

# Espectro de voz de pacientes diagnosticados com ELA bulbar

Análise do declive em bandas de energia de baixa frequência

## *Voice spectrum of patients diagnosed with bulbar ALS*

*Slope analysis in the lower energy bands*

Hugo Cordeiro

Depart. de Engenharia Eletrónica, Telecomunicações e de Computadores, Instituto Superior de Engenharia de Lisboa  
Instituto Politécnico de Lisboa  
Lisboa, Portugal  
hcordeiro@deetc.isel.ipl.pt

Carlos Meneses

Depart. de Engenharia Eletrónica, Telecomunicações e de Computadores, Instituto Superior de Engenharia de Lisboa  
Instituto Politécnico de Lisboa  
Lisboa, Portugal  
cmeneses@deetc.isel.ipl.pt

**Resumo** — Este estudo apresenta uma comparação entre a análise espectral de vozes de pacientes diagnosticados com Esclerose Lateral Amiotrófica (ELA) bulbar e oradores saudáveis. O principal objetivo é entender de que modo esta doença afeta a voz e consequentemente o espectro das vozes dos pacientes. Os espectros são analisados em duas bandas de energia, onde são caracterizados picos de energia e consequentemente também o declive entre estas bandas de energia. Os resultados obtidos permitem concluir que os sujeitos diagnosticados com ELA bulbar apresentam diferenças significativas nos valores médios e variâncias dos picos de energia em relação aos oradores saudáveis.

**Palavras Chave** – *Análise espectral; Vozes patológicas; Esclerose lateral amiotrófica.*

**Abstract** — This study presents a comparison between spectral analysis of voices of patients diagnosed with bulbar amyotrophic lateral sclerosis (ALS) and healthy speakers. The main objective is to understand how this disease affects the voice and, consequently, the spectrum of patients' voices. The spectra are analyzed in two energy bands, where energy peaks are characterized and, consequently, also the slope between these energy bands. The results obtained allow to conclude that subjects diagnosed with ALS bulbar present significant differences in the mean values and variances of the peak energy compared to the healthy speakers.

**Keywords** - *Spectral analysis, Pathological voices; Amyotrophic lateral sclerosis.*

### I. INTRODUÇÃO

A Esclerose Lateral Amiotrófica (ELA) é uma doença neurológica que evolui de forma progressiva. Não é uma doença contagiosa, mas é mortal para 50% dos pacientes nos primeiros 30 meses [1].

Esta doença é caracterizada pela morte prematura dos neurónios motores que conduzem a informação do cérebro para os músculos. A consequência da morte precoce desses neurónios é que esses músculos, que permitem realizar os mais variados movimentos (andar, falar, mastigar, engolir...), vão ficando gradualmente mais fracos, originando cansaço e podendo desenvolver atrofia muscular. Frequentemente os pacientes referem pequenos saltos involuntários nos músculos (fasciculações), fadiga e câibras [2].

A forma medular da ELA, a mais frequente, envolve músculos dos braços ou das pernas e pode provocar dificuldades ao nível da marcha. No caso de envolver os músculos das mãos, os pacientes podem ter dificuldade em, por exemplo, abotoar o casaco. A Esclerose Lateral Amiotrófica pode designar-se por “bulbar” quando está associada à dificuldade em articular palavras (disartria) [3], mastigar e engolir (disfagia) [4].

Apesar de ser uma doença rara, devido à sua gravidade a sua deteção precoce e monitorização é extremamente importante. Neste contexto a deteção usando um método não evasivo, como a fala, pode ser uma mais valia no aumento da qualidade de vida dos pacientes. Em [5] surge um primeiro trabalho que sugere o uso do processamento da fala para deteção preliminar de ELA bulbar. Mais recentemente foram publicados vários trabalhos onde a análise de fala é usada para deteção de ELA com recurso a características acústicas como o *jitter* e o *shimmer* [6]–[8], sendo nestes casos usados classificadores de alto nível como discriminadores lineares e redes neuronais. As taxas de acertos rondam os 90%.

