

BOLSA DE INVESTIGAÇÃO

Relatório de atividades do período – 1 de Outubro de 2018 a 31 de Março de 2019

Bolseira: Sara Isabel Pereira Castilho

Projeto: Dysphonic to Natural Voice Reconstruction (DyNaVoiceR) -R&D Project AAC nº 02/SAICT/2017

Palavras-chave: DynaVoiceR; Fala sussurrada; Base de dados; Voz.

Local de trabalho: Instituto de Engenharia Eletrónica e Informática de Aveiro e Escola Superior de Saúde da Universidade de Aveiro

No âmbito do projeto "Dysphonic to Natural Voice Reconstruction (DyNaVoiceR)" -R&D Project AAC nº 02/SAICT/2017, a bolseira desenvolveu o seu trabalho na construção de uma base de dados de amostras de vozes normais em Português Europeu (PE) em fala normal e fala sussurrada. Para atingir este objetivo desenvolveram-se as seguintes tarefas:

1 - Elaboração do pedido de autorização e parecer a uma Comissão de Ética independente, relativamente à criação da base de dados de amostras de vozes normais em PE em fala normal e fala sussurrada, com os seguintes documentos:

- Formulário para submissão de pedidos de apreciação à Comissão de Ética relativos a estudos de investigação;
- Apresentação do estudo;
- Consentimento informado;
- Guião para recolha de amostras de fala;
- Protocolo de Anamnese Vocal (Mini) da Universidade de Aveiro;
- Curriculum Vitae dos proponentes.

2 - Definição do corpus de gravação para a base de dados.

Cada gravação de cerca de 25 minutos por informante inclui 13 sons sustentados (correspondentes às sibilantes e vogais orais do PE), 28 palavras (com diferentes contextos vocálicos e consonânticos, apesar da prevalência por consoantes fricativas e oclusivas), 6 frases (CAPE_V (1)) e 1 texto

foneticamente balanceado -"O Vento Norte e o Sol"(2) (Apêndice 1). As mesmas tarefas foram sempre gravadas em fala normal e em fala sussurrada.

O *corpus* de gravação da base de dados foi sujeito a uma gravação-teste com respetiva segmentação e anotação para assegurar viabilidade e pertinência das tarefas de fala propostas de acordo com os objetivos do projeto. Após esta gravação-teste e reunião com toda a equipa do projeto, fizeram-se as adaptações necessárias para a proposta definitiva de gravação.

3 -Gravação

Para a gravação da base de dados, a bolsreira tratou da logística das gravações (apresentação do estudo, tarefas de fala a gravar e consentimento informado). Fez também o recrutamento de informantes voluntários e respetivo agendamento.

O recrutamento de voluntários para colaboração no projeto dentro e fora da Universidade de Aveiro foi feito de acordo com os seguintes critérios de inclusão:

- Idade compreendida entre os 18 e os 35 ou entre os 46 e 65, procurando informantes de duas faixas etárias diferentes (3).
- Ausência de historial de patologia vocal e com qualidade vocal sem alterações, segundo avaliação por um Terapeuta da Fala;
- Ausência de infeção ou inflamação das vias aéreas superiores no dia da gravação;
- Língua Portuguesa como língua materna e serem naturais de localidades em que se fala o dialeto do centro-litoral;

Foram excluídos os informantes que, apesar de reunirem as condições acima mencionadas, apresentaram uma das seguintes condições:

- Em processo de correção ortodôntica;
- Patologia respiratória ou refluxo gastro-esofágico com consequências na qualidade vocal;
- Alterações de fluência;
- Realizaram algum tratamento vocal anteriormente;
- Não conseguira executar as tarefas vocais propostas.

As gravações decorreram numa sala ABS-AUD.45.1 produzida pela Absorsor Portugal, com redução de 45 dB, na Escola Superior de Saúde da

Universidade de Aveiro. Os informantes estavam sentados e foram gravados com um microfone de cabeça Sennheiser Ear Set 1 com uma frequência de amostragem de 48000 Hz, 16 bits por amostra. A fala sussurrada foi gravada com um volume 2.25 vezes superior à fala normal.

