



Analise acústica de vozes com e sem edema de Reinke

Acoustic analysis of voices with and without Reinke's edema

Alexandre Carelli Borsoi*, Marcelo de Oliveira Rosa†,

RESUMO

Esse estudo tem como objetivo ajudar a identificar as anomalias na fala, em específico o edema de Reinke, através de algoritmos de processamento digital de sinais de fala, com o intuito de auxiliar fonoaudiólogos a identificar as anomalias na voz. Para isso foi utilizado um banco de dados contendo arquivos de fala de 58 mulheres na faixa etária de 45 a 78 anos, em que metade delas possuem o edema de Reinke (grupo de pesquisa) enquanto a outra metade não possui anomalias na voz (grupo de controle), e para processamento desses sinais de fala foi utilizado o software Praat, usado para a análise e síntese da fala. Também foi utilizado o AVQI, que é um índice que ajuda a classificar a qualidade dos áudios. Para analisar os resultados utilizou-se de testes estáticos como o teste t de Student e diagramas de bloco (*box plots*). Os resultados se mostraram promissores na identificação do edema, através dos testes t e da avaliação visual dos diagramas de bloco, em que somente a medida acústica de proeminência do pico cepstral suavizado (CPPs) não se mostrou suficiente para uma separação eficaz do grupo de controle do grupo de pesquisa.

Palavras-chave: Índice de Qualidade da Voz Acústica, edema de Reinke, medidas acústicas, processamento digital de sinais da fala.

ABSTRACT

This study aims to help identifying speech anomalies, specifically Reinke's edema, through digital speech signal processing algorithms, with the aim of helping speech therapists to identify voice anomalies. For this, a database containing speech files of 58 women aged from 45 to 78 years was used, where half of them have Reinke's edema (study group) while the other half have no voice anomalies (control group), and to process these speech signals the Praat software was used, a computer program for speech analysis and synthesis. The AVQI was also used, which is an index that helps to classify the quality of audios. To confirm the results, statistic tests such as the Student t test and box plots were used. The results were promising in the identification of edema, through the t tests and the visual evaluation of the block diagrams, where only the acoustic measures of smoothed cepstral peak prominence (CPPs) was not sufficient for an effective separation of the control group from the study group.

Keywords: Acoustic Voice Quality Index, Reinke's edema, acoustic measures, digital processing of speech signal.

1 INTRODUÇÃO

O edema de Reinke consiste em uma lesão da prega vocal que pode ocorrer devido ao uso intenso da voz e, principalmente, pelo tabagismo. Como consequência, indivíduos com esse edema tem como sintomas uma voz mais grave, e rouquidão.

O processamento digital de sinais de fala para obter medidas acústicas é utilizado por pesquisadores como forma de identificar e classificar disfonias na fala [1], em que os pesquisadores utilizaram 6 medidas acústicas que se mostraram mais proeminentes a indicar anomalias na voz, para desenvolver um índice chamado de "Acoustic Voice Quality Index" (AVQI), que em português seria Índice de Qualidade de Voz Acústica.

* Engenharia de Controle e Automação, Universidade Tecnológica Federal do Paraná, Curitiba, Brasil; aleborsa8@gmail.com

† Universidade Tecnológica Federal do Paraná, Campus Curitiba mrosa@utfpr.edu.br



O processamento e obtenção de medidas acústicas através de sinais de fala a fim de classificar e distinguir disfonias na voz pode também ser observado em [2], que analisa as medidas acústicas com uma máquina de vetores suporte, e com uma rede neural afim de encontrar padrões nas medidas acústicas provenientes dos distúrbios na fala.

Esse trabalho tem intuito de auxiliar fonoaudiólogos a identificar os distúrbios provocados na fala por pelo edema de Reinke, através do processamento digital e análise de sinais de fala.

2 MÉTODO

Para esse experimento foi utilizado o software Praat, e uma base de dados contendo 116 arquivos de áudio, em que 58 são arquivos de uma contagem de 1 até 10, e os outros 58 são arquivos de áudio da vogal sustentada, tanto a vogal sustentada quanto a contagem pertencem a mulheres na faixa etária de 45 até 78 anos, em que 58 deles, sendo divididos igualmente entre contagem e vogal sustentada, são de mulheres que contém o edema de Reinke, esse grupo é chamado de grupo de pesquisa, enquanto a outra metade não possui o edema, sendo chamada de grupo de controle.

Criou-se de uma rotina no Praat, que tinha como entrada os arquivos de áudio da vogal sustentada e da contagem, e como saída o valor do AVQI segundo a formula (1), bem como o valor das 6 medidas acústicas utilizadas para calcular o AVQI. A rotina foi aplicada em todas as falas do banco de dados e os valores de AVQI e das medidas acústicas de cada um dos sujeitos foram armazenados em uma tabela.