O objetivo deste trabalho é avaliar se o declive espectral entre bandas de energia consegue ter informação relevante na caracterização das vozes de pacientes diagnosticados com ELA bulbar. Este parâmetro proposto em [9] e cujo algoritmo

otimizado em [10] demonstrou resultados em torno de 90% na detecção de vozes com patologias nas pregas vocais.

O ponto seguinte apresenta a base de dados usada neste trabalho. O ponto 3 descreve os métodos usados e o ponto 4 apresenta os resultados e a discussão. O artigo termina com as conclusões.

## II. BASE DE DADOS

A base de dados utilizada neste trabalho [7] é constituída por 15 oradores diagnósticos com ELA bulbar e 39 oradores saudáveis, totalizando 54 oradores. Os oradores diagnosticados com ELA bulbar estão repartidos entre 6 oradores masculinos e 9 femininos, com uma média de idades de 57,7 anos. Os oradores saudáveis estão repartidos por 23 oradores masculinos e 16 oradores femininos, com uma média de idades de 41,9 anos. O sinal analisado é a vogal /a/ sustentada, adquirido com vozeamento confortável. Os sinais foram adquiridos com um *smartphone* e *headset*, a uma frequência de amostragem de 44,1 kHz e 16 bit por amostra. A duração média dos sinais é cerca de 4,1 segundos.

## III. MÉTODO

Proposto em [9] o *LBST* (*Low Band Spectral TILT*) pretende caracterizar vozes patológicas através do declive entre duas bandas de baixa frequência na vogal /a/. A primeira banda contém, tipicamente, as duas primeiras harmónicas. A segunda banda é caracterizada pelo intervalo entre a terceira e a décima harmónica, onde se situa o típico primeiro formante da vogal /a/.

O declive das duas bandas é dado pelo quociente entre as diferenças de amplitude máxima das duas bandas, *FBME* (*First Band Maximum Energy*) e *SBME* (*Second Band Maximum Energy*), e a diferença das respetivas frequências, Eq. (1),

$$LBST = \frac{SBME - FBME}{SBME_{freq} - FBME_{freq}} \quad (1)$$

Verifica-se que tipicamente o *LBST* é positivo para o caso do orador saudável e negativo nos sujeitos não saudáveis, independentemente da evolução da doença.

No algoritmo inicialmente proposto a estimativa de *FBME*, *SBME* e consequentemente do limiar de separação entre bandas era baseado na estimativa da frequência fundamental. Esta tarefa pode ser particularmente difícil em vozes com maior degradação vocal, levando neste caso o algoritmo inicial a estimativas incorretas da frequência fundamental nas vozes patológicas e consequentemente a erros de classificação. O algoritmo proposto em [10] remove a estimativa da frequência fundamental do algoritmo e o limiar para a separação das bandas *FBME* e *SBME* para a ter um valor fixo. Verificou-se que nos resultados obtidos que a nova abordagem permite um aumento de 5% na discriminação entre vozes com patologia das pregas vocais e saudáveis.

Neste trabalho toda a base de dados apresentada no ponto anterior foi considerada para análise. Nesta fase da investigação não se pretende construir um sistema de classificação baseado em treino e teste, mas verificar se o *LBST* permite discriminar entre vozes diagnosticadas com ELA bulbar e saudáveis. Assim, toda a base de dados foi analisada.

Os sinais de fala foram processados em tramas com 30 ms de duração com um andamento de 10 ms. O limiar de separação entre bandas usado é de 420 Hz e a FFT tem uma dimensão de 4096 pontos. As Fig. 1 e 2 mostram, para uma trama do sinal da vogal sustentada /a/, declive negativo num orador diagnosticado com ELA e declive positivo para um orador saudável, assim como as zonas que são consideradas para o cálculo do *FBME* entre [0-420] Hz e o *SBME* entre [420-1260].

