amostragem utilizada para recolha de sinais de ECG foi de mil amostras por segundo.

Como os sinais captados são de amplitude muito baixa, o dispositivo dispõe de capacidade de amplificação. Além disso, filtra ruído proveniente do sistema elétrico e de outros sinais de interferência, fazendo a conversão necessária para que o software de captação consiga exibir as formas de onda [22]. As especificações detalhadas deste dispositivo podem ser consultadas no anexo A.1. A figura 3.1 ilustra este dispositivo.



Figura 3.1: Aparelho Biopac MP35 [22].

3.1.2 Biopac Student Lab (BSL)

Foi o software utilizado para aquisição do sinal de eletrocardiografia. Foi fornecido juntamente com o aparelho Biopac, uma vez que é necessária uma licença válida para utilização deste num contexto de captação.

Com uma interface simples mas bastante funcional, este software já incluía rotinas parametrizadas para o tipo de captação que iria ser feita [23]. Esta característica, aliada ao uso de um manual com indicações importantes para colocação de elétrodos e outras informações úteis, facilitou bastante o processo de aprendizagem.

3.1.3 Adobe Audition CS5.5

Foi o software utilizado para aquisição do sinal de voz. Apesar de todas as funcionalidades de edição e processamento em tempo real, a única da qual se tirou partido foi a gravação em formato WAV, a uma frequência de amostragem de 48000 kHz e com 16 bits. Durante este processo, foi feita uma monitorização em tempo real da captação. Depois, foi necessária a segmentação manual de várias porções de registo de voz, para as quais o software se revelou muito prático, diminuindo o tempo da tarefa [24].

3.1.4 Microfone Sennheiser EARSET 1

Para captação dos sinais de voz, o dispositivo utilizado foi este microfone de *headset* de elevada precisão, especialmente projetado para captação de sinais de voz. Para além da inteligibilidade dos sinais dele resultantes, não capta ruído proveniente de movimentações de ar e apresenta dimensões muito pequenas, adaptando-se anatomicamente à orelha e à cara. Com um design bastante flexível, torna-se possível a utilização de diversos componentes neste mesmo dispositivo [25]. A fisionomia deste dispositivo pode ser vista na figura 3.2 e as suas especificações podem ser consultadas no anexo A.2.



Figura 3.2: Microfone Sennheiser Earset 1 [25].

3.1.5 Roland UA 25EX

Este foi o dispositivo utilizado para fazer a pré-amplificação entre o microfone e o computador, na obtenção do sinal de voz. O seu desenho foi feito especialmente de forma a garantir elevada mobilidade. Tem capacidade para conversões A/D de 24 bits / 96 kHz, fazendo uma compressão/limitação apropriada para uma utilização num contexto de estúdio/profissional. A pré-amplificação é feita com um sistema de *Phantom Power* de +48V [26]. As especificações detalhadas deste dispositivo podem ser consultadas no anexo A.3.

A grande variedade de entradas e saídas é também uma característica que reforça a sua utilidade. No caso desta captação em específico, foi necessária a utilização de um adaptador XLR para Jack 3.5mm para ligar o microfone ao sistema. A escolha recaiu sobre o MZA 9009 da *Sennheiser*.

3.1.6 *International Affective Picture System (IAPS)*

Esta base de dados de imagens foi utilizada para a escolha dos estímulos a serem exibidos aquando da recolha das amostras. Baseado num enorme conjunto de exemplos (cerca de mil) com classificação proveniente de um elevado número de pessoas (de várias idades e ambos os sexos), o resultado final é um conjunto de classificações segundo três critérios: domínio, excitação e prazer. Estes três critérios culminam numa classificação única que reflete o domínio, intensidade e valência das emoções induzidas por cada imagem.

O modo como estes estímulos devem ser aplicados para potenciar as reações emocionais desejadas está explícito no manual técnico deste sistema. Torna-se importante que os sujeitos não tenham sido expostos previamente às imagens, para que não se crie qualquer tipo de habituação [27].

3.1.7 Material adicional

Para além dos recursos anteriormente mencionados, foi necessário reunir material adicional para que a recolha pudesse ser feita com sucesso. Para sinais de eletrocardiograma e voz foi utilizado um computador portátil pessoal. Para a apresentação dos estímulos foi necessário um

monitor em frente ao sujeito, cedido pela Faculdade de Psicologia. Os elétrodos utilizados no sinal de eletrocardiograma são descartáveis e foram fornecidos juntamente com o dispositivo.

Durante a maior parte das recolhas foi utilizada uma câmara de filmar para fazer despiste de eventuais artefactos no sinal, designadamente os motores (movimentos voluntários ou involuntários que interferissem no sinal).

3.2 Processamento da Informação

Na fase de processamento, a reunião de recursos foi mais simplificada, uma vez que grande parte destes eram software. Numa primeira fase, foram utilizados em paralelo o Matlab e o Praat (ou Voice Studio) para processamento de sinais de eletrocardiografia e voz, respetivamente. Depois de obter todos os parâmetros destes sinais, foi utilizado o SPSS (*Statistic Package for the Social Sciences*) para uma análise estatística.

3.2.1 Algoritmia de Processamento de Sinal

Para efetuar as medições e comparações mencionadas ao longo deste documento, foi necessário recorrer a um conjunto de algoritmos que lidam com análise de sinal. Neste caso, os sinais analisados são fisiológicos e, como tal, contínuos no tempo.

Foi prática corrente converter os sinais para o domínio das frequências e fazer uma análise espectral, pelos motivos anteriormente referidos. O software utilizado para o efeito foi o *Matlab*, que está especialmente vocacionado para cálculo numérico e possui várias ferramentas para análise de sinais em geral. Além disso, existe um conjunto de ferramentas especialmente direcionadas à análise de sinais de voz – a *Speech Processing Toolbox* [28]. No entanto, como a análise de voz foi feita maioritariamente no *Praat* (mais vocacionado para análise de sinais de fala), não se tirou tanto partido destes recursos. Portanto, apenas nos sinais de ECG se recorreu a esta algoritmia.

Para sinais de eletrocardiografia, existe também um conhecido *toolkit*, o *Physionet*. Nele encontra-se presente um número considerável de funções que lidam com sinais de eletrocardiografia e que extraem os parâmetros tipicamente medidos na sua análise [29].

3.2.2 Praat

Palavra holandesa para “falar”, o *Praat* é um software livre para análise de sinais de fala. Foi desenvolvido por Boersma e Weenick, da Universidade de Amsterdão. A sua versão mais recente é a 5.1, datada do final de janeiro de 2009. A possibilidade de criação de *scripts* para análise e extração de parâmetros da voz (*jitter*, *shimmer*, entre outros) faz com que o seu uso se revele mais prático relativamente a outras opções [30].

3.2.3 Voice Studio

Também direcionado para processamento de sinais de voz, oferece um conjunto de ferramentas intuitivas e expeditas numa interface amigável. Neste software é possível a gravação através de um

microfone e a escrita instantâneas em formatos comprimido e não comprimido. É também possível a obtenção dos principais parâmetros objetivos de qualidade da voz, nomeadamente frequência fundamental, *jitter*, *shimmer*, HNR, entre outros – todos eles descritos anteriormente. Para além disto, destaca-se o facto de estas funções de análise conseguirem correr em tempo real [31].

Este software não foi usado para a extração de parâmetros, mas serviu como um auxiliar para confrontação de parte dos resultados obtidos.

3.2.4 Statistic Package for the Social Sciences (SPSS)

O SPSS possibilita um estudo estatístico, transformando e organizando toda a recolha feita de um modo que a torne relevante para o estudo. Este ambiente possibilitará também a realização de testes estatísticos (de que são exemplo os testes de correlação, de t-Student ou de qui-quadrado) ou cálculo de médias.

Apresenta também outras funcionalidades relacionadas com previsão de valores numéricos, de que é exemplo a regressão linear. Este tipo de técnicas são um auxílio importante em contextos de apoio à decisão.

A versão mais recente do software é a 20.0, datada de Agosto de 2011. Um exemplo da interface deste software pode ser visto na figura 3.3.

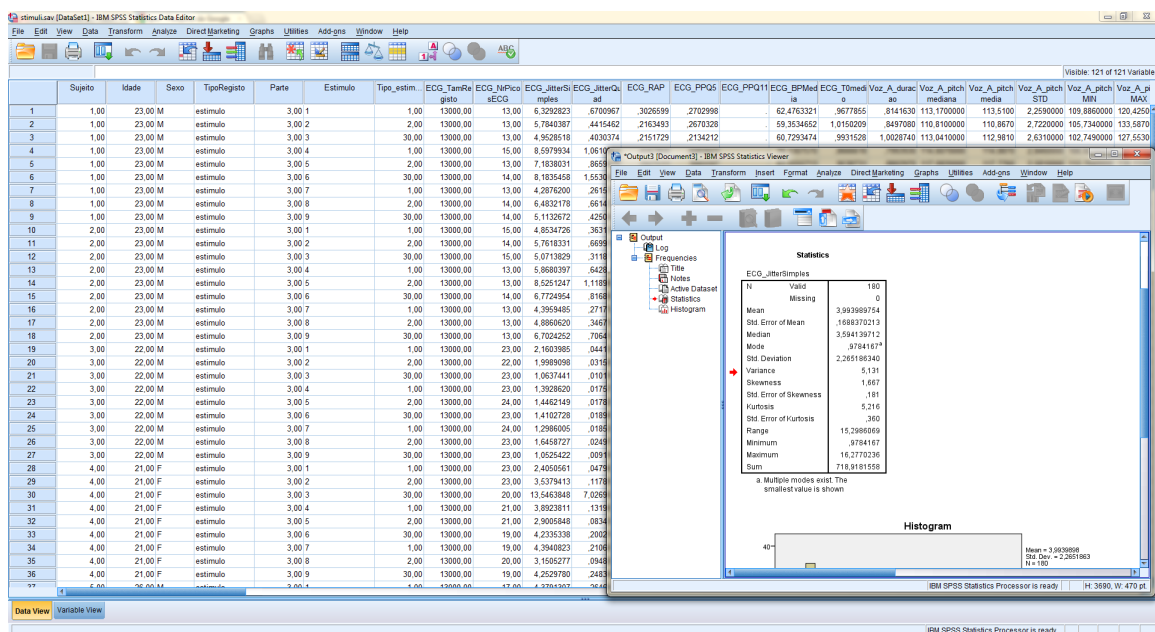


Figura 3.3: Exemplo de interface no software SPSS.

3.3 Considerações finais

Neste capítulo apresentou-se o conjunto de ferramentas utilizadas ao longo deste estudo. Importa destacar aquelas que tiveram um papel mais relevante na fase de processamento de sinal,

de que são exemplo o Matlab ou o Praat. No entanto, a utilização de algumas ferramentas complementares e até então desconhecidas (de que é exemplo o sistema IAPS) trouxe o complemento ideal à qualidade deste conjunto de recursos. De destacar também a fiabilidade do material de captação de som, conferindo maior segurança à análise.

Capítulo 4

Procedimento Experimental

Identificadas as principais ferramentas utilizadas, importa saber de que modo foram utilizadas para alcançar os resultados que aqui se apresentam. Neste capítulo, o enfoque será na fase de recolha. Será descrito o local onde esta ocorreu, o modo como se recrutaram os participantes para o estudo e o seu consentimento, o critério para escolha de estímulos e o protocolo experimental. Este trabalho foi realizado com a colaboração de um colega da Faculdade de Engenharia, criando um procedimento comum que satisfizesse os requisitos de ambos os estudos.

O projeto deste procedimento foi feito de um modo cuidadoso, uma vez que a complexidade inerente ao elevado número de estímulos e marcações a efetuar exigia o máximo rigor. Um simples erro conduziria irremediavelmente à realização de uma nova aquisição. Esta repetição não poderia ser feita no mesmo sujeito, uma vez que se assim fosse já estaria criada habituação aos estímulos visuais, comprometendo a validade das respostas fisiológicas obtidas.

Por último, tirar-se-ão as devidas conclusões relativamente a esta fase do projeto.

4.1 Local de Recolha

A recolha das amostras no âmbito deste estudo foi feita no Laboratório de Neuropsicofisiologia, da Faculdade de Psicologia e Ciências da Educação da Universidade do Porto. Foi necessário o transporte de algum material da Faculdade de Engenharia, sendo necessário um termo de responsabilidade de modo a prevenir qualquer problema associado ao transporte ou eventual danificação do material. O espaço utilizado na recolha de sinais está ilustrado na figura 4.1.

4.2 Recolha de Voluntários e caracterização da amostra

Para a recolha de amostras, foi definido um limiar mínimo de participações e repetições para o qual o poder estatístico dos resultados seria previsivelmente satisfatório. Esta explicação será dada na secção 6.2. Assim sendo, definiu-se que as recolhas seriam feitas em vinte participantes, metade do sexo feminino e metade do sexo masculino.



Figura 4.1: Local de recolha de amostras.

As idades dos participantes estão compreendidas entre os 21 e os 52 anos, com uma média de 24.9 anos, idade típica de um estudante universitário em final de curso. O método de recrutamento consistiu num e-mail dinâmico enviado a toda a comunidade estudantil da Faculdade de Engenharia da Universidade do Porto. Aqueles que demonstraram interesse preencheram um formulário com nome completo, idade, contacto e disponibilidade para colaborar no estudo. Um gráfico circular representativo da distribuição de idades dos voluntários pode ser consultado na figura 4.2.

As recolhas foram feitas entre os dias 23 de Abril e 4 de Maio de 2012, agendadas conforme a disponibilidade dada pelos voluntários.

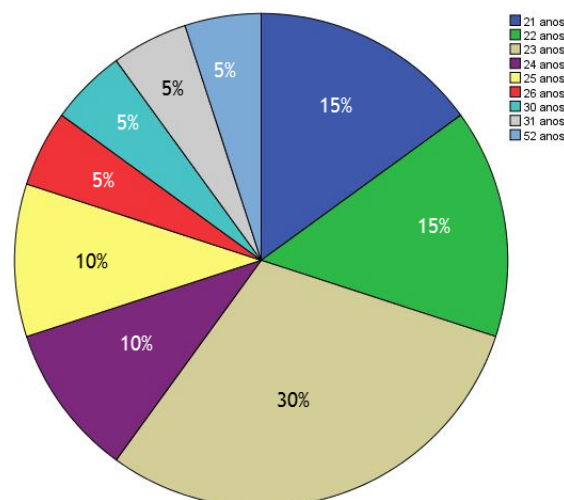


Figura 4.2: Distribuição das idades dos participantes.

4.3 Consentimento Informado

Antes do início de cada recolha, foi necessário apresentar um documento a cada participante para que existisse total conhecimento *à priori* da natureza do teste. Esta situação, para além de prevenir situações de desconforto inesperadas, dava a possibilidade a cada participante de colaborar ou não no estudo de forma informada, ou de o abandonar quando quisesse. Além de explicar as montagens dos elétrodos, indicava que o procedimento seria filmado para efeitos de despiste de interferências, sendo sempre garantido o total anonimato.

Foram também despistadas as contra indicações para a realização do ECG: embolia pulmonar, enfermidade aguda, limitação física ou psicológica e intoxicação medicamentosa (alguns deles já identificados na secção 2.1). Em todo o caso, a situação foi novamente verificada, perguntando-se diretamente a cada voluntário se sofria de alguma condição clínica que inviabilizasse o estudo.

Foi também dada a garantia que todo o equipamento e procedimentos usados obedeciam aos padrões de higiene e segurança e que o procedimento não causaria qualquer tipo de dor. Este consentimento informado foi criado com o auxílio do Laboratório de Neuropsicofisiologia. Foi também pedido um parecer a um médico cardiologista relativamente ao seu conteúdo. O consentimento pode ser visto no anexo E.