Em cada gravação esteve sempre presente a bolsreira, Terapeuta da Fala, responsável por explicar o estudo e respetivos procedimentos, orientar os voluntários nas tarefas a manter uma boa qualidade vocal durante a gravação, o orientador, responsável pela gestão dos equipamentos durante a gravação e o informante.

No total, foram gravados 35 informantes, dos quais 5 foram excluídos de acordo com os critérios acima referidos. A base de dados é composta por 30 informantes, 15 do sexo feminino e 15 do sexo masculino, de duas diferentes faixas etárias (Apêndice 2). A gravação em bruto de cada informante é composta por quatro ficheiros áudio, um para cada tarefa.

4 -Segmentação e anotação dos ficheiros de áudio

À gravação de áudio dos informantes, sucedeu-se a segmentação dos ficheiros em bruto, dividindo-os nos diferentes sons, palavras, frases e texto, em fala normal. O procedimento repetiu-se também para a fala sussurrada, totalizando-se 92 ficheiros de áudio para cada informante (48 em fala normal e 48 em fala sussurrada). Cada ficheiro segmentado foi nomeado com um código específico para indicar o sexo do informante, a tarefa e o modo (normal ou sussurrado) (Apêndice 3). Esta segmentação foi feita com recurso ao Audacity 2.3.0.

A anotação dos ficheiros que compõem a base de dados é um passo essencial para o desenvolvimento do projeto e esta foi feita através do Praat 6.0.43. Para esta base de dados foram anotados todos os fonemas e todas as palavras. Nas frases apenas foi anotada uma repetição de cada frase e no texto foram anotados os segmentos /s, z, S, Z, i, a, O, U/. A anotação dos informantes gerou 92 ficheiros por cada um deles.

Foram definidos critério de anotação dos ficheiros pela bolsreira, em conjunto com o orientador, tendo como base a literatura existente(4,5), para uma maior consistência e homogeneidade nas anotações (Apêndice 4).

Dada a diversidade de características vocais encontrada, mesmo dentro da “normalidade” e devido a algumas interferências não controláveis na gravação, surgiu a necessidade de criar um ficheiro onde foi referenciado pela bolsreira os ficheiros que apresentavam *clipping* e/ou distorções. Esse ficheiro foi partilhado com os restantes membros envolvidos no projeto.

Uma vez que o Projeto DynaVoiceR está a ser desenvolvido em conjunto com a Faculdade de Engenharia do Porto, a bolsreira preparou e participou em reuniões que decorreram no Porto, tendo sido estas importantes para a conciliação de práticas na criação da base de dados e definição de estratégias para partilha de conhecimentos e do material em desenvolvimento, bem como suporte mútuo à distância.

Bibliografia

1. Jesus L, Barney A, Couto P. Voice quality evaluation using CAPE-V and GRBAS in european Portuguese. Proc 6th 2009;3-6.
2. Jesus LMT, Valente ARS, Hall A. Is the Portuguese version of the passage “The North Wind and the Sun” phonetically balanced? J Int Phon Assoc. 2015;45(1):1-11.
3. Provisional Guidelines on Standard International Age Classifications. 1982;(74).
4. Jovičić ST, Šarić Z. Acoustic Analysis of Consonants in Whispered Speech. J Voice. 2008;22(3):263-74.
5. Jesus LMT, Shadle CH. A parametric study of the spectral characteristics of European Portuguese fricatives. 2002;437-64.

Apêndice 1 –Tarefas de fala do *corpus* da base de dados

São descritas neste apêndice, as tarefas tal como foram apresentadas aos participantes durante as sessões de gravação.

