SUPER-GAUSSIANITY OF SPEECH SPECTRAL COEFFICIENTS AS A POTENTIAL BIOMARKER FOR DYSARTHRIC SPEECH DETECTION

Ina Kodrasi, Hervé Bourlard

Idiap Research Institute, Speech and Audio Processing Group, Martigny, Switzerland
{ina.kodrasi, herve.bourlard}@idiap.ch

ABSTRACT

Parkinson's disease (PD) and Amyotrophic Lateral Sclerosis (ALS) are progressive neurodegenerative diseases which, among other symptoms, cause dysarthria of speech. To assist the clinical diagnosis and treatment of neurological diseases, several studies have addressed the characterization and classification of healthy and dysarthric speech. However, most contributions deal with PD speech, with significantly fewer results presented for ALS speech. The objective of this paper is to show that ALS speech has a similar statistical distribution as PD speech, with the complex spectral coefficients being significantly less super-Gaussian than healthy speech spectral coefficients. In addition, a method to exploit the super-Gaussianity of speech signals as a feature to classify healthy and dysarthric speech is presented and evaluated. The proposed approach is evaluated on a French database of healthy and dysarthric (PD and ALS) speech. Experimental results show that the use of the super-Gaussianity of speech signals yields a significantly higher classification accuracy than state-of-the-art features such as fundamental frequency, jitter, shimmer, harmonics-to-noise ratio, or Mel frequency cepstral coefficients.

Index Terms— super-Gaussianity, Weibull distribution, SVM, Parkinson's disease, Amyotrophic Lateral Sclerosis

1. INTRODUCTION

Due to increasing population numbers and aging, the prevalence of neurological diseases such as Parkinson's disease (PD) and Amyotrophic Lateral Sclerosis (ALS) is also increasing [1, 2]. The number of people requiring screening and treatment will continue to grow in the coming decades, likely putting a strain on the health care system. To diagnose neurological diseases and evaluate their progression, neurologists exploit several examinations which asses different motor and sensory skills. However, such examinations are subject to the expertise of the neurologist and might be affected by his familiarity with the patient. A common symptom of PD and ALS is dysarthria of speech, which results from disturbances to the muscular control of the movement mechanism necessary for the execution of speech [3]. Dysarthria affects several components of the speech production mechanism such as phonation and articulation, causing the speech of PD and ALS patients to be characterized by imprecise consonants, distorted vowels, harsh voice, monopitch, monoloudness, and breathiness [4]. To assist the clinical diagnosis and treatment of neurological diseases with objective tools, there has been a growing interest in the research community to develop reliable features to characterize and classify healthy and dysarthric speech [5–20]. While most contributions deal with PD speech [5–16], significantly fewer results on ALS speech have been reported [17–20].

Impacted phonation in PD speech has been characterized using features such as fundamental frequency f_0, jitter, shimmer, or harmonics-to-noise ratio (HNR) [5–8]. These features aim at quantifying the disturbances in the vocal fold vibration as well as the excessive turbulence due to incomplete closure of the vocal folds. Impacted articulation on the other hand has been characterized by assessing the trajectories of the formant frequencies and by computing measures such as the vowel space area (VSA), vowel articulation index, consonant spectral trend, or formant centralization ratio [8–12]. Additionally, vocal tract atypicalities have been successfully described using Mel frequency cepstral coefficients (MFCCs), linear prediction coefficients (LPCs), or perceptual LPCs in [8, 13–15]. Several of the above mentioned features have also been used to assess impacted phonation and articulation in ALS speech, such as jitter, shimmer, HNR, and MFCCs in [17, 18] or formant trajectories and VSA in [19, 20].

In [16] we have shown that due to impacted phonation and articulation, the statistical distribution of PD speech spectral coefficients differs from healthy speech spectral coefficients. Modeling the distribution of the spectral coefficients using a Weibull distribution [21], it has been shown that PD speech is less super-Gaussian than healthy speech. The objective of this paper is two-fold. First, we show that this result generalizes to dysarthric speech arising due to other neurological diseases such as ALS. By modeling the global distribution of the spectral coefficients in each frequency bin, it is shown that the distribution of PD and ALS speech is very similar and significantly less super-Gaussian than healthy speech. Second, we evaluate a method to exploit the super-Gaussianity of speech signals for automatic classification of healthy and dysarthric speech. It is shown that using the super-Gaussianity of speech spectral coefficients as a feature in a support vector machine (SVM) for classification yields a high performance, significantly outperforming using state-of-the-art features such as f_0, jitter, shimmer, HNR, or MFCCs.