4.4 Escolha de imagens do IAPS

Como foi referido na secção 3.1, as imagens usadas como estímulos são integrantes do *International Affective Picture System*. Recorrendo aos valores normativos que o estudo obteve para o grau de agradabilidade dos estímulos para a valência (classificação intervalar entre 1 para um estímulo mais desagradável e 9 para um estímulo mais agradável), foram selecionados 27 estímulos (este número será explicado na secção 4.6). O critério utilizado foi o de potenciar ao máximo reações fisiológicas nos sujeitos a analisar, pelo que se optou pelas classificações mais elevadas para estímulos agradáveis, nas mais baixas para desagradáveis e nas intermédias para neutros.

O tempo que seria utilizado no caso de uma escolha mais cuidada seria bastante, especialmente devido ao tamanho da base de dados. Este fator, aliado à tentativa de seguir escolhas o mais objetivas possível, levou a que se optasse por esta solução.

4.5 Exame de eletrocardiograma

O exame de aquisição eletrocardiográfica que foi feito neste estudo seguiu um procedimento diferente do descrito na secção 2.1. Baseia-se numa derivação com dois elétrodos ativos e uma ligação à terra, estando o de polaridade negativa no pulso direito e o positivo no tornozelo esquerdo, acrescidos de um elétrodo "terra" no tornozelo direito (ver figura 4.3). Pode-se justificar esta abordagem com uma montagem apenas com dois elétrodos ativos, uma vez que não se visava uma análise clínica do sinal, mas sim que o algoritmo a ser implementado detetasse os complexos

QRS, de fácil registo. Entendeu-se que, para evitar situações desconfortáveis junto dos participantes (nomeadamente nos do sexo feminino), seria suficiente realizar este tipo de montagem. O procedimento presente no manual do Biopac MP35 facilitou a tarefa, apresentando toda a sequência do procedimento. O tipo de ligação a ser feita no dispositivo pode ser visto na figura 4.3.

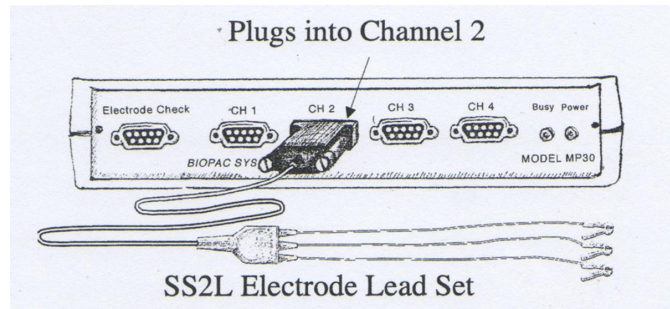


Figura 4.3: Ligação dos eléctrodos ao sistema Biopac MP35 [23].

A figura 4.4 representa as ligações dos eléctrodos utilizadas nas várias partes do corpo.

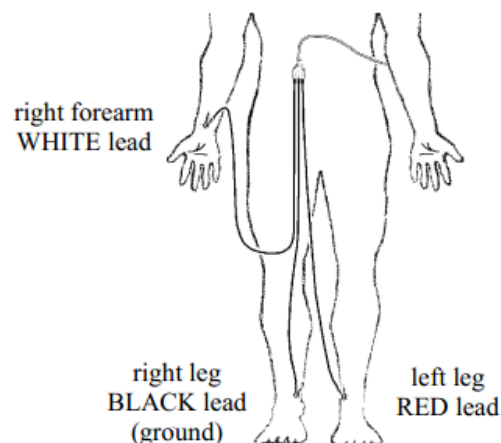


Figura 4.4: Ligação dos eléctrodos para cada sujeito [23].

4.6 Protocolo experimental

Nesta secção será explicada toda a sequência do protocolo de aquisição de dados. Tendo em conta que este é comum a dois estudos, será explicado na sua totalidade, uma vez que muitas das opções tomadas, especialmente no que diz respeito a duração de estímulos, estão limitadas pelo enquadramento geral.

Antes do início das aquisições, foi solicitado a cada participante que permanecesse sentado e com respiração normal, movimentando-se o mínimo possível (especialmente nos membros, onde se encontravam os eléctrodos). Outro fator importante a verificar foi a utilização de peças metálicas (anéis, pulseiras, etc) perto dos locais de medição. Além disso, tentou-se não expor os participantes a estímulos distratores, incluindo a observação dos próprios registos.

4.6.1 Fases do protocolo

O protocolo foi constituído por quatro partes: uma de calibragem e três blocos de estimulação. Na primeira parte foi feita uma calibração, em que se captaram sinais de eletrocardiografia, eletrodermografia e voz. Primeiro em situação de repouso (sem estímulos) sem vozeamento, depois com vozeamento da palavra “imagem” e das vogais /a/ e /i/. Foi testado se os sinais estavam de acordo com o esperado, antes de se prosseguir com o teste.

Na segunda parte, foi apresentado o primeiro bloco de estimulação, constituído por nove estímulos apresentados sequencialmente na ordem agradável-neutro-desagradável (três vezes). Nesta parte, não era pedido qualquer vozeamento, sendo apenas necessária a observação dos estímulos para captação de sinais fisiológicos.

Na terceira parte, apresentou-se novo bloco de estímulos, solicitando-se uma resposta: o sujeito teria de dizer a palavra “imagem” (devido ao número de vogais presentes nesta), seguida de uma classificação dada a esse estímulo. Neste processo, houve uma adaptação da escala relativamente à original (que estava definida de 1 a 9), sendo o valor -5 associado a um estímulo desagradável, 0 a um neutro e 5 a um agradável.

Na última parte, apresentou-se um último bloco de estímulos. No final de cada estímulo, solicitava-se o vozeamento das vogais /a/ e /i/. A figura 4.5 ilustra a sequência de todas as partes, bem como a sua duração. Cada "M" representado corresponde a um marcador, pelo que este diagrama foi a principal referência para que as marcações na recolha fossem todas feitas com sucesso.

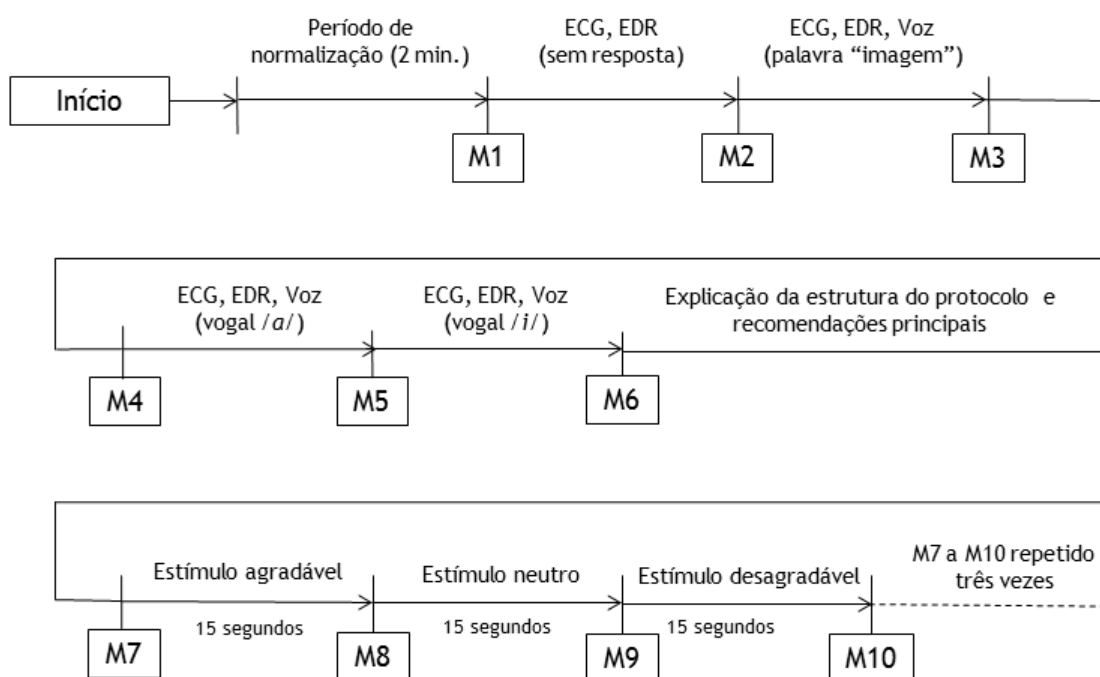


Figura 4.5: Sequência do protocolo, com indicação de marcadores.

De todas as partes referidas, apenas a terceira não foi utilizada no âmbito deste estudo. No total, o teste teve uma duração útil não superior a quinze minutos.

4.6.2 Tempo de exposição dos estímulos e intervalo interestímulos

A duração definida para cada ensaio foi de quinze segundos. Este valor justifica-se pelo fato da resposta mais lenta entre os sinais a serem captados (eletrocardiografia e eletrodermografia) rondar, no pior dos casos, cerca de 6/7 segundos [32]. A este valor acrescenta-se uma margem de segurança para cobrir erros nos processos de marcação da ocorrência de estímulos no registo e de segmentação dos registos.

Para cada ensaio foi utilizado um intervalo interestímulo com ponto de fixação. Portanto, cada ensaio estruturava-se na apresentação do estímulo durante um período de dois segundos, seguido de um ecrã preto durante os seguintes treze. É durante este período que o participante fornece a resposta pedida, dependendo da fase do teste em que se encontre (palavra /imagem/ seguida da classificação ou vogais /a/ e /i/).

4.6.3 Marcação e segmentação de estímulos

Entre cada estímulo, eram acionados marcadores nos registos dos sinais de eletrodermografia e eletrocardiografia. Uma vez que os dispositivos utilizados para a captação eram diferentes, esta foi feita de modo manual, sendo despistados eventuais erros de marcação após a recolha estar finalizada. O instante temporal exato onde os marcadores foram aplicados correspondia ao momento em que cada imagem era apresentada.

Esta segmentação, criada para todos os estímulos, partes e sinais, foi feita manualmente para cada participante. Este processo, apesar de demorado, (27 estímulos x 3 sinais x 20 participantes = 1620 segmentos) justifica-se para que não existisse qualquer tipo de erro que comprometesse os resultados finais deste estudo. Ao verificar cada segmento de sinal, tal cenário era garantido com elevado grau de confiança.

4.7 Considerações finais

Ao longo deste capítulo foi descrito todo o processo de recolha de amostras, desde a obtenção do espaço e de todo o material, até à segmentação dos sinais. Na fase inicial, nem sempre foi fácil conseguir reunir todas as condições para que as aquisições se iniciassem. Para que tal acontecesse, a contribuição de todas as pessoas envolvidas foi determinante, sem a qual não seria possível dar sequência a este conjunto de eventos. No entanto, uma vez terminada esta fase que envolveu maior logística com fatores externos incontrolláveis, o processo desenrolou-se de uma forma controlada e dinâmica.

Capítulo 5

Processamento de Sinais

Depois de feita a recolha, este capítulo irá descrever as abordagens utilizadas para o tratamento e análise dos sinais recolhidos. O software utilizado para o tratamento desta informação encontra-se descrito no Capítulo 3. Assim sendo, este irá conter duas partes fundamentais: uma que abrange o tratamento dos sinais de eletrocardiografia e outro que abrange os sinais de voz. No caso dos primeiros, será confrontado um conjunto de algoritmos desenvolvidos, de natureza automática e semi-automática. Esta comparação será feita, principalmente, em termos da sua eficácia, uma vez que foi este o fator diferenciador para a escolha da abordagem final.

No que diz respeito aos sinais de voz, será descrito como foram analisados e que parâmetros foram extraídos. Para cada um dos tipos de sinal, será apresentado um exemplo do conjunto de descritores estatísticos (como média, máximo, mínimo ou desvio padrão) para os parâmetros extraídos, organizados segundo tabelas.

5.1 Sinais de eletrocardiografia

Neste sub-capítulo mostrar-se-á o processamento feito nestes sinais após a sua recolha, nomeadamente das principais funcionalidades dos diferentes algoritmos de extração de picos desenvolvidos, das fórmulas aplicadas na extração de parâmetros e do modo como toda a informação ficará organizada para ser analisada posteriormente.

5.1.1 Principais funcionalidades

Todas as funcionalidades implementadas tinham como objetivo final a extração de características destes sinais, analisando os gráficos caracterizadores da sua evolução. A escolha destas características vai de encontro ao objetivo primordial do estudo, que é o estudo de alterações relacionadas com o período do sinal e a sua variabilidade – o *jitter*. Atendendo a esse facto, os tipos de *jitter* extraídos do sinal foram os seguintes: *jitter* simples, *Relative Average Perturbation* (RAP), *Five-Point Period Perturbation Quotient* (PPQ5) e *Eleven-Point Period Perturbation Quotient* (PPQ11). Este último apenas foi possível de obter em parte dos segmentos, dada a reduzida duração de alguns registos e, por consequência, o baixo número de picos.

As fórmulas utilizadas para o cálculo destas valores (*jitter*, PPQ e período médio) foram as que se encontram nas equações 5.1, 5.2 e 5.3, respetivamente [9]. Na medida de PPQ_x, o valor de x é um número inteiro ímpar, que representa variações do período em porções do sinal de voz.

$$Jitter(\%) = \frac{1}{T_0} \times \frac{1}{n_T - 1} \times \sum_{i=1}^{n_T - 1} \|T_{i-1} - T_i\| \times 100\% \quad (5.1)$$

$$PPQ_x(\%) = \frac{1}{T_0} \times \frac{1}{n_T - x - 1} \times \sum_{i=t+1}^{n_T - 1} \left| \sum_{k=i-t}^{i+t} \frac{T_k}{x} * T_i \right| \times 100\% \quad (5.2)$$

$$T_0 = \frac{1}{n_T} \times \sum_{i=1}^{n_T} T_i \quad (5.3)$$

Para além do cálculo destas grandezas, extraíram-se também outras que serão úteis na análise a ser feita mais tarde, de que é exemplo o número médio de batimentos cardíacos por minuto (BPM).

Para uma visualização simultânea de todos os segmentos de ECG, gerou-se uma imagem que confrontasse todos os sinais, para cada parte do teste (9 estímulos). Deste modo, seria possível fazer uma inspeção visual preliminar e, eventualmente, encontrar alguma alteração mais significativa (ver figura 5.1).

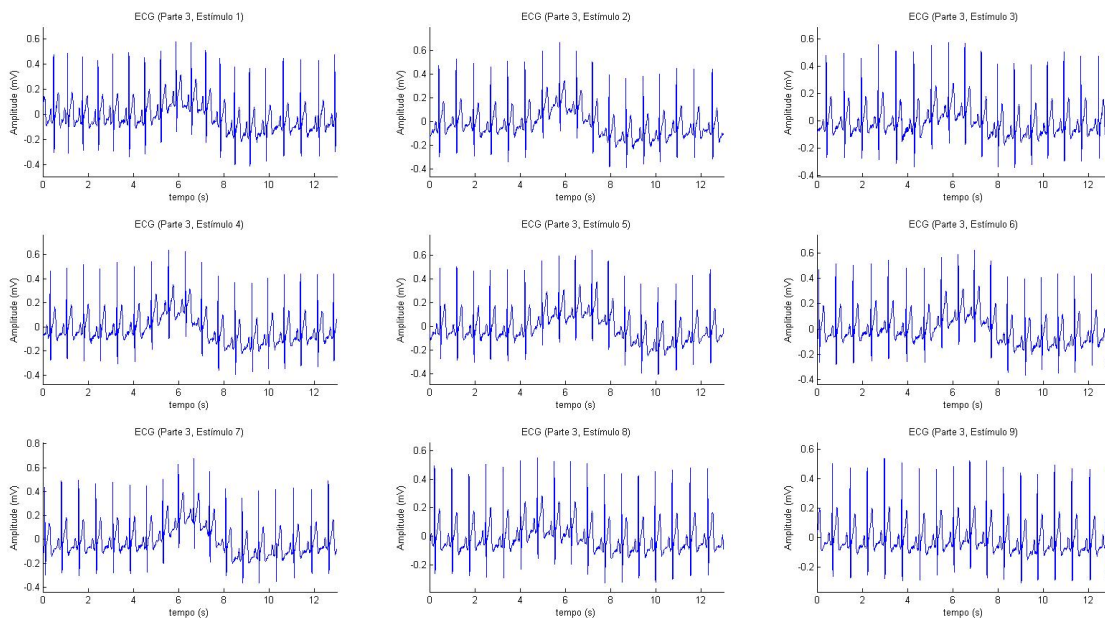


Figura 5.1: Sinais de ECG para a parte 3 do protocolo experimental (vozeamento de vogais).