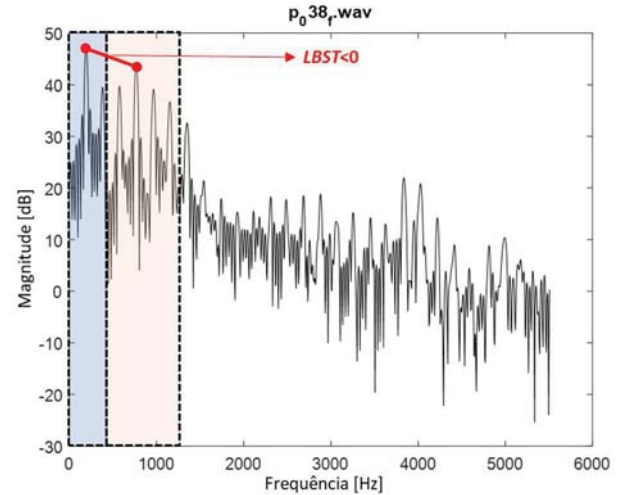


Figure 1. Espectro de um orador feminino diagnosticado com ELA.

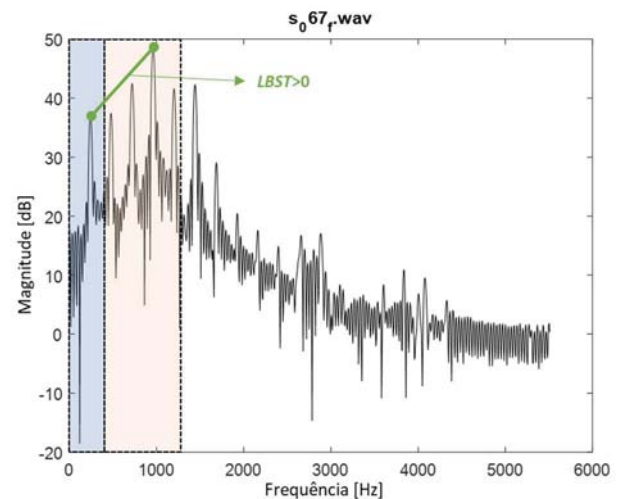


Figure 2. Espectro de um orador feminino saudável.

## IV. DISCUSSÃO E RESULTADOS

Numa primeira abordagem pretende-se verificar qual o impacto do *LBST* na detecção de oradores diagnosticados com ELA bulbar. Na Fig. 3, que apresenta dos valores médios do *LBST* para todos os oradores da base de dados, verifica-se que os oradores com patologia apresentam tipicamente valores de *LBST* menores que zero, ou seja, declive negativo. Nesta classe apenas 3 dos 15 oradores apresentam declive positivo neste parâmetro. Por outro lado, nos oradores saudáveis, 22 em 39 oradores saudáveis apresentam declive positivo, sendo que os restantes 17 apresentam declive negativo e poderiam ser classificados como oradores não saudáveis. De facto, não é eficiente propor apenas o valor médio do *LBST* como parâmetro de separação.

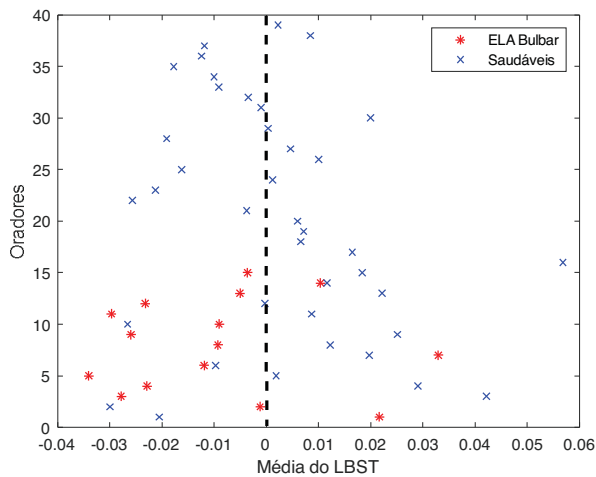


Figure 3. Valores médios do *LBST* por Orador.