2. SUPER-GAUSSIANITY OF SPEECH SPECTRAL COEFFICIENTS

Empirical observations have shown that the distribution of the complex healthy speech spectral coefficients is super-Gaussian [22–24]. Super-Gaussianity of the speech spectral coefficients arises due to pauses between phonemes and due to formant transitions in voiced sounds. However, dysarthric speech caused by PD and ALS is characterized by imprecise consonants, harsh voice, monopitch, monoloudness, and breathiness [4]. Our hypothesis is that *i) these characteristics result in a spectro-temporal smearing of the energy* *of the speech signal in the short-time Fourier transform (STFT) domain for both PD and ALS patients; ii) this excessive smeared energy fills in the pauses between phonemes and formant transitions, resulting in PD and ALS speech spectral coefficients being less super-Gaussian than healthy speech spectral coefficients; and iii) the super-Gaussianity of the spectral coefficients can be successfully used to automatically classify healthy and dysarthric (PD or ALS) speech*. To verify this hypothesis, one should model the distribution of the speech spectral coefficients and assess their super-Gaussianity.

To model the magnitude of speech spectral coefficients, several distributions have been used in the literature. In [22, 23, 25], the magnitude of spectral coefficients has been modeled using a Gamma distribution or a Chi distribution. In [16, 24], a Weibull distribution has been used. The Weibull distribution is characterized by a shape parameter and it has been shown that smaller shape parameter values describe more super-Gaussian distributed complex spectral coefficients [24]. In this paper, the global distribution of the magnitude of the speech spectral coefficients is modeled using the Weibull distribution. However, it should be noted that the presented results and conclusions can also be derived using the Gamma or Chi distributions. In the remainder of this section, the Weibull distribution is briefly reviewed and the maximum likelihood (ML) estimation procedure of the shape parameter is briefly discussed. For additional details on the distribution model and the ML estimation procedure, the reader is referred to [16].

Speech spectral coefficients are denoted by $S_k(l)$, with k the frequency index and l the frame index. In addition, λ_k^2 denotes the average speech power spectral density (PSD), i.e., $\lambda_k^2 = \mathcal{E}\{|S(k)|^2\}$, with $\mathcal{E}$ the expected value operator. Modeling the magnitude of the spectral coefficients $|S_k|$ using a Weibull distribution results in the probability density function [16, 21, 24]

$$p(|S_k|) = \frac{\beta_k}{\alpha_k}\left(\frac{|S_k|}{\alpha_k}\right)^{\beta_k - 1} e^{-\left(\frac{|S_k|}{\alpha_k}\right)^{\beta_k}}, \quad (1)$$

where β_k denotes the shape parameter and α_k denotes the scale parameter. The scale parameter α_k can be expressed in terms of the average PSD λ_k^2 and shape parameter β_k as

$$\alpha_k = \frac{\lambda_k}{\sqrt{\Gamma\left(1 + \frac{2}{\beta_k}\right)}}, \quad (2)$$

with $\Gamma(\cdot)$ denoting the gamma function [26]. For $\beta_k < 2$, the Weibull distribution models super-Gaussian distributed complex spectral coefficients, with lower values of β_k corresponding to more super-Gaussian distributed complex spectral coefficients [24]. In order to estimate the shape parameter β_k, an ML estimator is used in this paper. Given the speech spectral coefficients at frequency index k, i.e., $S_k(1),\ S_k(2),\ \ldots,\ S_k(L)$, with L the total number of frames, an ML estimate of the shape parameter β_k can be obtained by solving the optimization problem

$$\min_{\beta_k} \quad \Bigg[L \log \beta_k - L\beta_k \log \alpha_k + (\beta_k - 1)\sum_{l=1}^{L} \log |S_k(l)| - \sum_{l=1}^{L}\left(\frac{|S_k(l)|}{\alpha_k}\right)^{\beta_k}\Bigg], \quad (3)$$

with α_k given by (2). Since no analytical solution to (3) exists, an iterative optimization technique should be used. In this paper, the one-dimensional quasi-Newton method is used.