Depois da análise visual ser feita, os parâmetros quantitativos foram extraídos e escritos para um ficheiro XLS (*Microsoft Excel Format*) para depois serem usados em ambiente SPSS.

5.1.2 Extração de Picos (método automático)

Numa fase inicial, a detecção de picos no sinal de eletrocardiografia foi feita recorrendo a um algoritmo de detecção de picos (*peakfinder*), da autoria de Yoder (2009). A base de funcionamento deste algoritmo consiste em fazer uma correlação entre um conjunto P-QRS-T escolhido arbitrariamente com o resto do sinal de ECG. Para os valores de correlação acima de um limiar estipulado, a função deteta um pico e retorna à posição (amostra) em que este se encontra.

Com a localização dos picos e a frequência de amostragem do sinal (mil amostras por segundo), o cálculo do período instantâneo do sinal é simples de fazer. O esquema da figura 5.2 ilustra o seu funcionamento.

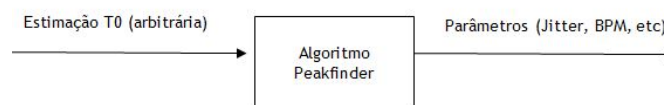


Figura 5.2: Funcionamento do algoritmo automático.

A principal limitação deste método reside no facto de utilizar uma estimativa totalmente automática (e, por consequência, sujeita a erros), do período do sinal de ECG para extração do conjunto P-QRS-T para se fazer a correlação. Esta situação, associada a sinais com maior quantidade de ruído nos quais a amplitude da onda T se aproxime da do complexo QRS, pode levar à detecção de “falsos picos” que contribuirão para uma degradação na fiabilidade dos resultados. Posto isto, tentou-se fazer uma adaptação deste algoritmo, tornando-o mais eficiente.

5.1.3 Extração de Picos (método semi-automático usando *peakfinder*)

Após alguns testes ao método automático, chegou-se à conclusão que a introdução de um método semi-automático iria trazer uma maior fiabilidade ao algoritmo. Se o período do sinal de ECG pudesse ser estimado manualmente (seleccionando os picos do sinal com o rato), faria com que a sua previsão não estivesse tão sujeita a erros como no caso anterior, em que o período era um valor constante, utilizado indiscriminadamente em todos os segmentos.

Para garantir melhores resultados no cálculo da correlação, introduziu-se um método semelhante para extração de pontos já descrita. Existe no gráfico do sinal de eletrocardiografia uma interface que solicita ao utilizador que indique, com o rato, o início e o fim de um complexo QRS. Fazendo a correlação desta porção de sinal com o sinal completo, torna-se mais eficiente o processo de detecção de picos. O esquema da figura 5.3 e a imagem representativa dos picos de correlação (figura 5.4) ilustram o funcionamento deste algoritmo, em que o eixo das abcissas corresponde à amostra atual e o das ordenadas ao valor da correlação.

Foram feitos testes preliminares deste algoritmo, chegando à conclusão que também não seria completamente eficaz. Deste modo, optou-se pelo desenvolvimento de um outro algoritmo com uma abordagem totalmente diferente.



Figura 5.3: Funcionamento do algoritmo semi-automático usando *peakfinder*.

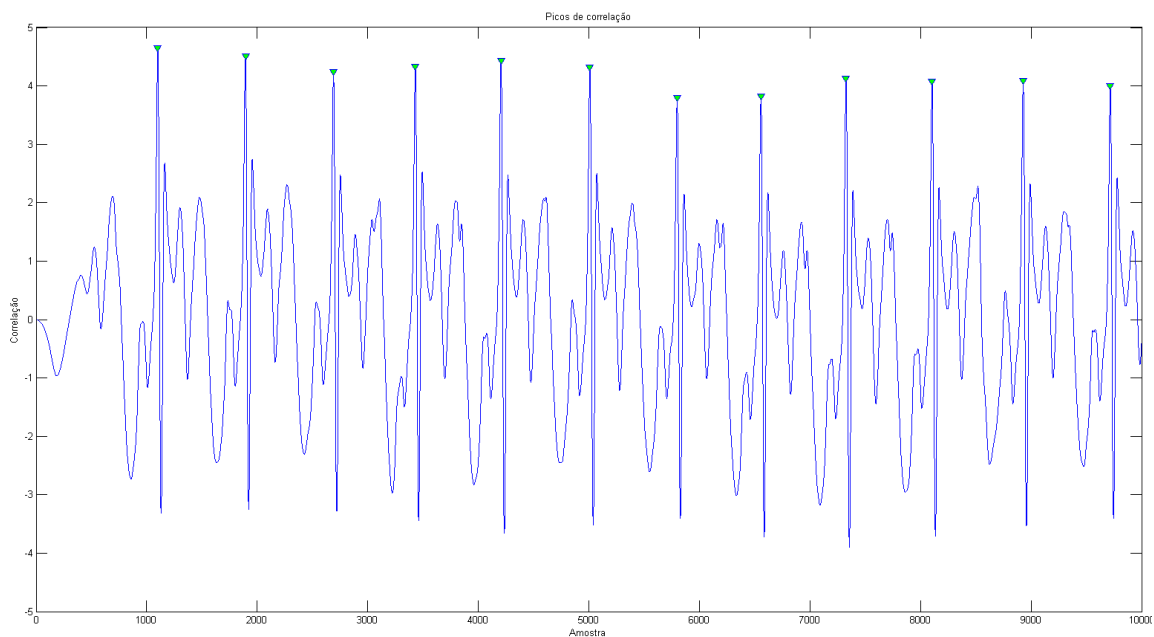


Figura 5.4: Representação de picos de correlação provenientes do algoritmo semi-automático *peakfinder* (caso de sucesso).

5.1.4 Extração de Picos (método semi-automático dos limiares)

Neste algoritmo, o paradigma é totalmente diferente relativamente às anteriores. Consiste na definição de um limiar de amplitude através do desenho de uma reta (definindo dois pontos que a formarão diretamente no sinal de ECG, com auxílio do rato) que, no caso ideal, “isole” os picos do sinal de ECG. Depois, para todos os pontos que estejam acima desse limiar, verifica-se a posição no eixo das abcissas daquele com maior amplitude, retornando o seu índice. Por último, segue-se uma abordagem semelhante às anteriores no que diz respeito a cálculo do período e dos outros parâmetros a ele associados. O seu princípio de funcionamento está esquematizado na figura 5.5. Após a fixação do limiar, é bem visível a detecção dos complexos QRS (ver figura 5.7, onde se apresenta um gráfico amostras *versus* tensão elétrica (mV)).



Figura 5.5: Funcionamento do algoritmo semi-automático dos limiares.

A figura 5.6 ilustra o funcionamento deste algoritmo. Os pontos a verde são os que têm amplitudes superiores às do limiar (a vermelho), logo potenciais picos.

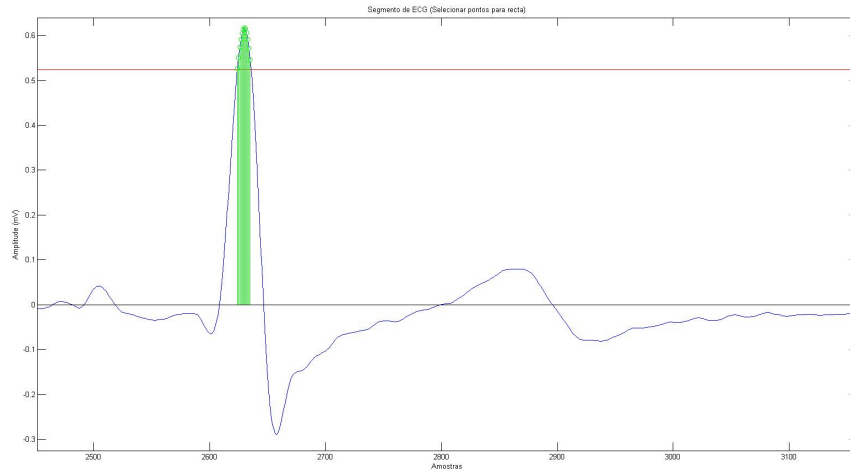


Figura 5.6: Detecção de picos pelo método dos limiares.

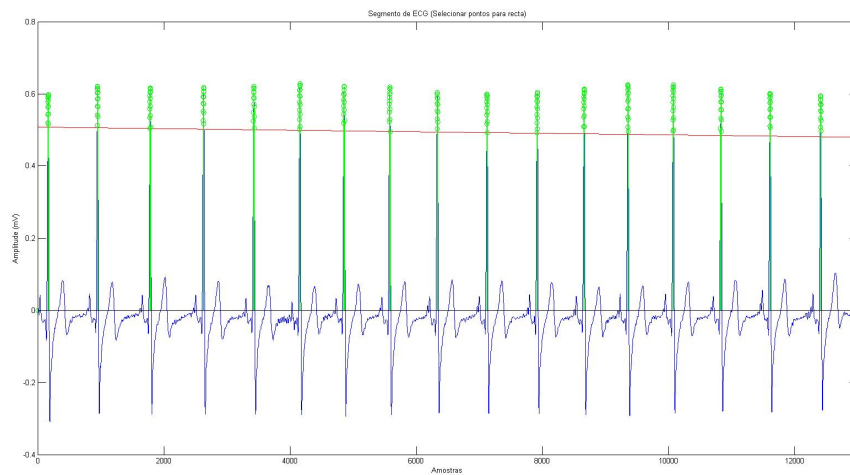


Figura 5.7: Complexos QRS após detecção.

No entanto, este algoritmo também apresenta uma limitação. Os sinais de ECG tinham, por vezes, variações que tornavam impossível desenhar uma reta em que os pontos de sinal de ECG que a superavam fossem exclusivamente pertencentes a complexos QRS. Por vezes ocorria a detecção de ondas T, que reduziam o período e, para aqueles segmentos em específico, faziam aumentar o *jitter*. Um exemplo deste caso pode ser visto na figura 5.8.

A resolução desta vulnerabilidade consistiu na criação do conceito de zona. Se cada zona correspondesse a um complexo QRS, bastaria incluir uma condição que fizesse o valor das abcissas aumentar de metade do período do sinal para que não fossem encontradas ondas T indevidas,

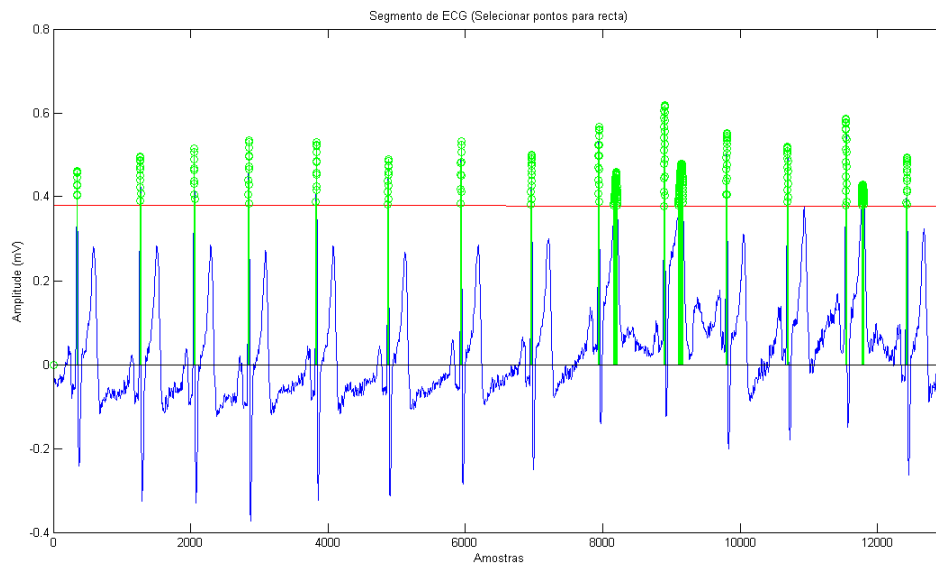


Figura 5.8: Exemplo de detecção indevida de ondas T no algoritmo da reta em sinal de ECG com ruído.

que ocorrem sempre nesta janela temporal. Esta abordagem assume o pressuposto que os dois primeiros complexos QRS estão isolados e acima da reta – o que, devido ao facto do método ser semi-automático, é uma condição de simples verificação.

Depois de extraídos os picos, a metodologia usada para cálculo dos parâmetros desejados é feita de um modo semelhante para os métodos anteriores.

Para efeitos de validação deste algoritmo, foi feito um conjunto de testes de sujeitos e estímulos aleatórios provenientes da amostra recolhida. Cruzando os resultados obtidos pelo algoritmo e pela visualização e leitura manuais do valor dos picos, conseguiu-se uma capacidade discriminativa de 100 % (zero falsos alarmes e zero omissões) para 10 % da amostra total, valor que conferiu confiança suficiente para avançar para a extração de características utilizando este algoritmo.

5.1.5 Parâmetros da análise de sinais de eletrocardiografia

Antes da análise preliminar de resultados nos sinais de ECG, importa conhecer os parâmetros exatos que foram extraídos destes sinais. Foram eles: o tamanho do registo (em amostras), número de picos (complexos QRS), *jitter* simples, *jitter* quadrático (em vez da diferença absoluta, foi calculado o quadrado dessa diferença), RAP, PPQ5, PPQ11, batimentos por minuto e período fundamental médio. Foi também ensaiada a fórmula de cálculo do PPQ55, no entanto como os segmentos não ultrapassavam os 13 segundos, o seu cálculo não foi possível uma vez que não havia picos suficientes no eletrocardiograma.

Uma vez que não foi possível extrair valores de *shimmer* ou de ruído, o conjunto de parâmetros apresentou-se como o mais apropriado dadas as limitações existentes.

Para os parâmetros de ECG extraídos, foi feita uma análise estatística, na qual foram calculados descritores como máximo, mínimo, média, desvio padrão e variância que serão descritos em maior pormenor no capítulo 6.

5.2 Sinais de voz

Neste sub-capítulo descreve-se o processamento aplicado nos sinais de voz após a sua recolha. A sua análise foi feita efetuando uma segmentação manual cuidada em cada uma das gravações, de modo a isolar os impulsos de cada vogal sustentada para proceder aos cálculos seguintes com rigor.

No caso dos sinais de voz, não houve muita algoritmia envolvida, uma vez que o objetivo era apenas o de analisar os sinais num ambiente apropriado, como é o caso do Praat. Este software é capaz de produzir um relatório com um grande conjunto de parâmetros para serem usados noutro ambiente de análise. Um exemplo de um sinal de voz (/a/), apresenta-se na figura 5.9.