Tarefa 1: Fonemas sustentados – dizer cada som 3 vezes em fala normal e 3 vezes em fala sussurrada, 2 a 3 segundos

S de sapato X de chapéu Z de casa J de janela
I de ilha E de peso E de ela A de água
A de amarelo O de óculos O de ovo U de uva
E de sede

Tarefa 2: Leitura de palavras –dizer cada palavra 3 vezes em fala normal e 3 vezes em fala sussurrada, pela ordem apresentada;

Col. 1	Col. 2	Col. 3	Col. 4	Col. 5	Col. 6
Nuca	Ripa	Sala	Zaro	Asa	Velho
Lupa	Laje	Juba	Assa	Rita	Hora
Chiba	Face	Chama	Luta	Haja	Buda
Minha	Pica	Vida	Jarra	Ache	Viga
Guga	Vaze	Acha	Fisga		

Tarefa 3: Leitura de frases – ler cada frase 3 vezes em fala normal e 3 vezes em fala sussurrada

A Marta e o avô vivem naquele casarão rosa velho.

Sofia saiu cedo da sala.

A asa do avião andava avariada.

Agora é hora de acabar.

A minha mãe mandou-me embora.

O Tiago comeu quatro peras.

Tarefa 4: Leitura do texto – ler 1 vez em fala normal e 1 vez em fala sussurrada

O Vento Norte e o Sol (Cruz-Ferreira 1999: 130)

O vento norte e o sol discutiam qual dos dois era o mais forte, quando sucedeu passar um viajante envolto numa capa. Ao vê-lo, põem-se de acordo em como aquele que primeiro conseguisse obrigar o viajante a tirar a capa seria considerado o mais forte. O vento norte começou a soprar com muita fúria, mas quanto mais soprava, mais o viajante se aconchegava à sua capa, até que o vento norte desistiu. O sol brilhou então com todo o esplendor, e imediatamente o viajante tirou a capa. O vento norte teve assim de reconhecer a superioridade do sol.

Apêndice 2 – Tabelas dos informantes da base de dados

Informantes femininas

Código	Gravação	Segmentação	Anotação	Observações
SPF01	Sim	Sim	Sim	
SPF02	Sim	Sim	Sim	
SPF03	Sim	Sim	Sim	
SPF04	Sim	Sim	Sim	
SPF05	Sim	Sim	Sim	
SPF06	Sim	Sim	Sim	
SPF07	Sim	Sim	Sim	
SPF10	Sim	Sim	Sim	
SPF11	Sim	Sim	Sim	<i>Whistle</i> muito frequente
SPF12	Sim	Sim	Sim	Dificuldade em manter f0 nos sons sustentados, com tendência a descer 30 hz entre o início e fim do som sustentado; Ligeira rouquidão nos sons sustentados; Voz com algumas características de patologia;
SPF13	Sim	Sim	Sim	
SPF14	Sim	Sim	Sim	<i>Creaky voice</i> muito frequente
SPF16	Sim	Sim	Sim	
SPF17	Sim	Sim	Sim	
SPF09	Sim	Sim	Sim	Ligeiro sigmatismo, apesar de não ser perceptível sempre Défice auditivo no OD
SPF15	Sim	Sim	Não	Dificuldade em manter f0 nos sons sustentados, com tendência a subir
SPF08	Não	Não	Não	Excluir da base de dados – voz patológica

Informantes masculinos

Código	Gravação	Segmentação	Anotação	Observações
SPM01	Sim	Sim	Sim	
SPM02	Sim	Sim	Sim	
SPM03	Sim	Sim	Sim	
SPM04	Sim	Sim	Sim	
SPM05	Sim	Sim	Sim	
SPM06	Sim	Sim	Sim	
SPM07	Sim	Sim	Sim	
SPM08	Sim	Sim	Sim	Dificuldade em manter os sustentados, com tendência a subir f0 cerca de 20 Hz; <i>Creaky voice</i>
SPM09	Sim	Sim	Sim	
SPM10	Sim	Não	Não	Voz áspera
SPM11	Sim	Não	Não	Sotaque; troca o b pelo v sp;
SPM12	Sim	Sim	Sim	
SPM13	Sim	Sim	Sim	Algumas características de voz patológica; Não diz o l – (_P) –aproximante labiodental; Sons sustentados com instabilidade SPM13_48_01 com muitos momentos de <i>clipping</i>
SPM14	Sim	Sim	Sim	SPM14_09_02 –inclui ataque glotal SPM14_10_01 efeito de aproximação
SPM15	Sim	Sim	Não	<i>Clipping</i> em todas as vogais sustentadas
SPM16	Sim	Sim	Sim	
SPM17	Sim	Sim	Sim	
SPM18	Sim	Sim	Sim	Instabilidade na f0 nos sons sustentados