3. RESULTS AND DISCUSSION

In this section, empirical analyses of the distribution of spectral coefficients for healthy and dysarthric speech arising due to PD and ALS are presented. In addition, statistical significance analyses are conducted to compare the mean shape parameter values across healthy speakers, PD patients, and ALS patients. Finally, classification results for healthy and dysarthric speakers are presented. We investigate two methods for constructing the feature vector for the classifier, i.e., i) using the Weibull shape parameter values from all frequencies, and ii) using the Weibull shape parameter values only from those frequencies where there is a statistically significant difference between the mean shape parameter values of healthy and dysarthric speakers. Frequencies with a statistically significant difference are determined on a training set.[1] Classification results using the proposed feature vectors in an SVM are compared to using state-of-the-art features such as the statistics of f_0, jitter, shimmer, HNR, and MFCCs.

3.1. Databases and preprocessing

For dysarthric speech recordings arising due to PD, we consider French recordings of 10 PD patients (6 males, 4 females) from [27], with all speakers being French native speakers.[2] The age of the PD patients ranges from 47 to 70 years old, with the average age being 60 years old.

For dysarthric speech recordings arising due to ALS, we consider French recordings of 10 ALS patients (5 males, 5 females) from the University of Geneva, with all speakers being French native speakers. The age of the ALS patients ranges from 56 to 81 years old, with the average age being 72 years old.

Finally, for healthy speech recordings, we consider French recordings of 20 healthy speakers (9 males, 11 females) from [28], with all speakers being French native speakers. The age of the healthy speakers ranges from 30 to 82 years old, with the average age being 56 years old.

The sampling frequency of all recordings is 44.1 kHz. The data has been recorded based on the MonPaGe speech protocol [28], where (among other tasks) patients are asked to read samples from a list of pseudo-words, i.e., strings of characters resembling real words but having no meaning. To discard possible differences in the computed features due to possibly different speech content, we consider only the common pseudo-words among all speakers, i.e., pseudo-words that are uttered by all speakers. The number of such common pseduo-words is 25. Concatenating all common pseudo-words for each speaker yields a signal with an average length of 20.5 s for the PD patients, 20.6 s for the ALS patients, and 19.1 s for the healthy speakers.

All recordings are downsampled to 16 kHz and manual voice activity detection is performed to extract the speech-only segments. Since the loudness level across different recording sessions might not have been kept constant (possibly influencing the computed features), each pseudo-word is normalized to a root mean square level of −30 dB. The signals are processed in a weighted overlap-add STFT framework using a tight analysis window with a frame size of 256 samples and an overlap of 50%.

[1] It should be noted that such a method for constructing the feature vector can be considered as a feature dimensionality reduction technique.

[2] Approval from Swissethics, number 2015-0002800 – (15-258)

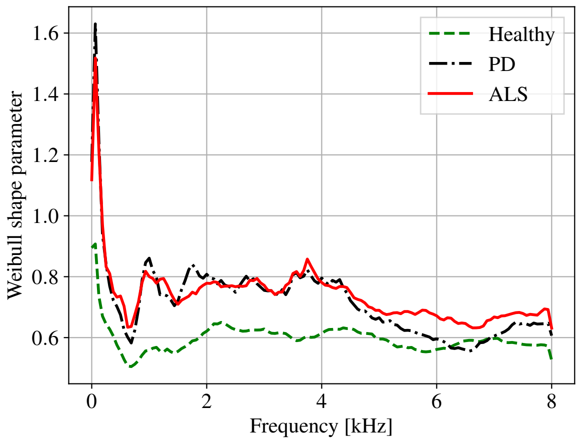

Fig. 1. Frequency-dependent Weibull shape parameter averaged across 20 healthy speakers, 10 PD patients, and 10 ALS patients. Shape parameters for each speaker are estimated from a signal of 25 pseudo-words, with an average length of 19.1 s for the healthy speakers, 20.5 s for the PD patients, and 20.6 s for the ALS patients.