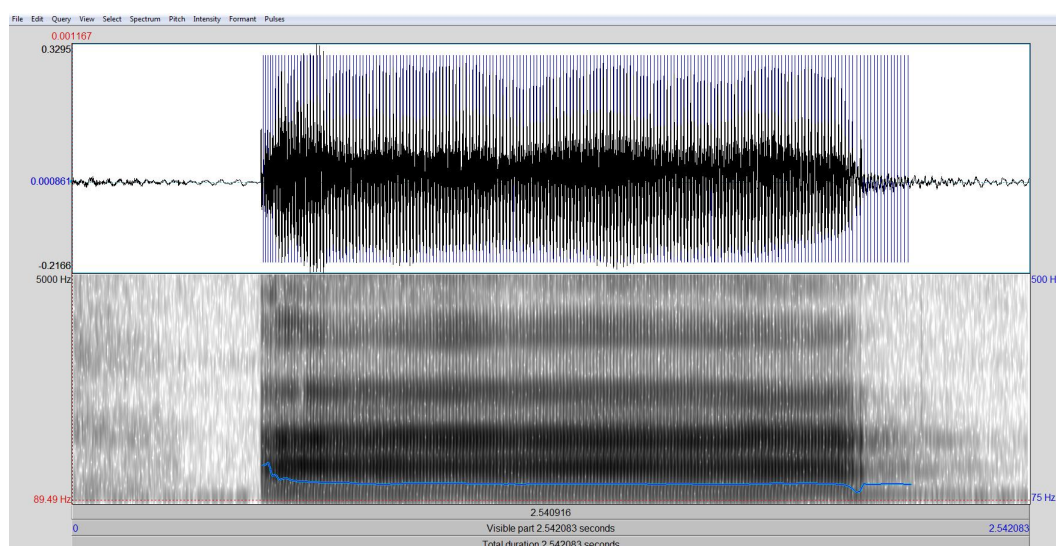


Figura 5.9: Sinal de voz para análise.

Para cada amostra, foi necessário fazer a seleção dos impulsos nos quais o sinal está estacionário (representados na imagem a azul), para obtenção do “Voice Report”. Um exemplo deste relatório pode ser visto na figura 5.10. Os parâmetros extraídos foram: duração do registo, mediana, média, desvio-padrão, mínimo e máximo da frequência fundamental, número de impulsos, média e desvio padrão do período fundamental, *jitter* simples, RAP, PPQ5, DDP, *shimmer* simples, APQ3, APQ5, APR11, DDA, autocorrelação média, NHR (*Noise-to-harmonic ratio*) e HNR (*Harmonic-to-noise ratio*). Estes valores foram extraídos para cada uma das vogais (/a/ e /i/).

Para otimizar todo este processo foi criado um *script* para leitura de todos os ficheiros, e uma rotina para extração de valores numéricos dos relatórios, de forma a ter todos os parâmetros organizados numa tabela. Estes parâmetros foram alvo de uma análise estatística semelhante à dos sinais eletrocardiográficos, devidamente explicada no capítulo 6.

```

-- Voice report for 163, Sound voz_sujeito9_parte3_est1_i --
Date: Wed May 23 14:16:36 2012

Time range of SELECTION
From 0,072062 to 1,226062 seconds (duration: 1,154000 seconds)
Pitch:
  Median pitch: 195,748 Hz
  Mean pitch: 195,338 Hz
  Standard deviation: 3,764 Hz
  Minimum pitch: 186,352 Hz
  Maximum pitch: 233,107 Hz
Pulses:
  Number of pulses: 225
  Number of periods: 224
  Mean period: 5,122550E-3 seconds
  Standard deviation of period: 0,094453E-3 seconds
Voicing:
  Fraction of locally unvoiced frames: 0 (0 / 462)
  Number of voice breaks: 0
  Degree of voice breaks: 0 (0 seconds / 1,154000 seconds)
Jitter:
  Jitter (local): 0,510
  Jitter (local, absolute): 26,143E-6 seconds
  Jitter (rap): 0,301
  Jitter (ppq5): 0,265
  Jitter (ddp): 0,902
Shimmer:
  Shimmer (local): 3,947
  Shimmer (local, db): 0,377 db
  Shimmer (apq3): 1,366
  Shimmer (apq5): 2,496
  Shimmer (apq11): 5,179
  Shimmer (dda): 4,098
Harmonicity of the voiced parts only:
  Mean autocorrelation: 0,991516
  Mean noise-to-harmonics ratio: 0,009982
  Mean harmonics-to-noise ratio: 24,821 db

```

Figura 5.10: Exemplo de *Voice Report* do Praat.

Para efeitos de comparação, foi obtido um conjunto de parâmetros nos sinais de voz utilizando o software Voice Studio. Tendo em conta que as fórmulas para cálculo de *jitter*, *shimmer* e outros parâmetros eram semelhantes, esta abordagem apenas serviu para confirmar que os resultados eram coincidentes.

5.3 Considerações finais

Ao longo deste capítulo foi descrito todo o processo de implementação da algoritmia deste estudo. Todo o conhecimento teórico adquirido previamente foi fundamental para que esta fase avançasse ao melhor ritmo possível. O facto do algoritmo que estava previsto utilizar inicialmente (*peakfinder*) não se ter relevado completamente funcional devido às variações dos sinais eletrocardiográficos obrigou a algumas adaptações que fizeram com que a fase de extração de características durasse um pouco mais que o previsto.

A fase de teste do algoritmo utilizado revelou-se importante para trazer maior rigor aos valores extraídos, apesar da demora verificada. Numa fase de extração de parâmetros, optou-se pelo princípio de extrair a maior quantidade possível, para que na análise estatística o número de possibilidades fosse maior. Nessa fase, seria feita uma filtragem, eliminando os casos que fizessem menos sentido.

Capítulo 6

Análise estatística de resultados

Todos os estudos de variáveis métricas necessitam de uma base estatística sólida para suportar as suas conclusões de um modo credível. Ao longo deste capítulo, serão explicados os principais mecanismos que contribuem para o tratamento estatístico da informação deste estudo. Inicialmente, será provada a validade da amostra, no que diz respeito à sua dimensão e variabilidade.

Geralmente, uma análise estatística é feita segundo duas abordagens diferentes: descritiva e inferencial. Numa análise descritiva, é obtido um conjunto de medidas que caracterizam as funções densidade de probabilidade para cada um dos parâmetros extraídos, como média, desvio padrão, variância, mediana, entre outros. A outra perspetiva consiste numa análise inferencial, na qual são levantadas hipóteses que, mediante o resultado do teste estatístico efetuado, são ou não confirmadas. Estas duas abordagens serão abordadas por esta ordem ao longo deste capítulo.

Para esta análise estatística utilizou-se a versão mais recente do software SPSS (*Statistic Package for the Social Sciences*) que se encontra descrito na secção 3.2. Antes de passar para a análise nesta ferramenta, é necessário garantir que toda a informação se encontra organizada no formato correto.

6.1 Testes da qualidade da amostra

Para avaliar a validade estatística da amostra a ser analisada, é necessário fazer testes de qualidade de ajuste. O teste geralmente escolhido para este fim é o qui-quadrado. Este teste permite, para uma amostra aleatória, aferir a proximidade entre o valor teórico esperado para uma distribuição e o seu valor real [33].

Depois de definidas as hipóteses do teste, deve proceder-se ao agrupamento das amostras em classes (ou categorias) não sobrepostas, cujos valores sejam qualitativos ou quantitativos discretos. No caso deste estudo em concreto, tal situação cumpre-se, uma vez que estamos perante elementos dos dois sexos, em condições em que existe (ou não) estímulo.

A estatística de teste é feita com base numa medida que avalia o ajuste entre as frequências observadas e as esperadas, através da fórmula presente na equação 6.1 [33]. Nesta fórmula, N corresponde ao valor observado e e_k ao valor esperado. Como é óbvio, a qualidade da amostra

será tanto melhor quanto menor o valor de Q . Os graus de liberdade da distribuição são calculados usando a equação 6.2, em que K é o número de classes e R o número de parâmetros da distribuição [33].

$$ET = Q = \sum_{k=1}^K \frac{(N_k - e_k)^2}{e_k} \quad (6.1)$$

$$GL = (K - 1) - R \quad (6.2)$$

Fixando um nível de significância, a rejeição ou não rejeição da hipótese faz-se comparando o valor de ET com o valor da distribuição em questão (para o ponto equivalente ao grau de liberdade calculado). Caso a hipótese não seja rejeitada, a proximidade do valor de ET relativamente ao da distribuição aumenta proporcionalmente ao tamanho da amostra analisada e com o número de observações esperadas.

Neste caso concreto, não é possível calcular o valor exato da estatística de teste uma vez que os valores teóricos não são conhecidos *à priori*. No entanto, pode-se aplicar uma regra prática para que o teste possa ser utilizado com confiança tendo em conta a dimensão da amostra ($N \geq 30$) e a frequência esperada para cada uma das classes ($e \geq 5$) [33]. Face aos vinte participantes existentes, às diferentes condições de aquisição de amostras (com e sem estímulo) e ao número de estímulos apresentado a cada participante (nove), verifica-se que estes dois requisitos são cumpridos, indiciando que a quantidade de amostras recolhidas assegura a qualidade das estatísticas visadas.

6.2 Análise descritiva

Nesta secção será feita uma análise estatística descritiva de todos os parâmetros extraídos, procurando caracterizá-los do modo mais completo possível. Para cada um dos parâmetros dos sinais de voz e eletrocardiograma, foi extraído o seguinte conjunto de descritores: número de amostras, média, mediana, desvio padrão, variância, enviesamento, curtose, mínimo e máximo. Não será feita uma análise exaustiva de todo este conjunto de descritores. As tabelas que reúnem toda esta informação podem ser consultadas no anexos B, C e D.

6.2.1 Sinais de eletrocardiograma

Para esta análise, os valores destes descritores foram obtidos para condições de repouso e com apresentação de estímulos. Um exemplo destes descritores pode ser visto na tabela 6.1, onde se apresenta o conjunto de descritores para a totalidade da amostra, com apresentação de estímulos. Os valores máximos tão elevados de *jitter* justificam-se pelo facto de, em algumas situações, as captações de sinal de ECG terem sido relativamente curtas. Isto conduziu a casos em que uma variação no período, devido ao baixo número de picos, tivesse maior impacto na percentagem da sua variabilidade total.

Dos parâmetros apresentados, são de destacar os conceitos de enviesamento e curtose. O primeiro é uma medida da simetria da função densidade, enquanto o segundo é uma medida caracterizadora do seu pico, na qual quanto maior for o seu valor, mais destacado este será. Os descritores para as restantes situações - sujeitos do sexo masculino e feminino, com ou sem estímulo, poderão ser consultados no anexo B.

Tabela 6.1: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a totalidade de amostra

Parâmetro	<i>Jitter</i> simples (%)	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	180	180	180	83	180
Média	3,994	0,232	0,267	0,036	81,361
Mediana	3,594	0,186	0,229	0,028	81,036
Desvio Padrão	2,265	0,173	0,173	0,032	12,081
Variância	5,131	0,030	0,030	0,001	145,952
Enviesamento	1,667	4,380	2,674	2,536	0,193
Curtose	5,216	27,924	12,159	8,373	-0,401
Mínimo	0,978	0,063	0,053	0,0038	59,353
Máximo	16,277	1,482	1,417	0,193	110,485

Cada um dos parâmetros avaliados apresenta uma função densidade de probabilidade caracterizada pelos parâmetros mostrados anteriormente. As funções de densidade de probabilidade para a variante do *jitter* de cinco pontos (PPQ5) em situação de repouso e perante a apresentação de estímulo podem ser vistas na figura 6.1. Analisando os dois gráficos, não é visível a diferença existente no valor médio do *jitter* (maior no caso da presença de estímulos), mas é possível verificar uma diferença na curtose dos dois gráficos, mais acentuada quando na presença de estímulo.

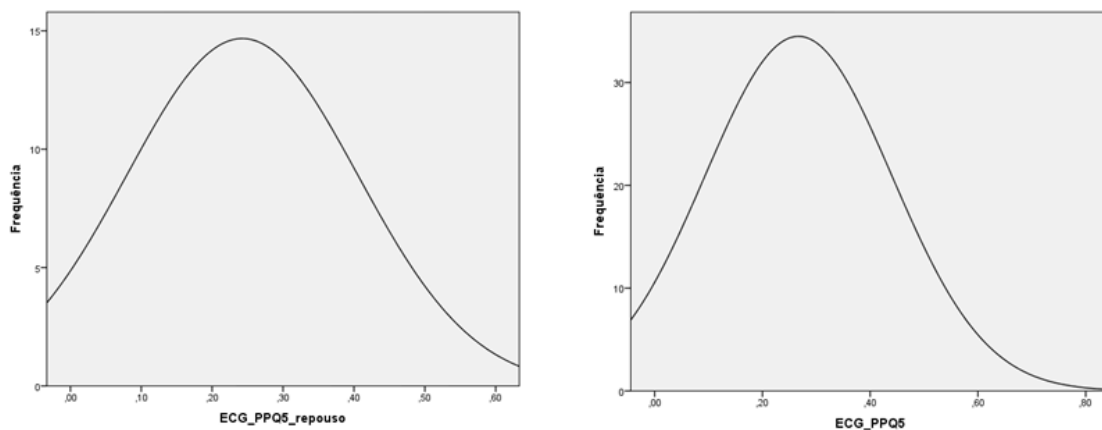


Figura 6.1: Funções de distribuição do PPQ5 para situação de repouso (esquerda) e com apresentação de estímulos (direita).

6.2.2 Sinais de voz

Para os parâmetros de voz extraídos, foi feita uma análise estatística descritiva, na qual foi extraído um conjunto de descritores. Estes valores foram obtidos para os seguintes casos:

- Vozeamento da vogal /a/, em situação de repouso;
- Vozeamento da vogal /a/, com apresentação de estímulo;
- Vozeamento da vogal /i/, em situação de repouso;
- Vozeamento da vogal /i/, com apresentação de estímulo;

Um exemplo de compilação destes descritores pode ser visto nas tabelas 6.2 e 6.3, no caso da análise da totalidade de amostra com apresentação de estímulo.

Tabela 6.2: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	177	176	176	174	174	174
Média	153,512	0,943	0,496	0,433	5,198	2,515
Mediana	152,083	0,617	0,318	0,323	4,708	2,258
Desvio Padrão	42,661	1,304	0,758	0,610	2,621	1,273
Variância	1819,945	1,700	0,574	0,372	6,869	1,620
Enviesamento	0,268	5,262	6,200	8,079	1,746	1,535
Curtose	-1,373	32,396	45,459	78,543	4,812	3,590
Mínimo	101,135	0,131	0,066	0,076	1,773	0,790
Máximo	245,020	11,076	7,258	6,917	17,203	8,171

Tabela 6.3: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	171	168	177	177	177
Média	3,164	4,132	0,936	0,107	17,653
Mediana	2,862	3,796	0,977	0,026	18,715
Desvio Padrão	1,6160293	2,0792488	0,1158767	0,2276207	6,0576330
Variância	2,612	4,323	0,013	0,052	36,695
Enviesamento	1,546	1,372	-2,796	2,906	-1,403
Curtose	3,401	2,267	6,640	7,200	1,960
Mínimo	1,046	1,415	0,515	0,0021	0,257
Máximo	10,018	13,232	0,998	0,957	28,395

As tabelas referentes a todos os casos analisados poderão ser consultadas nos anexos C e D.

6.3 Testes de correlação

Depois da amostra estudada estar validada e caracterizada estatisticamente, importa compreender quais os métodos a utilizar para provar as relações entre os vários parâmetros a serem analisados. Para tal, uma ferramenta indispensável são os testes de correlação. De um modo muito simplificado, pode definir-se o coeficiente de correlação como uma medida da relação linear entre duas variáveis numéricas.