As medidas se relacionam da seguinte forma para o cálculo do AVQI [1]:

$$AVQI = ((3.295 - (0.111 * CPPs) - (0.073 * HNR) - (0.213 * shim) + (2.789 * shdb) - (0.032 * Slope) + (0.077 * Tilt)) * 2.208) + 1.797 \quad (1)$$

Sendo:

CPPs: proeminência do pico cepstral suavizado;

HNR: razão harmônico-ruído;

shim: Shimmer local;

shdb: Shimmer local em dB;

Slope: inclinação do espectro médio de longo prazo;

Tilt: inclinação da linha de tendência através do espectro médio de longo prazo.

As medidas acústicas consistem em: CPPs, que é uma medida acústica definida em [3], que tem por objetivo mostrar quão alto é o pico cepstral em relação ao fundo do cepstrum; Shimmer local (%), que estima a variação período-a-período da amplitude pico-a-pico na amostra de voz analisada; Shimmer em dB, da mesma forma que o Shimmer local, porém em dB; HNR, que consiste em uma avaliação da presença de ruído no sinal; Slope, que consiste na inclinação do espectro médio de longo prazo; e Tilt, que consiste na inclinação da linha de tendência através do espectro médio de longo prazo.

As medidas Shimmer em dB e Shimmer local são definidas por essas equações [2]:

$$Shimmer (dB) = \frac{1}{N-1} * \sum_{i=1}^{N-1} 20 \log \left| \left(\frac{A_{i+1}}{A_i} \right) \right| \quad (2)$$



$$Shimmer Local (\%) = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |A_i - A_{i+1}|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (3)$$

Sendo:

N: número de períodos extraídos da frequência fundamental;

A_i : amplitude pico a pico.

Com os valores das medidas acústicas e do AVQI foi possível realizar o teste estatístico t de Student bicaudal para duas amostras presumindo variâncias diferentes. O objetivo do teste estatístico é verificar se os resultados das medidas do grupo de controle e pesquisa são diferenciáveis entre si. Além do teste estático os valores das medições também foram comparados graficamente através de diagramas de caixa (*box plots*), que utiliza valores mínimos e máximos, primeiro e terceiro quartil, mediana de uma base de dados como referência.

3 RESULTADOS

Com os valores das medidas acústicas foi possível realizar o teste t entre as medidas do grupo de controle e grupo de pesquisa, os resultados para um nível de significância (α) de 0,05 são mostrados na quadro 3, no qual o termo “diferente” significa a hipótese alternativa (H_1) deve ser aceita, enquanto que o termo “não diferente” significa que a hipótese nula (H_0) não pode ser rejeitada.

Quadro 1 – Resultados do teste t.

AVQI	CPPS	H/N (dB)	Shimmer Local (%)	Shimmer Local (dB)	Slope LTAS (dB)	Tilt of TrendLine LTAS (dB)
Diferente	Não Diferente	Diferente	Diferente	Diferente	Diferente	Diferente