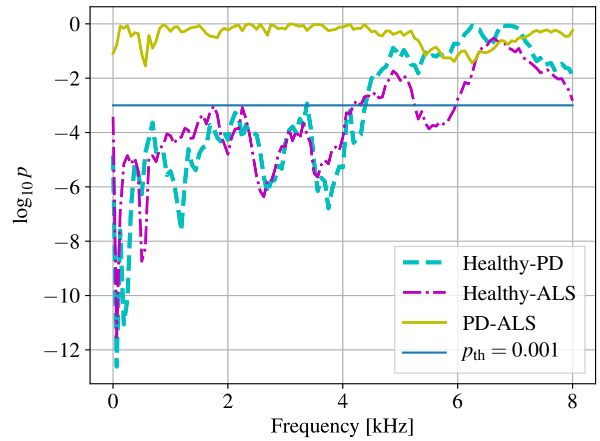

Fig. 2. Independent samples t-test comparing the mean Weibull shape parameters between healthy speakers and PD patients, healthy speakers and ALS patients, and PD patients and ALS patients. The selected statistical significance threshold $p_{\text{th}} = 0.001$ is also depicted.

3.2. Super-Gaussianity of healthy, PD, and ALS speech spectral coefficients

In this section, the super-Gaussianity of the spectral coefficients for different groups of speakers is compared. The Weibull shape parameters for each speaker are computed as following. For each frequency index k, spectral magnitudes $|S_k(1)|, |S_k(2)|, \cdots, |S_k(L)|$ are extracted and the average speech PSD λ_k^2 is computed. The scale parameter α_k is expressed in terms of the average speech PSD λ_k^2 and of the shape parameter β_k using (2). The ML estimate of the frequency-dependent shape parameter β_k is then obtained by solving (3). To initialize the iterative minimization procedure, $\beta_k = 2$ is used. Taking into account only half of the spectrum, the shape parameter vector is a 129-dimensional vector.

Fig. 1 depicts the frequency-dependent shape parameter values averaged across all healthy speakers, PD patients, and ALS patients. It can be observed that as expected, the distribution of speech spectral magnitudes is closer to an exponential distribution than to a Rayleigh distribution (i.e., $\beta_k < 1$), independently of whether healthy or dysarthric speech is considered. In addition, it can be observed that the shape parameters are smaller for healthy speech than for dysarthric PD or ALS speech, i.e., *healthy speech is more super-Gaussian than dysarthric speech*, independently of the neurological disease causing the dysarthria. This difference between the shape parameter values is particularly large at low frequencies. At higher frequencies, the spectral coefficients reflect mainly recording noise, and hence, the estimated shape parameters do not necessarily describe the distribution of the speech spectral coefficients. Finally, it can be observed that the shape parameter values for PD and ALS speech are very similar (particularly at low frequencies), showing that *dysarthric speech has a similar statistical distribution and a similar super-Gaussianity*, independently of the neurological disease causing the dysarthria.

To determine whether the previously discussed results are statistically significant, a statistical analysis is conducted. Shapiro-Wilk tests [29] indicate that the shape parameter values for the considered groups of speakers are normally distributed. To compare the difference in mean shape parameter values between two groups of speakers, an independent samples t-test is conducted. The difference for each frequency is assessed for healthy speakers and PD patients, healthy speakers and ALS patients, and PD patients and ALS patients. Fig. 2 depicts the obtained p-values (in logarithmic scale) for each frequency and each comparison. In addition, the selected threshold p-value $p_{\text{th}} = 0.001$ is also depicted. For $p < p_{\text{th}}$, we consider there to be a statistically significant difference between the compared mean shape parameters. It can be observed that *the difference in shape parameter values between healthy and dysarthric speakers is statistically significant* for frequencies below (approximately) 4 kHz. In addition, it can be observed that as expected, *the difference in shape parameter values between PD and ALS patients is not significant* for any frequency.

In summary, the presented results show that the distribution of PD and ALS speech is very similar and significantly less super-Gaussian than healthy speech. Furthermore, the presented results show that the frequency-dependent shape parameter (particularly at lower frequencies dominated by speech) can be a robust feature to discriminate between healthy and dysarthric speech arsing due to PD or ALS.

3.3. Classification of healthy and dysarthric speech

In this section we evaluate the performance of a classifier aiming to discriminate between healthy and dysarthric speech based on the super-Gaussianity of the spectral coefficients. It should be noted that dysarthric speech refers to both PD and ALS speech and that we do not aim to discriminate between the different neurological diseases.

As in [5, 6, 12–14, 17, 18], classification is done using an SVM with a radial basis kernel function. The validation strategy is a stratified 5-fold cross validation, ensuring that each fold has the same number of healthy and dysarthric speakers. Features are normalized

Table 1. Classification results of healthy and dysarthric (PD and ALS) speakers using different feature vectors.