Dentro das possibilidades existentes no software utilizado, podem-se destacar dois tipos de testes: de Pearson ou de Spearman. A correlação de Pearson é utilizada para dados numéricos com uma distribuição normal, sendo o coeficiente de correlação dado pela equação 6.3, em que X e Y são as amostras a serem relacionadas [34].

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}} \quad (6.3)$$

Para a correlação de Spearman, o método de cálculo é ligeiramente diferente, uma vez que entra com as diferenças de valores de X e Y . É mais utilizado para análise de correlação de dados não paramétricos e o seu coeficiente calcula-se segundo a equação 6.4, em que d_i é a diferença entre cada valor X e Y e n o número de pares de dados [35].

$$r_s = \frac{6\sum d_i^2}{n^3 - n} \quad (6.4)$$

Os valores para estes coeficientes de correlação variam entre -1 (associação negativa completa) e +1 (associação positiva completa). Este fator de correlação é uma medida da relação linear entre duas variáveis, sendo tanto maior quanto melhor esta for. O quadrado do valor deste coeficiente denomina-se por coeficiente de determinação (r^2).

O valor deste coeficiente, só por si, não é suficiente para exprimir a relevância da relação entre as variáveis. Para isso existe um outro parâmetro, a significância. Esta pode ser calculada através da equação 6.5, em que r é o coeficiente de correlação e n o número de graus de liberdade [35].

$$t_r = \frac{r}{\sqrt{1-r^2}} \sqrt{n-2} \quad (6.5)$$

Como se pode facilmente constatar, o tipo de correlação mais apropriado para o estudo é o de Pearson, uma vez que nenhum dos parâmetros a analisar é de natureza nominal, ordinal ou na forma de intervalos. Apenas são comparados valores numéricos.

No caso prático da análise em SPSS, os valores de significância considerados relevantes são os que se encontram abaixo de 0.05, existindo um destaque visual fornecido para estes casos. Um exemplo de uma das tabelas de correlações pode ser visto na figura 6.2. Como se pode ver, a relação linear entre os valores da frequência fundamental em situação de repouso é forte ($r = 0.802$, $p = 0.001$). No entanto, para os outros casos analisados, apesar do valor de significância cumprir o requisito para que o resultado seja estatisticamente relevante, o valor do coeficiente de

correlação é sugestivo de relações fracas ou moderadas ($r < 0.5$). Como seria de esperar, esta matriz de correlações é simétrica e apresenta o valor 1 para relações entre os mesmos parâmetros.

		ECG_RAP	ECG_PPQ5	ECG_BPMedia	Voz_A_pitchmedia	Voz_A_JitterRAP	Voz_A_JitterPPQ5
ECG_RAP	Pearson Correlation	1	,932**	-,166*	,123	-,039	-,037
	Sig. (2-tailed)		,000	,026	,104	,604	,624
	N	180	180	180	177	176	174
ECG_PPQ5	Pearson Correlation	,932**	1	-,131	,240**	-,064	-,075
	Sig. (2-tailed)	,000		,079	,001	,400	,323
	N	180	180	180	177	176	174
ECG_BPMedia	Pearson Correlation	-,166*	-,131	1	,297**	-,108	-,102
	Sig. (2-tailed)	,026	,079		,000	,153	,179
	N	180	180	180	177	176	174
Voz_A_pitchmedia	Pearson Correlation	,123	,240**	,297**	1	-,209**	-,213**
	Sig. (2-tailed)	,104	,001	,000		,005	,005
	N	177	177	177	177	176	174
Voz_A_JitterRAP	Pearson Correlation	-,039	-,064	-,108	-,209**	1	,840**
	Sig. (2-tailed)	,604	,400	,153	,005		,000
	N	176	176	176	176	176	174
Voz_A_JitterPPQ5	Pearson Correlation	-,037	-,075	-,102	-,213**	,840**	1
	Sig. (2-tailed)	,624	,323	,179	,005	,000	
	N	174	174	174	174	174	174

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Figura 6.2: Exemplo de matriz de coeficientes de correlação (para relação entre parâmetros de ECG e voz).

Após uma análise de correlação exaustiva para os contextos possíveis - relação entre sinais de voz e eletrocardiográficos em condição de repouso e estimulação - chegou-se à conclusão que os valores de correlação e significância não eram significativos ao ponto de conferir o suporte estatístico necessário, pelo que se optou por enveredar por uma análise correlacional mais abrangente.

6.4 Análise de modelos de regressão

Analisando os dados provenientes da estatística descritiva feita anteriormente, pretende-se agora efetuar uma inferência estatística a partir da qual se estabeleça uma relação linear ou não linear envolvendo pares de parâmetros. Os modelos de regressão são uma ferramenta poderosa utilizada para estimar, através de uma função (linear, logarítmica, exponencial ou outra) uma variação de um conjunto de parâmetros. Para uma análise preliminar do melhor método de regressão a utilizar costuma-se recorrer a gráficos de dispersão de pontos.

Para o caso da regressão linear entre uma variável independente X e uma variável dependente Y , a relação entre as duas grandezas (a equação da reta) vem dada pela equação 6.6, em que n é o índice da observação, X_{med} a média aritmética das observações, α e β os parâmetros da reta a estimar e E_n o erro aleatório associado a cada observação. A equação 6.6 assume que não existe qualquer erro nos valores de X [33].

$$Y_n = \alpha + \beta(X_n - X_{med}) + E_n \quad (6.6)$$

Para a criação de um modelo deste género, considera-se que um conjunto de hipóteses se encontra devidamente verificado [33]:

- E_n terá de ter valor esperado nulo e variância constante;
- Os erros serão mutuamente independentes e apresentam uma distribuição normal.

Os valores dos parâmetros de regressão são tipicamente obtidos recorrendo ao método dos mínimos quadrados.

Este tipo de modelo de regressão é analisado recorrendo a um gráfico de dispersão. Este terá os pontos mais afastados para valores mais baixos do coeficiente de correlação. Para evitar esta situação, existem modelos de regressão mais complexos que recorrem a outro tipo de funções de aproximação. Em todo o caso, todos eles partem de modelos linearizados aos quais é aplicada uma transformação de variável independente.

Para o estudo em questão, era necessário primeiro criar um gráfico com a dispersão de pontos. Depois, face ao conjunto de funções disponíveis no software para fazer a aproximação (linear, logarítmica, quadrática, cúbica, exponencial, inversa, entre outras), encontrar aquela para a qual o valor do coeficiente de determinação (r^2) fosse maior e verificar se a curva resultante se apresentaria ou não como uma boa aproximação à relação a ser testada.

Após o estudo de regressão feito para as variáveis em questão, chegou-se à conclusão que a dispersão de pontos existente não garantia uma relação com a aproximação desejada. Um exemplo disso encontra-se no gráfico da figura 6.3, que apresenta uma dispersão de pontos em que o eixo das abcissas representa o PPQ5 do eletrocardiograma e o das ordenadas o PPQ5 da vogal /i/.

Pode então dizer-se que existe a relação matemática entre parâmetros de voz e de eletrocardiograma e que esta é caracterizada em cerca de 65-70 % da sua variância, mas que estes resultados não são suficientes para assumir que existe uma função que os relacione de forma inequívoca. Uma metodologia que poderia melhorar este resultado seria a identificação de pontos que se encontram bastante afastados do resto da distribuição (*outliers*), fazendo uma nova regressão.

6.5 Comparações entre distribuições

Um modo de verificar se existe alguma relação entre as distribuições de probabilidade dos sinais de voz e de eletrocardiograma é através do teste estatístico de Kruskal-Wallis. Ao contrário dos testes de análise e comparação de variância (ANOVA), este teste serve para comparar, de um modo não paramétrico, uma distribuição ao longo dos seus grupos.

Sabendo que os sinais de voz e eletrocardiograma foram recolhidos em duas condições diferentes definidas pela ausência ou presença de estímulo, pode-se facilmente ver se as distribuições diferem (nomeadamente em termos de mediana, que é o parâmetro que define a hipótese nula neste teste). Os parâmetros extraídos são o valor da estatística de teste (H), o número de graus de liberdade (df) e a significância (p). Para a realização deste teste, consideraram-se relevantes valores de significância para $p \leq 0.05$.

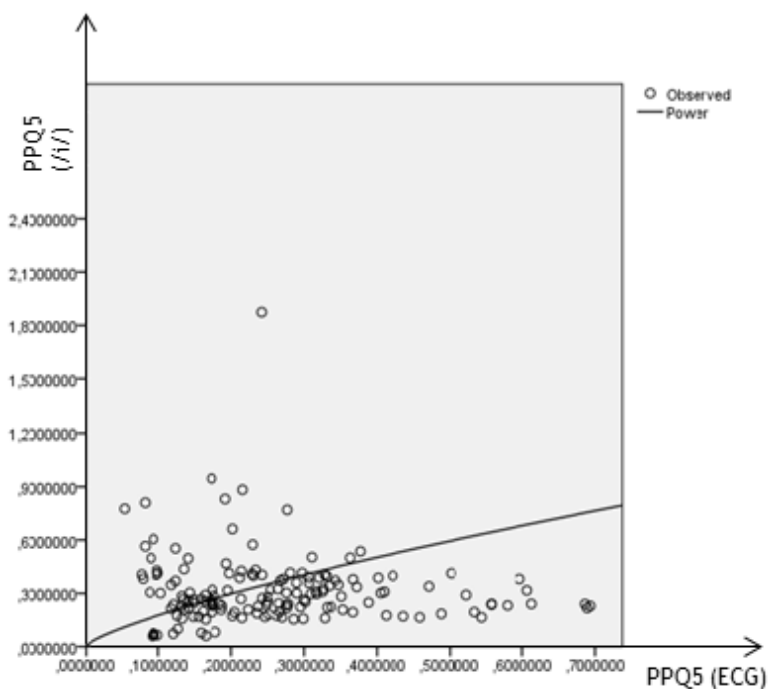


Figura 6.3: Exemplo de modelo de regressão para relação entre PPQ5 de eletrocardiograma e voz (no vozeamento da vogal /i/).

Esta situação de variação de mediana é efetivamente verificada para o caso do *jitter* do eletrocardiograma, como se pode ver pelo gráfico da figura 6.4. Para este caso, tem-se $H=5.577$, $df = 1$ e $p = 0.018$. Nestes gráficos, a linha preta representa a mediana da distribuição, e os conjuntos de valores acima e abaixo representam os quartis a 75 e 25 %, respetivamente. Os traços representam a sua gama de valores, e os outros pontos valores isolados na distribuição (*outliers*).

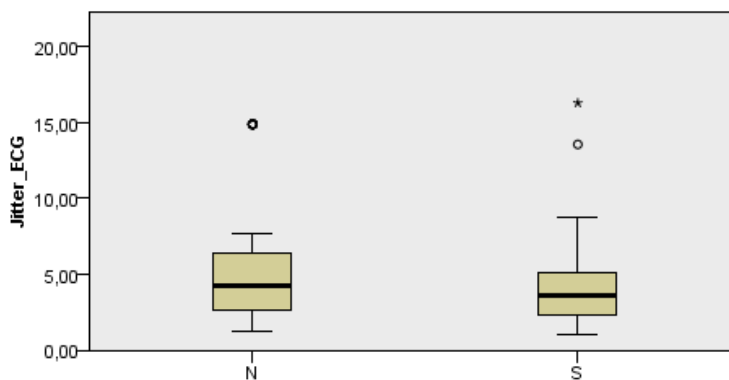


Figura 6.4: Comparação de funções de distribuição para *jitter*, sem (N) e com (S) apresentação de estímulo.

Para além desta variação no *jitter*, destaca-se ainda o aumento do valor médio dos batimentos cardíacos na situação em que se apresenta um estímulo. Este aumento não é apenas proveniente do estímulo em si, uma vez que a atividade muscular associada à fala também contribui para esta variação. No entanto, uma vez que em ambos os casos existe vozeamento semelhante, pode-se

estabelecer uma comparação justa. Para consultar valores mais pormenorizados que validam esta conclusão, ver tabelas B.1 (situação de repouso) e B.4 (com estímulo) do anexo B.

Numa comparação entre os dois géneros, conclui-se que existe uma maior variação nos batimentos cardíacos e no *jitter* na amostra do sexo feminino, especialmente no PPQ5. Ou seja, pode concluir-se que este conjunto da amostra é mais sensível a variações em sinais eletrocardiográficos perante a apresentação de estímulos. Estes resultados são confirmados pela comparação entre as tabelas B.2 (valores em repouso no sexo masculino), B.3 (valores em repouso no sexo feminino), B.5 (valores com estímulo para o sexo masculino) e B.6 (valores com estímulo para o sexo feminino) do anexo B.

No que diz respeito aos sinais de vozeamento para a vogal /a/, as principais diferenças nas distribuições estão presentes no caso da frequência fundamental e no *jitter* (RAP), com maior preponderância para o primeiro parâmetro, como se pode ver na figura 6.5 ($H=12.097$, $df=1$ e $p=0.01$).

Apesar da enorme diferença que se pode ver na mediana da distribuição dos valores da frequência fundamental, comparando as situações sem e com apresentação de estímulo (113 e 152 Hz, respetivamente), na média este valor difere um pouco menos (de 145 para 153 Hz). É também de assinalar um aumento em todas as variantes do *shimmer*. Os parâmetros de ruído (NHR e HNR) não sofrem alterações significativas. Para comparação entre valores de *jitter*, *shimmer*, HNR e NHR consultar tabelas C.1 (em repouso) e C.7 (com estímulo), do anexo C.

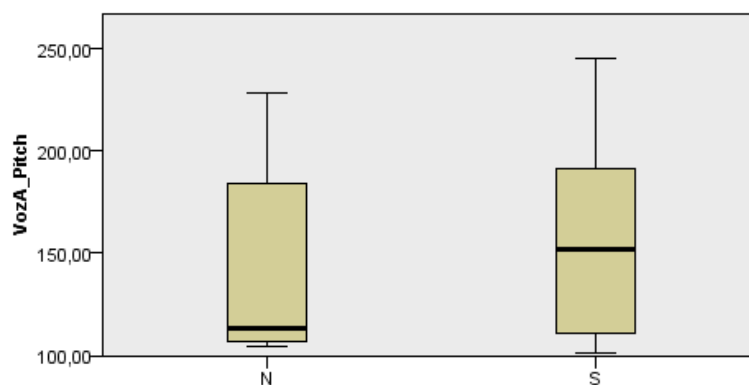


Figura 6.5: Comparação das distribuições de valores da frequência fundamental, sem (N) e com (S) apresentação de estímulo.

Comparando os sujeitos do sexo masculino e feminino, conclui-se que existe uma variação em dois sentidos diferentes na frequência fundamental, quando se está na presença de um estímulo. Enquanto que no caso masculino esta baixa (ficando a voz um pouco mais grave), nas mulheres sobe (ficando mais aguda). Esta variação é maior no caso feminino, ultrapassando mesmo os 20 Hz (de 166 para 188 Hz). Relativamente a variações no *jitter*, estas são mais acentuadas no sexo masculino, com RAP e PPQ5 a variarem 0.26 e 0.236 %, respetivamente). A comparação para este caso pode ser vista através da consulta das tabelas C.3 (valores em repouso no sexo masculino), C.5 (valores em repouso no sexo feminino), C.9 (valores com estímulo para o sexo masculino) e

C.11 (valores com estímulo para o sexo feminino) no anexo C.