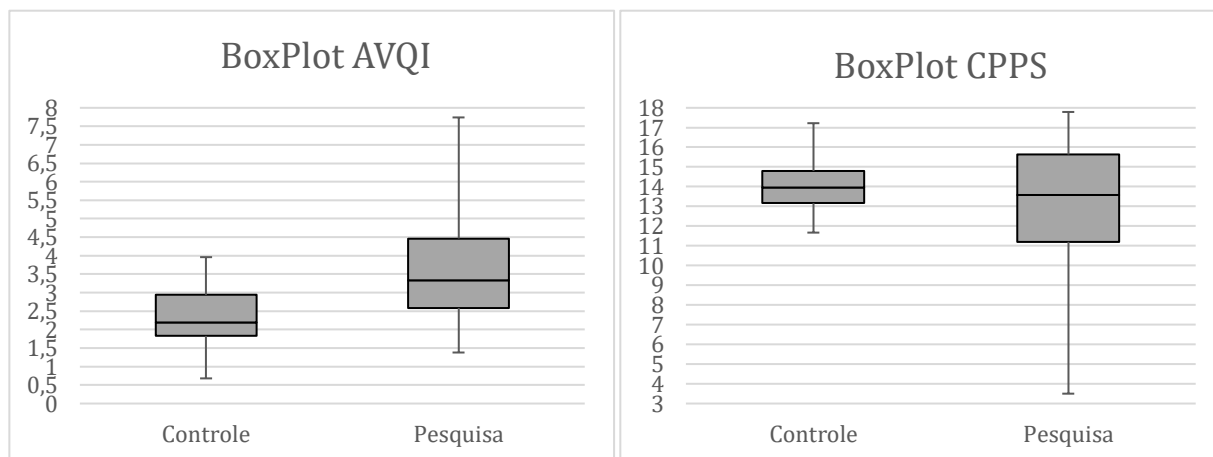
Fonte: Os Autores

Resultados denominados como “diferente” indicam que as medidas acústicas do grupo de controle e pesquisa possuem alta probabilidade de não terem ocorrido ao acaso, isso significa que há uma diferença no valor das medidas acústicas entre o grupo de controle e o de pesquisa. Já os resultados denominados como “não diferente” implicam que a probabilidade de os resultados não terem acontecido ao acaso seja menor, ou seja a um erro maior envolvido, sendo mais difícil afirmar que a uma diferença entre a medida acústica do grupo de controle e o grupo de pesquisa.

Além do teste t também foram feitos diagramas de caixa (*box plot*) como forma visual de observar as diferenças dos dois grupos.

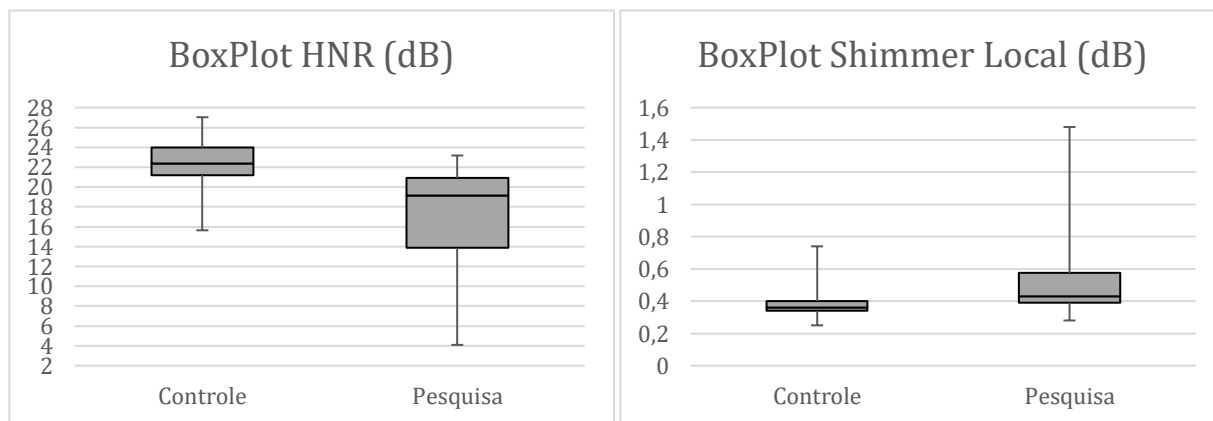


Figura 1 – Box plot AVQI e CPPs.



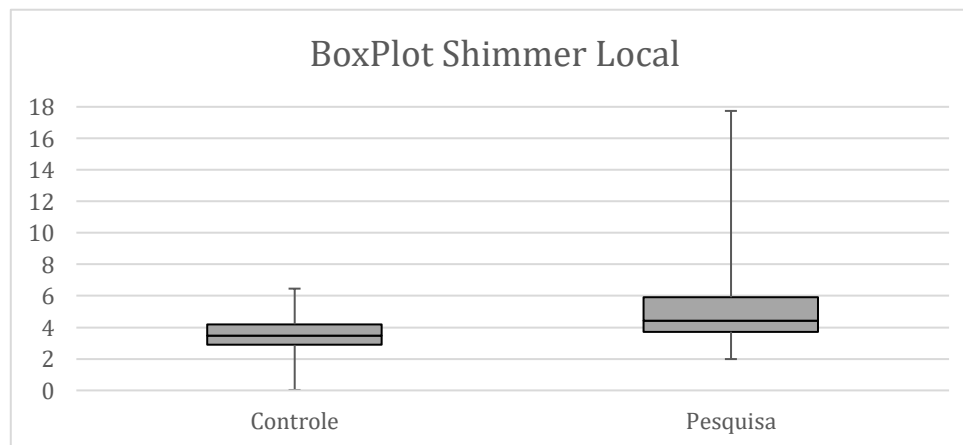
Fonte: Os Autores

Figura 2 – Box plot HNR e Shimmer Local (dB).