Performance [%]	Accuracy	Sensitivity	Specficty
f_0	65.0 ± 16.6	75.0 ± 27.4	55.0 ± 33.2
Jitter	62.5 ± 7.9	45.0 ± 18.7	80.0 ± 18.7
Shimmer	87.5 ± 0.0	85.0 ± 12.2	90.0 ± 12.2
HNR	80.0 ± 12.7	90.0 ± 12.2	70.0 ± 24.5
MFCCs	77.5 ± 9.4	90.0 ± 12.2	65.0 ± 20.0
$\boldsymbol{\beta}$	92.5 ± 10.0	90.0 ± 12.2	95.0 ± 10.0
$\boldsymbol{\beta}_p$	$\mathbf{95.0 \pm 6.1}$	$\mathbf{95.0 \pm 10.0}$	$\mathbf{95.0 \pm 10.0}$

using the mean and standard deviation of the training data in each fold. The classification performance is evaluated in terms of the mean and standard deviation of the accuracy, sensitivity, and specificity on the test set across all folds. Accuracy refers to the percentage of correctly classified speakers, sensitivity refers to the percentage of correctly classified dysarthric speakers, and specificity refers to the percentage of correctly classified healthy speakers. To select the soft margin constant C and the kernel width γ for the SVM, a grid search is performed with $C, \gamma \in \{10^{-1}, 10^0, 10^1, 10^2\}$. The final hyper-parameters are selected as the ones resulting in the highest mean test accuracy. Such a selection criterion might yield final performance estimates that are over-optimistically biased. However, our ultimate goal is to compare the relative performance differences between using different features, with the absolute performance values being less important.

As previously mentioned, the classification performance is analyzed for two different methods of constructing the feature vector based on the shape parameter. First, the feature vector is set to be the 129-dimensional shape parameter vector across all frequencies. Second, only shape parameters from those frequencies with a statistically significant difference between healthy and dysarthric speakers are used. The significance threshold is set to $p_{\text{th}} = 0.001$ and the significance analysis is done using the healthy and dysarthric speakers in the training set in each fold. As a result, the dimension of the shape parameter feature vector might slightly change depending on the speakers in the training set. For the results presented in the following, this vector is on average (across all folds) a 75-dimensional vector. To compare the proposed features to state-of-the-art features, we also compute f_0, jitter, shimmer, HNR, and 14 MFCCs using the open-source toolkit openSMILE [30]. Feature vectors are then constructed using four functionals for each quantity, i.e., mean, standard deviation, kurtosis, and skewness. Hence, the feature vectors for f_0, jitter, shimmer, and HNR are 4-dimensional vectors, whereas the feature vector for MFCCs is a 56-dimensional vector (14 MFCCs $\times$ 4 functionals).

Table 1 presents the accuracy, sensitivity, and specificity of the SVM using the different considered feature vectors. The vector $\boldsymbol{\beta}$ denotes the 129-dimensional shape parameter feature vector, whereas the vector $\boldsymbol{\beta}_p$ denotes the 75-dimensional shape parameter feature vector. It can be observed that the proposed features $\boldsymbol{\beta}$ and $\boldsymbol{\beta}_p$ result in a very high classification performance, with an accuracy of 92.5% and 95.0% outperforming all considered state-of-the-art features. In addition, using $\boldsymbol{\beta}_p$ outperforms all considered state-of-the-art features in terms of sensitivity and specificity as well. Using $\boldsymbol{\beta}_p$ instead of $\boldsymbol{\beta}$ yields an accuracy increase of 2.5% and a sensitivity increase of 5.0%, showing that reducing the dimension of the feature vector based on the statistically significant differences is beneficial for the classifier. It should be noted that further analyses (not presented here due to space constraints) have shown that for all considered feature vectors, the number of miss-classified PD or ALS patients is very similar. These analyses confirm that the considered features characterize dysarthric speech similarly, independently of the neurological disease causing it.

In summary, the presented results confirm that the classification methodology proposed in this paper is more accurate than the considered state-of-the-art approaches in classifying healthy and dysarthric French speech arising due to PD or ALS.