Por último, relacionaram-se os sinais de vozeamento para a vogal /i/, com e sem estímulos. Verificaram-se distribuições distintas para o caso do *jitter* (especialmente o PPQ5) e para o *shimmer* (APQ3) - ver figura 6.6 ($H=5.686$, $df=1$ e $p=0.017$). Como se pode constatar, existe um valor bastante superior na mediana para a apresentação de estímulo (aumento de 1.698 % para 1.928 %).

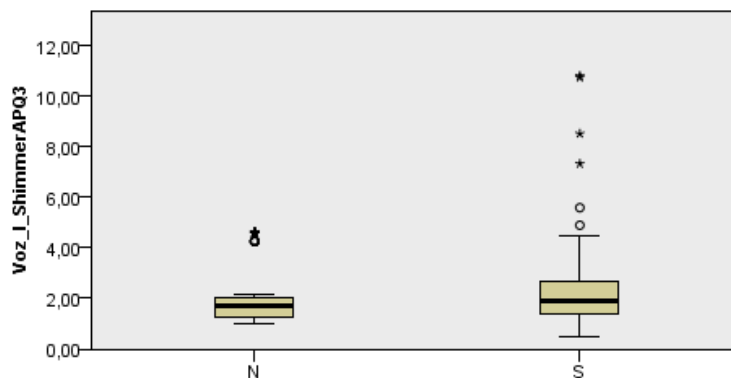


Figura 6.6: Comparação das distribuições de valores do *shimmer*, sem (N) e com (S) apresentação de estímulo

De salientar também, para a amostra total, a menor diferença na frequência fundamental, que apenas aumentou de 153 para 156 Hz. No caso masculino, houve uma descida da frequência fundamental em 6 Hz (de 127 para 121) na presença de estímulos, não tão acentuada como a subida no caso feminino (de 182 para 192). No entanto, as maiores variações do *jitter* (0.4029, 0.1616 e 0.1578 % para dois, três e cinco pontos) mantêm-se coerentes com a análise feita para a vogal /a/. Os valores dos parâmetros nos quais as conclusões tiradas se baseiam podem todos ser consultados nas tabelas do anexo D.

6.6 Considerações finais

Ao longo deste capítulo sobre análise estatística, chegou-se à conclusão que uma análise de correlação linear não chega para relacionar de forma suficientemente precisa os parâmetros de sinal de eletrocardiograma e de voz. Os modelos de regressão apresentaram-se como a alternativa mais discriminativa, abrindo um maior conjunto de possibilidades.

Através dos testes de análise de variância foi possível comparar as distribuições dos vários parâmetros através da confrontação entre condições de repouso e apresentação de estímulos, bem como de tendências coincidentes apresentadas nas distribuições de sinais de voz e eletrocardiografia. Deste modo, é possível concluir que o aumento das diversas variantes de *jitter* de sinal de eletrocardiografia é acompanhado por um aumento de *jitter* de voz, numa situação de exposição a estímulos visuais.

No que diz respeito ao gênero, conclui-se que, para o sinal eletrocardiográfico, os batimentos cardíacos e o *jitter* (PPQ5) têm aumentos mais significativos no sexo feminino quando se apresenta um estímulo, acontecendo o mesmo no caso da frequência fundamental dos sinais de voz. Quanto à frequência fundamental, os maiores aumentos dão-se igualmente no sexo feminino nesta mesma situação, sendo que as maiores variações em termo desta se encontram nos homens. Já no que diz respeito ao *jitter* de voz, o seu aumento é mais notado entre a amostra do sexo masculino. Destaque também para o aumento nos valores do *shimmer* para o caso da vogal /i/, na presença de estímulos.

Capítulo 7

Conclusões

Este estudo permitiu tirar um conjunto de conclusões que permitem relacionar parâmetros de sinal de voz e eletrocardiografia. Através de um conjunto de testes estatísticos, foi possível comprovar que a apresentação de estímulos visuais se traduz num aumento do *jitter* e que esta é uma tendência comum aos dois tipos de sinal. Neste capítulo final da dissertação, será feito um balanço de todo o trabalho realizado, analisando os resultados finais e confrontando-os com os objetivos inicialmente determinados. Será também enumerado um conjunto de dificuldades sentidas ao longo do trajeto, para além das indicações tendo em vista a continuidade deste estudo.

7.1 Satisfação dos objetivos

Para que este estudo fosse conduzido com sucesso, a compreensão de todos os conceitos associados ao funcionamento do coração, da voz e da sistema nervoso responsável pelo desencadear destes mecanismos foi determinante. Foi feita uma tentativa de explicação não demasiado exaustiva, procurando demonstrar sempre a aplicabilidade para cada caso. Posto isto, pode-se considerar que a revisão bibliográfica conferiu o conhecimento necessário à compreensão do estudo, enquadrando devidamente os leitores da dissertação.

No que diz respeito à recolha de amostras de sinal de voz e ECG, esta foi realizada igualmente com sucesso. Depois de reunido todo o material necessário e de definido o local para recolhas, foi estabelecido um procedimento pensado e estruturado que apresentava uma sequência para tudo o que era necessário fazer, minimizando a possibilidade da existência de erros. Nesta vertente, o trabalho colaborativo desenvolvido com o Mário Barbosa (cujo tema de dissertação é "Estudo e contextualização do polígrafo quanto à análise de voz, ECG e EDR") revelou-se uma ajuda importante que se refletiu na qualidade das recolhas. Importa também mencionar os voluntários que colaboraram com ambos os estudos, sem os quais este objetivo nunca teria sido alcançado.

Na fase de processamento de sinais, pode-se considerar que o objetivo associado à criação de algoritmos para extração de parâmetros destes sinais foi atingido com sucesso. Face à natureza dos sinais de eletrocardiografia, não seria possível a extração de um maior número de caracterizadores, especialmente tendo em conta as variações que alguns registos apresentavam. Teria

sido interessante ter obtido variantes de *shimmer* de ECG para confrontação com sinais de voz, mesmo atendendo a que provavelmente as relações encontradas pudessem não ser relevantes. A análise e extração de características efetuada no Praat foi também uma tarefa que conduziu a bons resultados.

Relativamente ao último objetivo de procura de relações entre sinais de voz e eletrocardiografia, pode-se considerar que foi atingido com parcial sucesso. Apesar das abordagens de correlação e regressão terem conduzido a resultados que se revelaram piores que o esperado, apresentam-se como uma abordagem válida. Os testes estatísticos de variância que comparavam as distribuições dos parâmetros extraídos revelaram-se mais conclusivos, permitindo confrontar as situações de repouso com as de apresentação de estímulo e avaliar tendências coincidentes para os dois tipos de sinais. A comparação feita para os dois sexos (masculino e feminino) também permitiu tirar conclusões relativamente à variabilidade dos resultados.

Fazendo uma comparação entre os objetivos propostos e atingidos, pode concluir-se que a maior parte destes foi superada com sucesso. Importa destacar que, numa fase inicial, estes objetivos eram um pouco diferentes. A análise de correspondências entre estes sinais seria feita apenas em situação de repouso e para um número menor de sujeitos. Além disso, o único parâmetro previsto para análise era a distribuição temporal dos impulsos glóticos e cardíacos - o *jitter*. Tendo em conta a adaptação feita, penso que o estudo ficou um pouco mais enriquecido, permitindo um estudo mais aprofundado e abrangente de conclusões.

7.2 Principais dificuldades

De todas as dificuldades sentidas ao longo desta dissertação, a principal foi a obtenção de todo o material necessário à realização dos testes. Tendo em conta a sua especificidade e a inexistência de alguns módulos na Faculdade de Engenharia, a colaboração da Faculdade de Psicologia através da cedência de local para recolha e de algum material foi determinante para ultrapassar este obstáculo. Numa fase anterior foram tentados outros contactos, nomeadamente com empresas do sector e com a Faculdade de Desporto. No entanto, o custo proibitivo da primeira solução e a indisponibilidade de recursos da segunda levaram à integração de dois módulos independentes e à definição de um sistema de marcação e sincronização próprios.

A recolha de voluntários também levantou alguns problemas, especialmente no que diz respeito ao sexo feminino. Este problema foi ultrapassado com o estabelecimento de contactos adicionais, para além do e-mail dinâmico enviado à comunidade estudantil. O processo de segmentação manual e organização da base de dados de sinais também foi uma tarefa que exigiu bastante paciência, mas indispensável ao sucesso do estudo.

O ruído nos eletrocardiogramas, na sua maior parte devido a movimentos indevidos durante as recolhas, dificultou a fase de processamento destes sinais. Foi necessária a criação de um novo algoritmo para lidar com situações pouco desejáveis em que existiam variações na morfologia das ondas do sinal que, embora não comprometessem em termos de fiabilidade os resultados, dificultaram no processo de extração de parâmetros.

7.3 Trabalho futuro

O tema deste estudo apresenta potencial para ser melhorado em algumas vertentes, das quais se destacam os seguintes casos:

- Um dos pontos no qual o estudo poderia melhorar em termos de validade estatística seria no número de voluntários. Fazendo-se também um aumento na duração do teste, poder-se-ia apresentar um maior número de estímulos, aumentando significativamente o número de casos a avaliar. Também se poderia aumentar o número de vogais analisadas (/a/, /i/ e /u/, por exemplo);
- Para evitar a quantidade de ruído existente nos sinais de eletrocardiografia, poder-se-ia aumentar o número de derivações, colocando-as no peito e nos membros. Este processo exigiria que todas as recolhas fossem refeitas em novos voluntários, uma vez que os que já fizeram o teste desenvolveram habituação relativamente aos estímulos apresentados. Além disso, seria necessário criar condições no local de recolha para que os sujeitos estivessem deitados e com a apresentação de estímulos no local correto (possivelmente direcionados para o teto);
- No caso do sinal de eletrocardiografia apresentar menor quantidade de ruído, poderia ser alvo de uma análise mais detalhada. Poderiam ser analisadas outras ondas (P, T ou U) e intervalos, caracterizando melhor o sinal como um todo. Esta análise, em paralelo com uma análise de voz mais detalhada e direcionada para a vertente patológica, poderia contribuir para um estudo de relações entre os dois sinais ao nível do diagnóstico médico.

Para além destes casos existiriam outros, com menor prioridade no que toca a melhoria. Numa primeira fase, o caminho a seguir seria o aumento do número de recolhas, uma vez que estas poderiam ser utilizadas juntamente com as anteriormente feitas.

7.4 Observações finais

É desejável que esta dissertação venha a ser útil no futuro, motivo pelo qual são dadas as indicações para futuros desenvolvimentos na secção anterior. Esta utilidade pode estar presente não só num estudo semelhante, como noutras temáticas nas quais existam pontos coincidentes.

Anexo A

Conjunto de especificações

Neste anexo constam as especificações técnicas do material utilizado no âmbito deste estudo.

A.1 Biopac MP35

Analog Inputs	Front panel DSUB 9f labeled "CH1"- "CH4"
Number of channels:	4 isolated, human-safe universal biopotential/transducer input amplifiers
A/D resolution:	24-bit
Gain ranges:	10x to 50,000x in 12 steps
Input voltage range:	Gain dependent 400 microvolts-2 V peak-peak (± 10 V with SS70L, page 39.)
Input Noise:	0.222 μV_{rms} (0.5 Hz-35 Hz)
CMRR:	85 dB minimum
Filters:	Programmable analog and digital (IIR) filters; automatic or user-adjustable
Analog Output	Back panel DSUB 9m labeled "Analog Out"
Voltage Output:	Range 0-4 V; Resolution 12-bit
Pulse Output:	Width: variable Repetition: variable
Pulse Level:	TTL-compatible with BSLSTM Stimulator: 0-10 V or 0-100 V with SS58L Stimulator: -10 V to +10 V
Input Triggering Options	
External Trigger:	Back panel BNC labeled "Trigger" TTL positive or negative edge
Analog Trigger:	Any of four input channels (front panel "CH1"- "CH4"); Threshold going positive or negative
Digital trigger:	Any of 8 input lines (back panel DSUB 25m)
Sample Rate	400 k samples/second aggregate across four channels - 100 k s/s on all four channels for short durations - 100 k s/s on three channels continuously - 50 k s/s on all four channels continuously
Electrode Check	
Impedance Range:	0-100 K Ω ; Checks electrode impedance from Vin+ and Vin- leads with respect to GND lead.
Serial InterfaceType	USB
Certification	Complies with IEC60601-1 EMC complies with IEC60601-1-2 CE Marked
Dimensions	7 cm x 29 cm x 25 cm
Weight	1.4 kg

Figura A.1: Especificações do aparelho Biopac MP35 [22].

A.2 Sennheiser Earset 1

Frequency response	20 to 20,000 Hz \pm 3 dB
Pick-up pattern	omni-directional
Nominal impedance (at 1 kHz)	1 k Ω
Min. terminating impedance	4.7 k Ω
Sensitivity (at 1 kHz)	5 mV/Pa \pm 2.5 dB
Diameter of mic capsule	3.3 mm
Max. sound pressure level	143 dB
Equivalent noise level	27 dB (A)
Supply current	approx. 240 μ A
Power supply	4.5 to 15 V
Operating temperature	-10°C to +50°C
Cable length	1.6 m
Diameter of boom arm	1.1 mm
Weight (without cable)	2.3 g
Connector	see "Variants" on page 11

Figura A.2: Especificações do microfone Sennheiser Earset 1 (Gerais) [36].

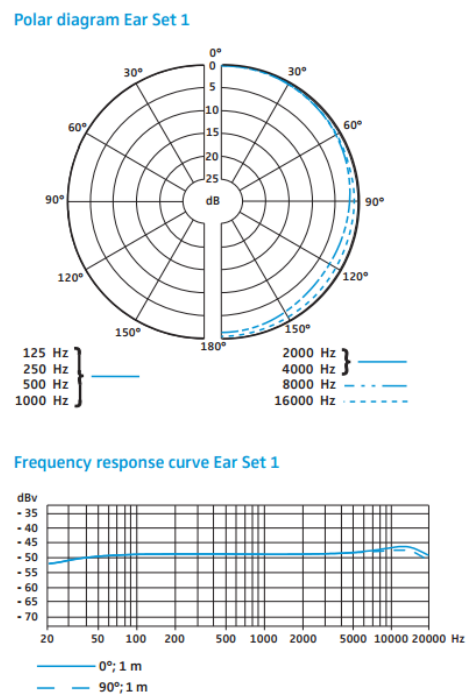


Figura A.3: Especificações do microfone Sennheiser Earset 1 (Resposta em frequência e direcionalidade) [36].

A.3 Roland UA 25EX

Number of Audio Record/Playback Channels	Residual Noise Level
Record	Output Jack
1 pair of stereo	-96 dBu (S/N ratio: 105 dB) (input terminated with 1 k Ω , OUTPUT VOLUME: max., INPUT MONITOR: min., IHF-A, typ.)
Playback	Mic Preamp
1 pair of stereo DIGITAL IN SWITCH = OFF: Full duplex (except for 96 kHz setting) DIGITAL IN SWITCH = ON: Recording only	PHANTOM POWER 48 V, ANALOG LIMITER/COMPRESSOR
Signal Processing	Connectors
PC interface	Front Panel
24 bits (ADVANCE SWITCH=OFF: 16 bits)	Input Jack L, R: XLR type (balanced, phantom power +48 V), 1/4 inch TRS phone type (balanced) * Input Jack R: high impedance is supported Headphone Jack: Stereo 1/4 inch phone type
AD/DA Conversion	Rear Panel
24 bits (linear)	Digital Input Connector: Optical type Digital Output Connector: Optical type Output Jack L, R: 1/4 inch TRS phone type (balanced), RCA phono type MIDI Connectors: IN/OUT USB Connector: B type
Sampling Frequency	Other
Digital output	Power Supply
44.1/48/96 kHz	USB Bus Power
Digital input	Current Draw
44.1/48/96 kHz	480 mA
AD/DA Conversion	Accessories
44.1/48/96 kHz * ADVANCE SWITCH=OFF: 44.1 kHz only	Owner's Manual USB cable CD-ROM
Frequency Response	Size and Weight
96.0 kHz	Width
20 Hz to 40 kHz (+0 dB/-2 dB)	164 mm 6-1/2 inches
48.0 kHz	Depth
20 Hz to 22 kHz (+0 dB/-1 dB)	133 mm 5-1/4 inches
44.1 kHz	Height
20 Hz to 20 kHz (+0 dB/-1 dB)	64 mm 2-9/16 inches
Nominal Input Level	Weight
Input Jack L, R (XLR type)	0.6 kg 1 lbs. 4 oz.
-60 to -20 dBu	* 0 dBu = 0.775 Vrms
Input Jack L, R (1/4 inch TRS phone type)	
-36 to +4 dBu * Maximum Input level: Nominal Input Level + 12 dB	
Nominal Output Level	
Output Jack L, R (1/4 inch TRS phone type)	
+4 dBu (balanced)	
Output Jack L, R (RCA phono type)	
-2 dBu * Maximum Output level: Nominal Output Level + 12 dB	

Figura A.4: Especificações de Roland UA 25EX [26].