Fonte: Os Autores

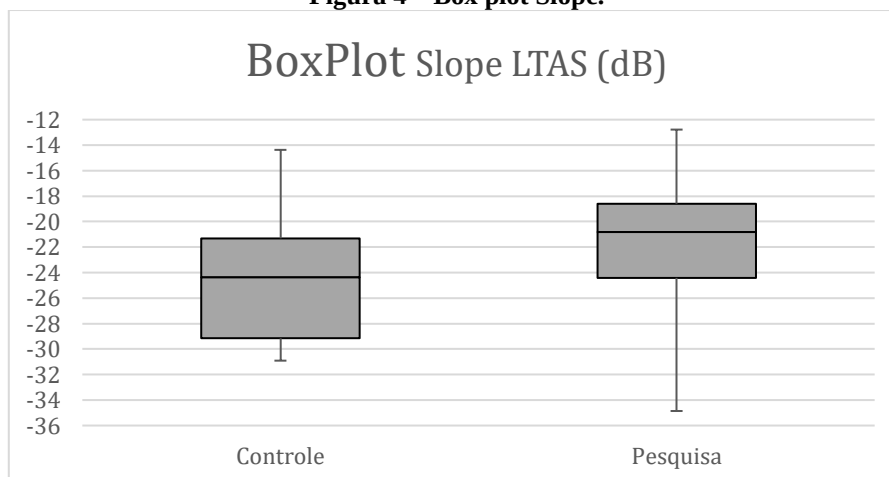
Figura 3 – Box plot Shimmer Local (%).



Fonte: Os Autores

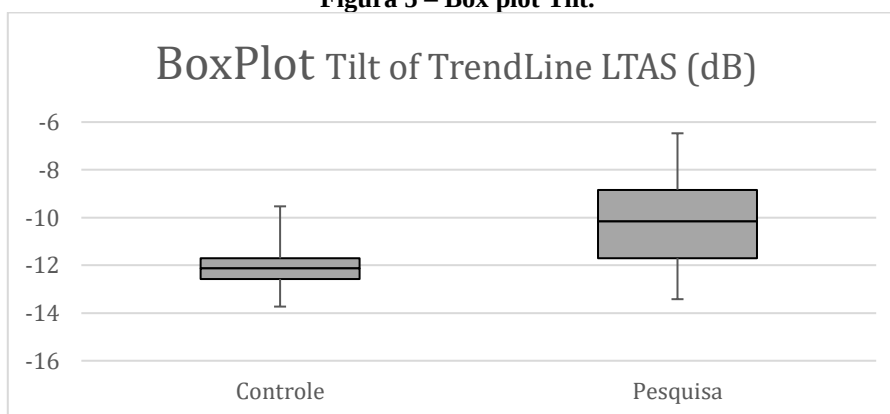


Figura 4 – Box plot Slope.



Fonte: Os Autores

Figura 5 – Box plot Tilt.



Fonte: Os Autores

4 CONCLUSÃO

Os resultados da análise visual dos diagramas de bloco em conjunto com os resultados dos testes t presentes no quadro 3 indicam que o AVQI em conjunto com as medidas acústicas se mostra promissor como uma ferramenta para fonoaudiólogos classificarem vozes com distúrbios na fala provenientes do edema de Reinke em relação a vozes sem distúrbios, com exceção do CPPs. Isso indica que no futuro podem ser desenvolvidos métodos de diagnosticar o edema de Reinke sem a necessidade de exames invasivos, como a laringoscopia, que são exames desconfortáveis ao paciente. Porém ainda existem limitações: o edema de Reinke não é a única doença que causa distúrbios na voz, ou seja, é necessário que mais pesquisas e trabalhos sejam desenvolvidos na área, pois somente esse trabalho, com esse banco de dados não garante que outro distúrbio na voz não acabe sendo classificado com um falso positivo.



AGRADECIMENTOS

Agradeço a Fundação Araucária pela bolsa, e também ao professor Marcelo de Oliveira Rosa que me auxiliou e instruiu-o para o desenvolvimento de trabalho.

REFERÊNCIAS

- [1] MARYN, Youri, et al. **Toward Improved Ecological Validity in the Acoustic Measurement of Overall Voice Quality: Combining Continuous Speech and Sustained Vowels.** Journal of Voice, Belgium, Vol. 24, No. 5, pp. 540-555, dez. 2008.
- [2] SODRÉ, Bruno Ribeiro. **Reconhecimento de padrões aplicados na identificação de patologias de laringe.** 2016. 107. Dissertação (Pós-graduação em Engenharia Elétrica e Informática Industrial) – Universidade Tecnológica Federal do Paraná, Curitiba, 2016.
- [3] HILLENBRAND, James.; CLEVELAND, Ronald A.; ERICKSON, Robert L. **Acoustic Correlates of Breathy Vocal Quality.** Journal of Speech and Hearing Research, Volume 37, pp. 769-778, August 1994.