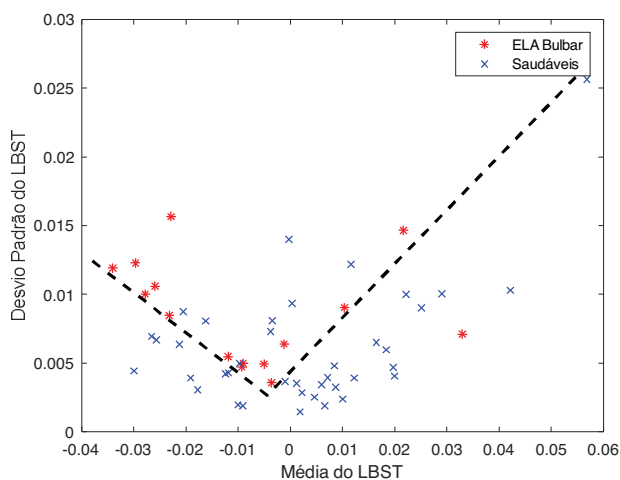


Figure 4. Valores Médios e Desvio Padrão do *LBST*

Contudo, observando a Fig. 4 que apresenta o desvio padrão em função dos valores médios do *LBST*, por orador, verifica-se que os valores de dispersão dos oradores diagnosticados com ELA bulbar é tipicamente superior aos dos oradores saudáveis. Tendo como referência o desvio padrão em função do valor médio do *LBST* verifica-se que apenas 1 orador não saudável seria incorretamente classificado e que apenas 8 oradores saudáveis não seriam corretamente classificados. De facto, verifica-se que os oradores saudáveis apresentam tipicamente um valor de *LBST* mais estável, ou seja, um desvio padrão inferior aos dos oradores não saudáveis que por sua vez apresentam menos estabilidade vocal.

As Fig. 5 e 6 apresentam os valores médios e o desvio padrão da energia respetivamente para a primeira (*FBME*) e a segunda (*SBME*) banda do espectro do sinal. Nesta análise pretende-se verificar quais dos parâmetros de *LBST*, dados na Eq. 1, contribuem para os resultados discutidos nos parágrafos anteriores. De notar que a primeira banda termina nos 420 Hz e a segunda banda termina nos 1260 Hz. Sendo o sinal em análise a vogal /a/ o primeiro formante encontra-se tipicamente dentro da segunda banda.

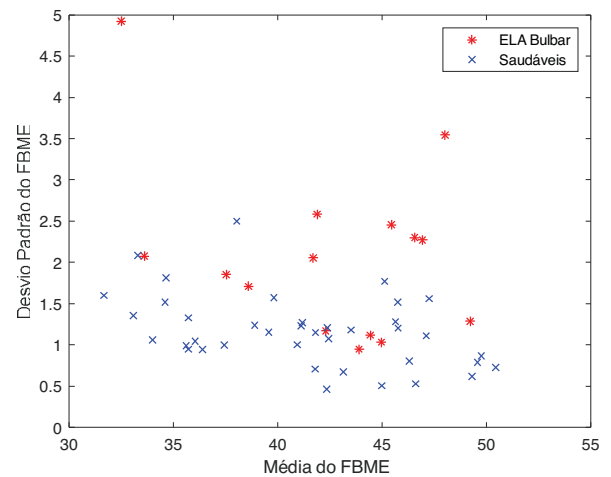


Figure 5. Valores médios e desvio padrão do *FBME*.

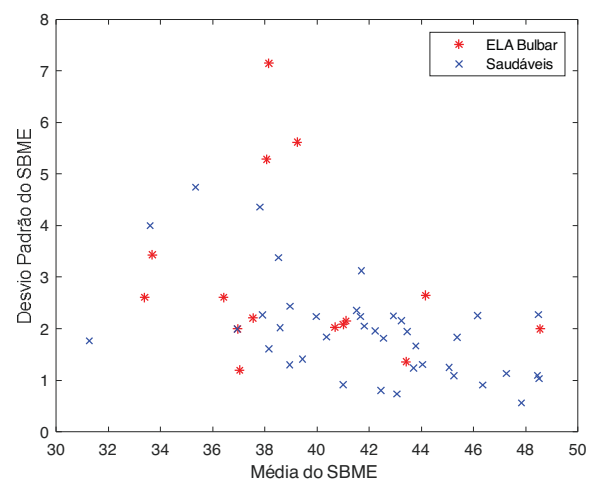


Figure 6. Valores médios e desvio padrão do *SBME*.