Anexo B

Descritores estatísticos de parâmetros de sinais de eletrocardiograma

Neste anexo estarão presentes todos os descritores estatísticos de sinal de voz obtidos para os resultados deste estudo. Servirá para uma consulta mais pormenorizada, procurando esclarecer situações esporádicas relativas a grupos específicos deste estudo.

Estes descritores foram obtidos para as seguintes situações:

- Em situação de repouso: para a totalidade da amostra, para o sexo masculino e para o sexo feminino;
- Com apresentação de estímulo: para a totalidade da amostra (tabela B.4), para o sexo masculino (tabela B.5) e para o sexo feminino (tabela B.6);

Tabela B.1: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a totalidade de amostra

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	180	180	180	27	180
Média	4,778	0,224	0,242	0,0385	79,878
Mediana	4,278	0,174	0,172	0,005	78,258
Desvio Padrão	3,036	0,162	0,163	0,049	12,271
Variância	9,219	0,026	0,026	0,002	150,589
Enviesamento	1,755	1,057	0,334	0,749	0,283
Curtose	3,687	0,52	-1,324	-1,560	-0,612
Mínimo	1,229	0,036	0,0217	0,005	61,203
Máximo	14,882	0,664	0,514	0,106	104,13

Tabela B.2: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a amostra do sexo masculino

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	90	90	90	18	90
Média	4,375	0,197	0,214	0,055	79,067
Mediana	4,544	0,174	0,172	0,055	77,32
Desvio Padrão	1,868	0,109	0,149	0,052	14,234
Variância	3,489	0,012	0,022	0,003	202,612
Enviesamento	0,143	0,613	0,698	0,000	0,544
Curtose	-0,763	-0,775	-0,504	-2,267	-0,821
Mínimo	1,229	0,056	0,028	0,005	61,579
Máximo	7,701	0,399	0,514	0,106	104,130

Tabela B.3: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, em situação de repouso, para a amostra do sexo feminino

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	90	90	90	9	90
Média	5,180	0,252	0,271	0,005	80,689
Mediana	3,786	0,181	0,257	0,005	80,998
Desvio Padrão	3,837	0,198	0,173	0	9,946
Variância	14,724	0,039	0,030	0,000	98,927
Enviesamento	1,507	0,769	-0,002	-	-0,252
Curtose	-0,763	-0,775	-0,504	-2,267	-0,821
Mínimo	2,074	0,036	0,022	0,005	61,203
Máximo	14,882	0,664	0,512	0,005	97,559

Tabela B.4: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a totalidade de amostra

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	180	180	180	83	180
Média	3,994	0,232	0,267	0,036	81,362
Mediana	3,594	0,186	0,230	0,028	81,036
Desvio Padrão	2,265	0,173	0,173	0,032	12,081
Variância	5,131	,030	0,030	0,001	145,952
Enviesamento	1,667	4,380	2,674	2,536	0,193
Curtose	5,216	27,924	12,159	8,373	-0,401
Mínimo	0,978	0,064	0,054	0,004	59,353
Máximo	16,277	1,481	1,417	0,193	110,486

Tabela B.5: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a amostra do sexo masculino

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	90	90	90	32	90
Média	3,770	0,210	0,223	0,033	79,027
Mediana	2,801	0,175	0,182	0,028	76,375
Desvio Padrão	2,410	0,165	0,141	0,019	13,988
Variância	5,809	0,027	0,020	0,000	195,651
Enviesamento	1,869	5,337	2,771	0,886	0,672
Curtose	6,714	39,467	12,483	-0,109	-0,475
Mínimo	0,978	0,064	0,054	0,008	59,353
Máximo	16,277	1,482	1,043	0,081	110,485

Tabela B.6: Descritores estatísticos dos parâmetros de sinal de eletrocardiograma, com apresentação de estímulos, para a amostra do sexo feminino

Parâmetro	<i>Jitter simples (%)</i>	RAP (%)	PPQ5 (%)	PPQ11 (%)	BPM
Nº amostras	90	90	90	51	90
Média	4,218	0,254	0,311	0,037	83,697
Mediana	3,832	0,215	0,276	0,026	83,978
Desvio Padrão	2,099	0,179	0,192	0,038	9,320
Variância	4,409	0,032	0,037	0,001	86,863
Enviesamento	1,534	3,820	2,594	2,362	-0,455
Curtose	3,557	22,533	11,509	6,213	0,664
Mínimo	1,201	0,072	0,092	0,004	61,157
Máximo	13,547	1,459	1,417	0,193	107,491

Anexo C

Descritores estatísticos de parâmetros de sinais de voz (vogal /a/)

Neste anexo estarão presentes todos os descritores estatísticos de sinal de voz (vogal /a/) obtidos para os resultados deste estudo. Servirá para uma consulta mais pormenorizada, procurando esclarecer situações esporádicas relativas a grupos específicos deste estudo.

Estes descritores foram obtidos para as seguintes situações:

- Em situação de repouso: para a totalidade da amostra e para elementos dos sexos masculino e feminino;
- Com apresentação de estímulo: para a totalidade da amostra e para elementos dos sexos masculino e feminino;

Tabela C.1: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a totalidade de amostra (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	171	171	171	171	171	171
Média	145,482	0,563	0,289	0,310	4,867	2,453
Mediana	113,127	0,573	0,288	0,319	4,871	2,342
Desvio Padrão	42,972	0,150	0,083	0,088	1,312	0,742
Variância	1846,574	0,023	0,007	0,008	1,722	0,550
Enviesamento	0,427	-0,466	-0,193	-0,307	-0,013	0,133
Curtose	-1,416	-0,466	-0,628	-0,614	-0,776	-0,909
Mínimo	104,504	0,220	0,113	0,124	2,782	1,316
Máximo	228,335	0,824	0,430	0,442	7,245	3,721

Tabela C.2: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a totalidade de amostra (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	171	162	171	171	171
Média	2,990	3,960	0,917	0,147	16,573
Mediana	3,021	3,673	0,977	0,024	18,610
Desvio Padrão	0,799	1,040	0,145	0,283	6,333
Variância	0,638	1,081	0,021	0,080	40,107
Enviesamento	0,114	0,233	-2,016	1,997	-1,666
Curtose	-0,135	-0,482	2,344	2,228	1,620
Mínimo	1,708	2,197	0,540	0,008	0,712
Máximo	4,783	6,240	0,992	0,878	22,489

Tabela C.3: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo masculino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	81	81	81	81	81	81
Média	122,038	0,559	0,285	0,307	5,144	2,618
Mediana	108,101	0,522	0,281	0,319	5,686	2,647
Desvio Padrão	27,785	0,133	0,079	0,069	1,683	0,955
Variância	771,979	0,018	0,006	0,005	2,831	0,911
Enviesamento	1,737	-0,088	0,082	-0,685	-0,315	-0,255
Curtose	1,550	-1,647	-1,213	-0,545	-1,501	-1,611
Mínimo	104,839	0,362	0,174	0,168	2,782	1,316
Máximo	190,524	0,715	0,417	0,386	7,245	3,721

Tabela C.4: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo masculino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	81	72	81	81	81
Média	2,957	3,876	0,943	0,098	17,429
Mediana	3,175	3,574	0,9766	0,033	18,493
Desvio Padrão	0,989	1,201	0,089	0,181	4,120
Variância	0,638	1,081	0,021	0,080	40,107
Enviesamento	0,275	0,670	-2,384	2,414	-1,274
Curtose	-0,828	-0,465	4,033	4,124	1,170
Mínimo	1,708	2,322	0,697	0,010	7,596
Máximo	4,783	6,240	0,990	0,600	22,489

Tabela C.5: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo feminino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	90	90	90	90	90	90
Média	166,581	0,568	0,293	0,314	4,618	2,304
Mediana	182,834	0,608	0,288	0,318	4,710	2,338
Desvio Padrão	43,392	0,165	0,087	0,102	0,782	0,430
Variância	1882,864	0,027	0,008	0,010	0,612	0,185
Enviesamento	-0,461	-0,666	-0,400	-0,239	-0,849	-0,355
Curtose	-1,185	-0,088	-0,217	-0,959	0,654	0,260
Mínimo	104,504	0,220	0,113	0,124	2,815	1,392
Máximo	228,335	0,824	0,430	0,442	5,778	3,036

Tabela C.6: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), em situação de repouso, para a amostra do sexo feminino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	90	90	90	90	90
Média	3,020	4,028	0,894	0,191	15,802
Mediana	2,979	3,732	0,980	0,022	18,616
Desvio Padrão	0,581	0,892	0,178	0,346	7,751
Variância	0,338	0,795	0,032	0,119	60,079
Enviesamento	-0,343	-0,437	-1,521	1,523	-1,364
Curtose	0,619	-0,574	0,327	0,331	0,127
Mínimo	1,755	2,197	0,540	0,008	0,712
Máximo	4,058	5,022	0,993	0,878	22,093

Tabela C.7: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	177	176	176	174	174	174
Média	153,512	0,943	0,496	0,433	5,198	2,515
Mediana	152,083	0,617	0,318	0,323	4,708	2,258
Desvio Padrão	42,661	1,304	0,758	0,610	2,621	1,273
Variância	1819,945	1,700	0,574	0,372	6,869	1,620
Enviesamento	0,268	5,262	6,200	8,079	1,746	1,535
Curtose	-1,373	32,396	45,459	78,543	4,812	3,590
Mínimo	101,135	0,131	0,066	0,076	1,773	0,790
Máximo	245,020	11,076	7,258	6,917	17,203	8,171

Tabela C.8: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a totalidade da amostra (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	171	168	177	177	177
Média	3,163942	4,131679	0,936266	0,107435	17,653141
Mediana	2,862000	3,795500	0,977356	0,026354	18,715000
Desvio Padrão	1,6160293	2,0792488	0,1158767	0,2276207	6,0576330
Variância	2,612	4,323	0,013	0,052	36,695
Enviesamento	1,546	1,372	-2,796	2,906	-1,403
Curtose	3,401	2,267	6,640	7,200	1,960
Mínimo	1,0460	1,4150	0,5146	0,0021	0,2570
Máximo	10,0180	13,2320	0,9979	0,9566	28,3950

Tabela C.9: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo masculino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	88	87	87	85	85	85
Média	118,434	1,255	0,642	0,543	5,749	2,743
Mediana	112,986	0,666	0,311	0,328	4,800	2,261
Desvio Padrão	15,776	1,778	1,045	0,847	2,976	1,451
Variância	248,884	3,160	1,092	0,717	8,859	2,104
Enviesamento	1,336	3,748	4,447	5,899	1,900	1,620
Curtose	0,626	15,556	22,442	40,434	4,421	3,178
Mínimo	101,135	0,320	0,170	0,192	2,162	0,905
Máximo	157,394	11,076	7,258	6,917	17,203	8,171

Tabela C.10: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo masculino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	82	79	88	88	88
Média	3,343	4,300	0,938	0,102	17,325
Mediana	2,869	3,898	0,977	0,027	18,709
Desvio Padrão	1,653	1,919	0,095	0,184	5,316
Variância	2,731	3,683	0,009	0,034	28,257
Enviesamento	1,693	1,873	-2,781	3,032	-1,446
Curtose	3,560	5,549	7,771	9,310	1,695
Mínimo	1,100	1,600	0,535	0,008	0,719
Máximo	10,018	13,232	0,993	0,929	24,993

Tabela C.11: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo feminino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	89	89	89	89	89	89
Média	188,196	0,637	0,354	0,328	4,671	2,297
Mediana	191,198	0,577	0,324	0,315	4,443	2,257
Desvio Padrão	30,769	0,319	0,182	0,154	2,115	1,038
Variância	946,734	0,102	0,033	0,024	4,473	1,078
Enviesamento	-1,252	0,843	0,904	0,647	0,808	0,792
Curtose	2,344	0,932	1,361	0,440	0,589	0,548
Mínimo	106,702	0,131	0,066	0,076	1,773	0,790
Máximo	245,020	1,665	1,013	0,793	12,021	5,710

Tabela C.12: Descritores estatísticos dos parâmetros de sinal de voz (vogal /a/), com apresentação de estímulo, para a amostra do sexo feminino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	89	89	89	89	89
Média	2,999	3,982	0,935	0,113	17,977
Mediana	2,760	3,599	0,978	0,026	19,099
Desvio Padrão	1,573	2,212	0,134	0,265	6,726
Variância	2,474	4,892	0,018	0,070	45,240
Enviesamento	1,438	1,157	-2,670	2,699	-1,423
Curtose	3,429	0,653	5,361	5,489	1,986
Mínimo	1,046	1,415	0,515	0,002	0,257
Máximo	9,982	9,845	0,998	0,957	28,395

Anexo D

Descritores estatísticos de parâmetros de sinais de voz (vogal /i/)

Neste anexo estarão presentes todos os descritores estatísticos de sinal de voz (vogal /i/) obtidos para os resultados deste estudo. Servirá para uma consulta mais pormenorizada, procurando esclarecer situações esporádicas relativas a grupos específicos deste estudo.