Legenda:

	Linhas em branco são pessoas da faixa etária mais baixa (18-35A)
	Linhas em cinza são pessoas da faixa etária mais alta (46-65A)
	Informantes excluídos da base de dados

Apêndice 3 – Formato de nomeação dos ficheiros de áudio

Nome –tipo: SPF00_00_00

- Letras iniciais: SPF- informante feminino ou SPM- informante masculino;
- 00 -nº atribuído ao participante, de acordo com ordem de gravação;
- 00 -tarefa

SUSTENTADOS

vogais orais -01 -i; 02 -e; 03-E; 04-a; 05-6; 06-O; 07-o; 08-u; 09-1;
sibilantes- 10-s; 11-z; 12-S; 13-Z;

PALAVRAS

14-Nuca	19-Ripa	24-Sala	29-Zaro	34-Asa	38-Velho
15-Lupa	20-Laje	25-Juba	30-Assa	35-Rita	39-Hora
16-Chiba	21-Face	26-Chama	31-Luta	36-Haja	40-Buda
17-Minha	22-Pica	27-Vida	32-Jarra	37-Ache	41-Viga
18-Guga	23-Vaze	28-Acha	33-Fisga		

FRASES

42-A Marta e o avô vivem naquele casarão rosa velho.

43-Sofia saiu cedo da sala.

44-A asa do avião andava avariada.

45-Agora é hora de acabar.

46-A minha mãe mandou-me embora.

47-O Tiago comeu quatro peras.

48-TEXTO

O Vento Norte e o Sol (Cruz-Ferreira 1999: 130)

O vento norte e o sol discutiam qual dos dois era o mais forte, quando sucedeu passar um viajante envolto numa capa.

Ao vê-lo, põem-se de acordo em como aquele que primeiro conseguisse obrigar o viajante a tirar a capa seria considerado

o mais forte. O vento norte começou a soprar com muita fúria, mas quanto mais soprava, mais o viajante se aconchegava

à sua capa, até que o vento norte desistiu. O sol brilhou então com todo o esplendor, e imediatamente o viajante tirou

a capa. O vento norte teve assim de reconhecer a superioridade do sol.

- 00 -fala normal ou sussurrada (01 -vozeado/ 02 -sussurrado)

Exemplo:

SPF02_15_02 corresponde ao ficheiro de áudio da informante feminina nº2, da palavra lupa, em fala sussurrada;

SPM14_45_01 corresponde ao ficheiro de áudio do informante masculino nº14, da frase “Agora é hora de acabar”, em fala normal.

Apêndice 4 –Critérios de anotação

Para a anotação, manteve-se o Praat com as definições standard. O espectro esteve maioritariamente a 5000 Hz e transitoriamente a 16 KHz apenas para definir fronteiras das fricativas;

Âncoras para o anotação:

- Forma do espectro das fricativas (ruído);
- *Burst* das oclusivas;
- Forma da espectro das formantes nas vogais

VOGAIS –

- Fala normal: ter em consideração o traçado de F2 e F3 em conjunto com forma de onda e periodicidade;
- Fala sussurrada: mesmos critérios, contudo ataques vocais bruscos no sussurrado são excluídos na anotação
- Vogais curtas ou não vozeadas anotadas como reduzidas/não vozeadas (_0);
- Perceção auditiva.

SIBILANTES

- Anotadas de acordo com o espectrograma, mesmo que f0 já não seja detetado;
- Fricativas desvozeadas ou com menos de 1/3 de vozeamento anotadas como não vozeadas com (_0);
- Perceção auditiva.

OUTRAS ANOTAÇÕES:

- (_o) fonemas fricativizados;
- (l) sempre sem velarização;
- (_k) creaky voice (registo basal)
- (_^) *whistled voice* (assobio)
- (_v) *voiced* (quando em fala sussurrada)
- (_P) aproximante labiodental