4. CONCLUSION

In this paper the statistical distribution of the speech spectral coefficients in ALS patients has been compared to the distribution of PD and healthy speech spectral coefficients. It has been shown that ALS speech has a similar statistical distribution as PD speech. In addition, it has been shown that due to imprecise consonants, distorted vowels, harsh voice, monopitch, monoloudness, and breathiness, both ALS and PD speech are significantly less super-Gaussian than healthy speech. To classify healthy and dysarthric (PD or ALS) speech, it has been proposed to construct a feature vector using the Weibull shape parameters from frequencies with a statistically significant difference between the mean shape parameter values of healthy and dysarthric speech. On a French database of healthy, PD, and ALS speech, it has been shown that using the proposed feature vector in an SVM yields a high classification accuracy, significantly outperforming considered state-of-the-art feature vectors such as the fundamental frequency f_0, jitter, shimmer, HNR, or MFCCs.

5. ACKNOWLEDGMENT

The authors would like to acknowledge the support of the Swiss National Science Foundation project no CRSII5_173711 "MoSpeeDi" on "*Motor Speech Disorders: characterizing phonetic speech planning and motor speech programming/execution and their impairments*". They would also like to thank the project partners from University of Paris III: Sorbonne Nouvelle, Geneva University Hospitals, and University of Geneva for a fruitful collaboration.

6. REFERENCES

[1] C. Marras, J. C. Beck, J. H. Bower, E. Roberts, B. Ritz, G. W. Ross, R. D. Abbott, R. Savica, S. K. Van Den Eeden, A. W. Willis, and C. M. Tanner, "Prevalence of Parkinson's disease across North America," *Nature Partner Journals Parkinson's Disease*, vol. 4, no. 21, Dec. 2018.

[2] K. C. Arthur, A. Calvo, T. R. Price, J. T. Geiger, A. Chiò, and B. J. Traynor, "Projected increase in Amyotrophic Lateral Sclerosis from 2015 to 2040," in *Nature Communications*, Aug. 2016, vol. 7.

[3] J. R. Duffy, *Parkinsons Disease and Movement Disorders: Diagnosis and Treatment Guidelines for the Practicing Physician*, chapter Motor speech disorders: Clues to neurologic diagnosis, pp. 35–53, Humana Press, Totowa, NJ, 2000.

[4] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246–269, July 1969.

[5] M. A. Little, P. E. McSharry, E. J. Hunter, J. Spielman, and L. O. Ramig, "Suitability of dysphonia measurements for telemonitoring of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1015–1022, Apr. 2009.

[6] A. Tsanas, M. A. Little, P. E. McSharry, J. Spielman, and L. O. Ramig, "Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 5, pp. 1264–1271, May 2012.

[7] D. Sztahó, G. Kiss, and K. Vicsi, "Estimating the severity of Parkinson's disease from speech using linear regression and database partitioning," in *16th Annual Conference of the International Speech Communication Association*, Dresden, Germany, Sept. 2015, pp. 498–502.

[8] D. Hemmerling, J. R. Orozco-Arroyave, A. Skalski, J. Gajda, and E. Nöth, "Automatic detection of Parkinson's disease based on modulated vowels," in *17th Annual Conference of the International Speech Communication Association*, San Francisco, USA, Sept. 2016, pp. 1190–1194.

[9] S. Sapir, L. O. Ramig, J. L. Spielman, and C. Fox, "Formant centralization ratio: a proposal for a new acoustic measure of dysarthric speech," *Journal of Speech, Language and Hearing Research*, vol. 53, no. 1, pp. 114–125, Feb. 2010.

[10] S. Skodda, W. Visser, and U. Schlegel, "Vowel articulation in Parkinson's disease," *Journal of Voice*, vol. 25, no. 4, pp. 467–472, July 2011.

[11] J. Rusz, R. Cmejla, T. Tykalova, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, and E. Ruzicka, "Imprecise vowel articulation as a potential early marker of Parkinson's disease: Effect of speaking task," *Journal of the Acoustical Society of America*, vol. 134, no. 3, pp. 2171–2181, Sept. 2013.

[12] M. Novotný, J. Rusz, R. Čmejla, and E. Růžička, "Automatic evaluation of articulatory disorders in Parkinson's disease," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 9, pp. 1366–1378, Sept. 2014.

[13] J. R. Orozco-Arroyave, F. Hönig, J. Arias-Londoño, J. Bonilla, S. Skodda, J. Rusz, and E. Nöth, "Voiced/unvoiced transitions in speech as a potential bio-marker to detect Parkinson's disease," in *16th Annual Conference of the International Speech Communication Association*, Dresden, Germany, Sept. 2015, pp. 95–99.