Estes descritores foram obtidos para as seguintes situações:

- Em situação de repouso: para a totalidade da amostra e para elementos dos sexos masculino e feminino;
- Com apresentação de estímulo: para a totalidade da amostra e para elementos dos sexos masculino e feminino;

Tabela D.1: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a totalidade da amostra (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	171	171	171	153	171	153
Média	153,222	1,117	0,605	0,296	8,840	2,063
Mediana	139,486	0,523	0,284	0,277	3,739	1,698
Desvio Padrão	40,750	1,487	0,838	0,166	12,474	1,168
Variância	1660,557	2,211	0,703	0,028	155,603	1,364
Enviesamento	0,279	2,041	1,988	1,586	2,401	1,457
Curtose	2,344	0,932	1,361	0,440	0,589	0,548
Mínimo	102,330	0,242	0,097	0,128	2,351	1,031
Máximo	232,963	5,053	2,788	0,783	44,297	4,614

Tabela D.2: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a totalidade da amostra (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	144	144	171	171	171
Média	3,119	4,265	0,958	0,064	21,073
Mediana	2,347	2,738	0,992	0,009	23,559
Desvio Padrão	2,109	3,877	0,081	0,134	6,464
Variância	4,450	15,028	0,007	0,018	41,786
Enviesamento	2,245	2,982	-2,309	2,385	-1,354
Curtose	4,531	8,205	3,757	4,034	0,868
Mínimo	1,611	1,952	0,732	0,002	5,622
Máximo	9,981	18,280	0,998	0,444	28,511

Tabela D.3: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo masculino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	81	81	81	81	81	81
Média	122,038	0,559	0,285	0,307	5,144	2,618
Mediana	139,486	0,523	0,284	0,277	3,739	1,698
Desvio Padrão	27,785	0,133	0,079	0,069	1,683	0,955
Variância	771,979	0,018	0,006	0,005	2,831	0,911
Enviesamento	1,737	-0,088	0,082	-0,685	-0,315	-0,255
Curtose	1,550	-1,647	-1,213	-0,545	-1,501	-1,611
Mínimo	104,839	0,362	0,174	0,168	2,782	1,316
Máximo	190,524	0,715	0,417	0,386	7,245	3,721

Tabela D.4: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo masculino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	81	72	81	81	81
Média	2,957	3,876	0,943	0,098	17,423
Mediana	3,175	3,574	0,977	0,033	18,493
Desvio Padrão	0,989	1,201	0,089	0,181	4,120
Variância	0,979	1,442	0,008	0,033	16,976
Enviesamento	0,275	0,670	-2,384	2,414	-1,274
Curtose	-0,828	-0,465	4,033	4,124	1,170
Mínimo	1,708	2,322	0,697	0,010	7,596
Máximo	4,783	6,240	0,990	0,600	22,489

Tabela D.5: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo feminino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	81	81	81	63	81	63
Média	181,567	1,491	0,825	0,251	13,460	1,809
Mediana	186,585	0,523	0,284	0,283	3,685	1,476
Desvio Padrão	32,186	1,921	1,061	0,083	16,878	1,033
Variância	1035,955	3,689	1,124	0,007	284,863	1,068
Enviesamento	-0,537	1,342	1,330	-0,134	1,251	1,886
Curtose	-0,428	-0,172	-0,185	-1,466	-0,313	2,004
Mínimo	128,008	0,242	0,097	0,128	2,351	1,031
Máximo	232,963	5,053	2,788	0,366	44,297	4,261

Tabela D.6: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), em situação de repouso, para a amostra do sexo feminino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	63	63	81	81	81
Média	3,275	5,101	0,931	0,109	19,611
Mediana	2,349	2,737	0,992	0,009	23,559
Desvio Padrão	2,786	5,475	0,108	0,181	8,424
Variância	7,761	29,979	0,012	0,033	70,958
Enviesamento	2,010	2,010	-1,299	1,331	-0,814
Curtose	2,264	2,260	-0,238	-0,190	-0,994
Mínimo	1,611	1,952	0,732	0,003	5,622
Máximo	9,981	18,280	0,998	0,444	28,511

Tabela D.7: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a totalidade da amostra (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	174	169	164	161	162	160
Média	156,058	1,047	0,469	0,423	4,934	2,246
Mediana	167,760	0,578	0,289	0,280	4,188	1,928
Desvio Padrão	41,763	1,912	0,965	0,802	3,612	1,471
Variância	1744,173	3,655	0,932	0,644	13,047	2,165
Enviesamento	0,065	5,190	7,437	6,677	6,486	3,315
Curtose	-1,596	30,801	66,508	46,101	60,372	15,296
Mínimo	101,525	0,108	0,052	0,059	1,149	0,468
Máximo	227,375	15,289	10,193	6,788	40,756	10,803

Tabela D.8: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a totalidade da amostra (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	159	152	174	174	174
Média	3,067	3,867	0,971	0,040	22,162
Mediana	2,573	3,463	0,989	0,013	23,219
Desvio Padrão	2,294	1,909	0,060	0,092	5,943
Variância	5,261	3,643	0,004	0,008	35,328
Enviesamento	5,533	1,907	-3,851	4,746	-1,130
Curtose	42,857	4,659	16,595	28,039	2,692
Mínimo	0,819	1,197	0,570	0,0003	1,255
Máximo	23,398	12,163	0,999	0,776	36,261

Tabela D.9: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo masculino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	88	84	82	79	80	79
Média	120,855	1,182	0,568	0,486	5,520	2,661
Mediana	115,536	0,601	0,254	0,287	4,482	2,244
Desvio Padrão	19,229	2,072	1,284	0,967	4,705	1,868
Variância	369,789	4,292	1,650	0,935	22,141	3,488
Enviesamento	1,419	4,805	5,992	5,921	5,677	2,742
Curtose	1,163	27,278	40,815	35,322	40,380	8,873
Mínimo	101,525	0,275	0,108	0,154	2,130	0,965
Máximo	172,568	15,289	10,193	6,788	40,756	10,803

Tabela D.10: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo masculino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	78	72	88	88	88
Média	3,318	3,528	0,968	0,043	21,19
Mediana	2,739	3,333	0,987	0,015	22,735
Desvio Padrão	2,959	1,090	0,053	0,075	5,303
Variância	8,756	1,189	0,003	0,006	28,126
Enviesamento	5,027	0,614	-2,795	2,854	-1,734
Curtose	30,270	0,265	6,890	7,546	2,423
Mínimo	1,106	1,630	0,758	0,002	5,538
Máximo	23,398	6,989	0,998	0,361	27,669

Tabela D.11: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo feminino (Parte I)

Parâmetro	Pitch (Hz)	Jitter (%)	RAP (%)	PPQ5 (%)	Shimmer (%)	APQ3 (%)
Nº amostras	86	85	82	82	82	81
Média	192,081	0,914	0,369	0,362	4,364	1,841
Mediana	194,547	0,536	0,302	0,254	4,054	1,792
Desvio Padrão	23,989	1,741	0,454	0,603	1,916	0,752
Variância	575,492	3,032	0,206	0,364	3,670	0,566
Enviesamento	-1,849	5,808	5,306	7,478	0,870	0,483
Curtose	5,249	37,892	30,729	61,198	0,504	0,019
Mínimo	103,197	0,108	0,052	0,059	1,149	0,468
Máximo	227,375	13,712	3,258	5,383	9,713	3,749

Tabela D.12: Descritores estatísticos dos parâmetros de sinal de voz (vogal /i/), com apresentação de estímulos, para a amostra do sexo feminino (Parte II)

Parâmetro	APQ5 (%)	APQ11 (%)	Autocorrelação	NHR (dB)	HNR (dB)
Nº amostras	81	80	86	86	86
Média	2,825	4,171	0,974	0,036	23,157
Mediana	2,463	3,558	0,991	0,011	23,520
Desvio Padrão	1,357	2,387	0,066	0,107	6,414
Variância	1,842	5,696	0,004	0,011	41,134
Enviesamento	1,252	1,524	-4,461	5,232	-0,999
Curtose	1,251	2,090	21,209	30,368	2,932
Mínimo	0,819	1,197	0,570	0,0003	1,255
Máximo	7,031	12,163	0,999	0,776	36,261

Anexo E

Consentimento informado

Este anexo contém o consentimento informado que se apresentou a cada voluntário antes da recolha efetuada na Faculdade de Psicologia e Ciências da Educação. A sua elaboração foi feita com a colaboração de elementos do Laboratório de Neuropsicofisiologia da referida faculdade.



[EEC0020] – Dissertação

Correspondência entre alterações da voz e do eletrocardiograma
Estudo e contextualização do Polígrafo quanto a Análise de Voz, ECG e EDR

Orientador: Prof. Anibal Ferreira

Consentimento Informado

A investigação para a qual contamos com a sua colaboração desenrola-se no âmbito da Unidade Curricular Dissertação do Mestrado Integrado em Engenharia Electrotécnica e de Computadores da Faculdade de Engenharia da Universidade do Porto. As dissertações para as quais os testes a serem feitos irão contribuir pretendem estudar variações em sinais de voz, eletrocardiografia e eletrodermografia em situação de repouso e como reação a um conjunto de estímulos validados pela comunidade científica.

Caso aceite colaborar nesta investigação, o que muito agradecemos, a sua tarefa consistirá simplesmente em observar um conjunto de imagens de carácter agradável, desagradável e neutro. Em alguns dos casos pode ser necessário atribuir-lhes uma classificação, devidamente justificada. Nesta tarefa não existem respostas certas nem erradas.

Durante esta tarefa de observação de imagens, a experiência envolve a realização simultânea de uma eletrocardiografia (colocação de um elétrodo no pulso e um em cada tornozelo, previamente humedecidos com gel), uma eletrodermografia (colocação de dois elétrodos na mão esquerda) e captação de sinal de voz com um microfone de *headset*. Será também feita uma filmagem para futuro despiste de distorções inesperadas em algum dos sinais.

Em relação ao exame de eletrocardiografia, as seguintes contra-indicações são eliminatórias: embolia pulmonar, enfermidade aguda, febril ou grave, limitação física ou psicológica, intoxicação medicamentosa.

O equipamento utilizado, bem como os procedimentos adotados, obedecem a todos os padrões de higiene e segurança e não causam qualquer desconforto aos participantes. No entanto, poderá desistir da experiência, se e quando entender. Todos os dados serão registados, tratados e arquivados de forma anónima e confidencial. Não serão analisados dados individuais e apenas se poderá proceder à divulgação científica de resultados de grupo, sendo que as eventuais publicações que decorrerem do estudo poderão ser oferecidas aos participantes que as solicitarem.

A decisão de participar nesta investigação é totalmente livre e voluntária, não resultando dessa decisão, qualquer que ela seja, nenhum prejuízo ou benefício pessoal. Caso tenha alguma questão ou pretenda esclarecimentos complementares, poderá obtê-los junto de Ricardo Dias (ee07135@fe.up.pt) ou Mário Barbosa (ee06157@fe.up.pt).

Aproveitamos, novamente, para agradecer a sua colaboração, que muito contribuirá para o sucesso deste estudo.

Eu, _____ (nome completo em letra legível), declaro que li integralmente o texto acima e que me foi dada a oportunidade de esclarecer todas as dúvidas que considere importantes para tomar uma decisão sobre a minha colaboração na investigação em apreço, pelo que me considero esclarecido (a) sobre a mesma e aceito nela participar de forma voluntária.

Porto, ___ de _____ de 2012

 (Assinatura legível)

Figura E.1: Consentimento informado apresentado a cada voluntário antes do início do teste

Referências

- [1] Douglas P. Zipes e Peter Libby Eugene Braunwald, Robert O. Bonow. *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine*. Elsevier Inc, 8ª edição, 2008.
- [2] Andrew Houghton e David Gray. *Making Sense of the ECG – Cases for self assessment*. Hodder Arnold, 2009.
- [3] Comparison of four different mobile devices for measuring heart rate and ecg with respect to aspects of usability and acceptance by older people. *Applied Ergonomics*, 43(3):582 – 587, 2012.
- [4] Oscar Castillo et al. Hybrid intelligent system for cardiac arrhythmia classification with fuzzy k-nearest neighbors and neural networks combined with a fuzzy system. *Expert Systems with Applications – An International Journal*, 2011.
- [5] Fernando Pereira et al. *Comunicações Audiovisuais – Tecnologias, Normas e Aplicações*. IST Press, 1ª edição, 2009.
- [6] the online encyclopedia Britannica. tone (sound) – britannica online encyclopedia. Disponível em <http://www.britannica.com/EBchecked/topic/599085/tone> [Último acesso a Fev.2012].
- [7] the free encyclopedia Wikipedia. Spectrogram – wikipedia, the free encyclopedia. Disponível em <http://en.wikipedia.org/wiki/Spectrogram> [Último acesso a Fev.2012].
- [8] Jorge Jescheniak et al. Word frequency effects in speech production: Retrieval of syntactic information and of phonological form. *Journal of Experimental Psychology: Learning, Memory and Cognition*, páginas 103–114, 1994.
- [9] Ricardo Sousa. *Metodologias de avaliação perceptiva e acústica do sinal de voz em aplicações de ensino do canto e diagnóstico/reabilitação da fala (Tese de Doutoramento)*. Faculdade de Engenharia da Universidade do Porto, 2011.
- [10] Aníbal Ferreira. *Técnicas não invasivas de avaliação vocal (pós-graduação em voz profissional)*. 2009.
- [11] Voiceproblem.org: Anatomy and physiology of voice production. Disponível em <http://www.voiceproblem.org/anatomy/learning.php> [Último acesso a Jun.2012].
- [12] Keith L. Moore et al. *Clinically Oriented Anatomy*. Lippincott Williams Wilkins, 5th edição, 2006.

- [13] Sympathetic nervous system. Disponível em http://www.sciencedaily.com/articles/s/sympathetic_nervous_system.htm [Último acesso a Jun.2012].
- [14] Parasympathetic nervous system. Disponível em http://www.sciencedaily.com/articles/p/parasympathetic_nervous_system.htm [Último acesso a Jun.2012].
- [15] Autonomic nervous system. Disponível em <http://genericlook.com/img/uploads/anatomy/autonomic-nervous-system.gif> [Último acesso a Jun.2012].
- [16] Dmitriyi Skopin. Heartbeat feature extraction from vowel speech signal using 2d spectrum representation.
- [17] Ying-Chien Wei et al. A portable ecg signal monitor and analyser. 2008.
- [18] Nivedita Deshpande et al. Assessment of systolic and diastolic cycle duration from speech analysis in the state of anger and fear. 2012.
- [19] Ana Leonor Pereira. As cores da voz - expressão das emoções no timbre da voz cantada. Tese de mestrado, Universidade Católica Portuguesa, 2007.
- [20] José Moreira et al. Assessment of systolic and diastolic cycle duration from speech analysis in the state of anger and fear. 2009.
- [21] Characterization of healthy and pathological voice through measures based on nonlinear dynamics. *Trans. Audio, Speech and Lang. Proc.*, 17(6):1186–1195, Agosto 2009.
- [22] Update do mp36 system. Disponível em <http://www.biopac.com/upgrade-mp36-system-mac> [Último acesso a Mai.2012].
- [23] BIOPAC Systems Inc. *Physiology Lessons for use with the Biopac Student Lab*. BIOPAC Systems Inc., 2007.
- [24] Adobe. Audio editing, mastering, mixing software | adobe audition. Disponível em <http://www.adobe.com/products/audition/features.html> [Último acesso a Mai.2012].
- [25] Sennheiser. Sennheiser ear set 1 - microphone. Disponível em <http://en-de.sennheiser.com/ear-set-1> [Último acesso a Mai.2012].
- [26] Roland. Ua 25ex :: Products :: Roland. Disponível em <http://www.roland.com/products/en/UA-25EX/> [Último acesso a Mai.2012].
- [27] P. J. Lang et al. International affective picture system (iaps): Technical manual and affective ratings. *Advances in Engineering Software*, 2005.
- [28] Voicebox: Speech processing toolbox for matlab. Disponível em <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html/> [Último acesso a Mar.2012].
- [29] Physionet - the research resource for complex physiologic signals. Disponível em <http://www.physionet.org/> [Último acesso a Mar.2012].
- [30] Praat: doing phonetics by computer. Disponível em <http://www.fon.hum.uva.nl/praat/> [Último acesso a Mar.2012].

- [31] Seegnal > voice studio. Disponível em <http://www.seegnal.pt/products/voicestudio.php> [Último acesso a Mar.2012].
- [32] Cornelia Kappeler-Setz et al. Towards long term monitoring of electrodermal activity in daily life. *Personal and Ubiquitous Computing*, páginas 1–11.
- [33] Rui Campos Guimarães et al. *Estatística*. McGraw-Hill, 1997.
- [34] Richard J. Harris. *A primer of multivariate statistics*. Lawrence Erlbaum Associates, 3ª edição, 2001.
- [35] William Rodrigues. *Estatística Aplicada*. W. C. Rodrigues, 6ª edição, 2004-08.
- [36] Sennheiser. *Instruction Manual EAR SET 1*. Sennheiser, 2009.