Na Fig. 5 verifica-se que os valores de *FBME* têm um desvio padrão inferior nos oradores saudáveis. Os oradores saudáveis têm também valores de energia mais baixos que os oradores não saudáveis, mas é de facto neste caso ao nível da dispersão que se verificam as maiores diferenças. Por outro lado, na Fig. 6, relativa ao *SBME*, verifica-se que ao nível da dispersão não existe um padrão distinto, mas já ao nível dos valores médios verifica-se que os oradores saudáveis apresentam valores tipicamente superiores aos oradores não saudáveis.

As Fig. 7 e 8 apresentam os desvios padrão das frequências do *FBME* e *SBME* em relação ao seu valor médio. Verifica-se que não apresentam padrões diferentes entre oradores saudáveis não saudáveis. Deste modo, de momento, é inconclusiva a existência de informação para classificação nestes parâmetros.

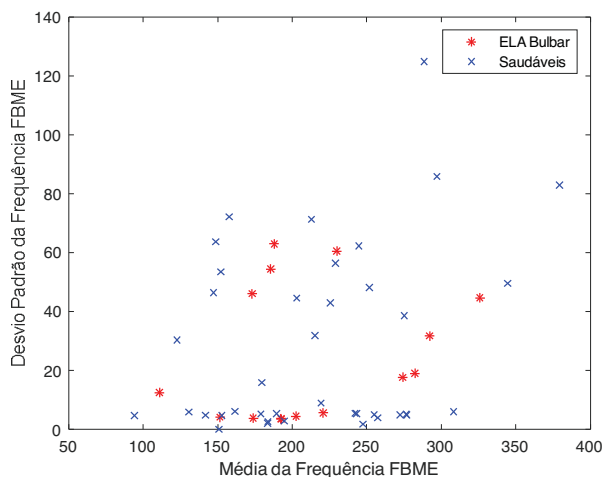


Figure 7. Valores médios e desvio padrão da frequência na FBME.

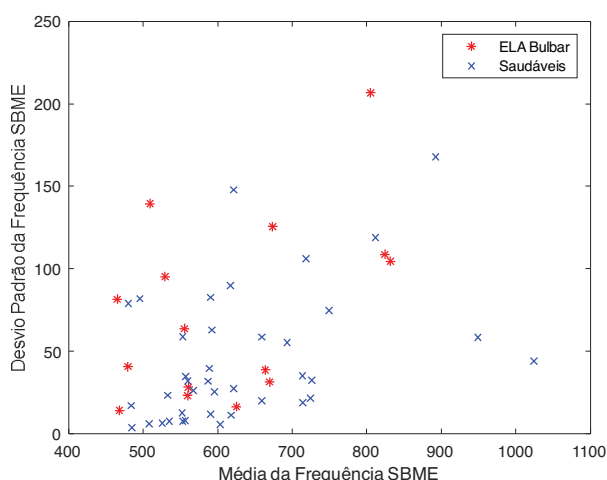


Figure 8. Valores médios e desvio padrão da frequência na SBME.

Através das análises anteriores verifica-se que o *LBST* apresenta potencial na classificação de oradores com ELA bulbar através da análise da vogal /a/. Os desvios no *LBST* são tipicamente maiores nos oradores com esta patologia e devem-se às variações de energia na primeira banda do espectro. Esta primeira banda caracteriza as primeiras harmónicas do sinal, relacionadas com a frequência fundamental. Pode-se, portanto, concluir que os sujeitos com maior degradação vocal têm também uma maior dificuldade em manter constante a energia de vibração das pregas vocais (*shimmer*). Pelo contrário, os oradores saudáveis apresentam melhor qualidade vocal, com o primeiro formante mais bem definido e com energia superior na segunda banda relativamente à primeira banda.