[14] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londo no, J. F. Vargas-Bonilla, and E. Nöth, "Spectral and cepstral analyses for Parkinson's disease detection in Spanish vowels and words," *Expert Systems: The Journal of Knowledge Engineering*, vol. 32, no. 6, pp. 688–697, Dec. 2015.

[15] L. Moro-Velázquez, J. A. Gómez-García, J. I. Godino-Llorente, J. Villalba, J. R. Orozco-Arroyave, and N. Dehak, "Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect Parkinson's Disease," *Applied Soft Computing*, vol. 62, pp. 649–666, Jan. 2018.

[16] I. Kodrasi and H. Bourlard, "Statistical modeling of speech spectral coefficients in patients with Parkinsons disease," in *Proc. ITG conference on Speech Communication*, Oldenburg, Germany, Oct. 2018.

[17] J. Wang, P. V. Kothalkar, B. Cao, and D. Heitzman, "Towards automatic detection of Amyotrophic Lateral Sclerosis from speech acoustic and articulatory samples," in *17th Annual Conference of the International Speech Communication Association*, Sept. 2016, pp. 1195–1199.

[18] A. Illa, D. Patel, B. K. Yamini, S. S. Meera, N. Shivashankar, P. Veeramani, S. Vengalii, K. Polavarapui, S. Nashi, A. Nalini, and P K. Ghosh, "Comparison of speech tasks for automatic classification of patients with Amyotrophic Lateral Sclerosis and healthy subjects," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Calgary, Canada, Apr. 2018, pp. 6014–6018.

[19] G. Weismer, R. Martin, R. D. Kent, and J. F. Kent, "Formant trajectory characteristics of males with Amyotrophic Lateral Sclerosis," *Journal of the Acoustical Society of America*, vol. 91, no. 2, pp. 1085–1098, Feb. 1992.

[20] G. S. Turner, K. Tjaden, and G. Weismer, "The influence of speaking rate on vowel space and speech intelligibility for individuals with Amyotrophic Lateral Sclerosis," *Journal of Speech and Hearing Research*, vol. 38, no. 5, pp. 1001–1013, Oct. 1995.

[21] W. Weibull, "A statistical distribution function of wide applicability," *Journal of Applied Mechanics*, vol. 18, pp. 293–297, Sept. 1951.

[22] R. Martin, "Speech enhancement using MMSE short time spectral estimation with Gamma distributed speech priors," in *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, USA, May 2002, pp. 253–256.

[23] T. Gerkmann and R. Martin, "Empirical distributions of DFT-domain speech coefficients based on estimated speech variances," in *Proc. International Workshop on Acoustic, Echo, and Noise Control*, Tel Aviv, Israel, Sept. 2010.

[24] I. Tashev and A. Acero, "Statistical modeling of the speech signal," in *Proc. International Workshop on Acoustic, Echo, and Noise Control*, Tel Aviv, Israel, Sept. 2010.

[25] R. C. Hendriks, R. Heusdens, and J. Jensen, "Log-spectral magnitude MMSE estimators under super-Gaussian densities," in *Proc. 10th Annual Conference of the International Speech Communication Association*, Brighton, United Kingdom, Sept. 2009, pp. 1319–1322.

[26] I. S. Gradshteyn and I. M. Ryzhik, *Table of integrals, series, and products*, Academic Press, San Diego, CA, USA, 6 edition, 2000.

[27] M. Fournet, S. C. Chiuvé, S. Momjian, P. R. Burkhard, and M. Laganaro, "Modulation of voice and speech parameters after subthalamic nucleus deep brain stimulation in Parkinson's disease," *Brain Stimulation*, vol. 10, no. 2, pp. 470, Mar. 2017.

[28] C. Fougeron, V. Delvaux, L. Ménard, and M. Laganaro, "The MonPaGe_HA database for the documentation of spoken French throughout adulthood," in *Proc. 11th International Conference on Language Resources and Evaluation*, Miyazaki, Japan, May 2018, pp. 4301–4306.

[29] S. S. Shapiro and M. B. Wilk, "An analysis of variance test for normality (complete samples)," *Biometrika*, vol. 52, no. 3/4, pp. 591–611, Dec. 1965.

[30] F. Eyben, F. Weninger, F. Gross, and B. Schuller, "Recent developments in openSMILE, the Munich open-source multimedia feature extractor," in *Proc. ACM Multimedia (MM)*, Barcelona, Spain, Oct. 2018, pp. 835–838.