## V. CONCLUSÕES

Este artigo apresenta um estudo dos espectros de vozes de oradores diagnosticados com ELA bulbar. Através da análise do espectro da vogal /a/ sustentada verifica-se que os oradores

diagnosticados com ELA bulbar apresentam tipicamente um declive espectral negativo (*LBST*), entre duas bandas de baixa frequência. A variação do *LBST* apresenta tipicamente uma maior variabilidade quando comparado com oradores saudáveis. Note-se que são usadas medidas simples e análise visual sem recursos a classificadores de alta complexidade.

O estudo apresenta uma potencial taxa de acerto de cerca 83%, com erros de 9 oradores em 54, havendo apenas um erro nos oradores diagnosticados com ELA bulbar, ou seja, uma taxa de falsos negativos reduzida. Note-se que é preferível aumentar a taxa de oradores saudáveis diagnosticados erradamente, mas que será sujeito a outras análises, do que não diagnosticar um orador com patologia.

Como trabalho futuro propõem-se a continuação do estudo dos espectros e a procura de novos parâmetros capazes de separar oradores saudáveis de oradores com ELA bulbar, assim como a implementação de um reconhecedor automático usando estes e outros parâmetros espectrais, que permita o rastreio, avaliação e acompanhamento da doença.

## REFERÊNCIAS BIBLIOGRÁFICA

- [1] C. Ingre, P. M. Roos, F. Piehl, F. Kamel, and F. Fang, "Risk factors for amyotrophic lateral sclerosis," *Clin. Epidemiol.*, vol. 7, pp. 181–193, 2015.
- [2] Associação Portuguesa de Esclerose Lateral Amiotrófica, "Quais os sintomas que a caracterizam e quais as formas de apresentação inicial da ELA." [Online]. Available: <https://www.apela.pt/page/51/quais-os-sintomas-que-a-caracterizam-e-quais-as-formas-de-apresentacao-inicial-da-ela>.
- [3] B. Tomik and R. J. Guiloff, "Dysarthria in amyotrophic lateral sclerosis: A review," *Amyotroph. Lateral Scler.*, vol. 11, no. 1–2, pp. 4–15, 2010.
- [4] R. T. Pontes, M. Orsini, M. R. G. De Freitas, R. De Souza Antonoli, and O. J. M. Nascimento, "Alterações da fonação e deglutição na esclerose lateral amiotrófica: Revisão de literatura," *Rev. Neurociencias*, vol. 18, no. 1, pp. 69–73, 2010.
- [5] J. R. Green *et al.*, "Bulbar and speech motor assessment in ALS: Challenges and future directions," *Amyotroph. Lateral Scler. Front. Degener.*, vol. 14, no. 7–8, pp. 494–500, 2013.
- [6] R. Norel, M. Pietrowicz, C. Agurto, S. Rishoni, and G. Cecchi, "Detection of amyotrophic lateral sclerosis (ALS) via acoustic analysis," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2018-Sept, no. May, pp. 377–381, 2018.
- [7] M. Vashkevich, A. Petrovsky, and Y. Rushkevich, "Bulbar ALS Detection Based on Analysis of Voice Perturbation and Vibrato," in *Signal Processing - Algorithms, Architectures, Arrangements, and Applications Conference Proceedings, SPA*, 2019, vol. 2019-Sept, pp. 267–272.
- [8] K. H. An *et al.*, "Automatic early detection of amyotrophic lateral sclerosis from intelligible speech using convolutional neural networks," *Proc. Annu. Conf. Int. Speech Commun. Assoc. INTERSPEECH*, vol. 2018-Sept, no. August, pp. 1913–1917, 2018.
- [9] H. Cordeiro and C. Meneses, "Low band spectral tilt analysis for pathological voice discrimination," in *2019 IEEE 6th Portuguese Meeting on Bioengineering (ENBENG)*, 2019, pp. 1–4.
- [10] H. Cordeiro and C. Meneses, "An improved algorithm for the Low Band Spectral Tilt estimation for pathological voice detection," in *Signal Processing - Algorithms, Architectures, Arrangements, and Applications Conference Proceedings, SPA*, 2019, pp. 202–207.