Automatic Detection of Parkinson's Disease from Words Uttered in Three Different Languages

J.R. Orozco-Arroyave[1,2], F. Hönig[2], J.D. Arias-Londoño[1], J.F. Vargas-Bonilla[1], S. Skodda[3]
J. Rusz[4,5], E. Nöth[2,6]

[1]Faculty of Engineering, Universidad de Antioquia, Medellń, Colombia
[2]Pattern Recognition Lab, Universität Erlangen-Nürnberg, Germany
[3]Department of Neurology, Knappschaftskrankenhaus, Ruhr-University Bochum, Germany
[4]Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic
[5]First Faculty of Medicine, Charles University, Prague, Czech Republic
[6]King Abdulaziz University, Jeddah, Saudi Arabia
Corresponding author: juan.rafael.orozco@fau.de

Abstract

About 90% of the people with Parkinson's disease (PD) develop speech impairments such as monopitch, monoloudness, imprecise articulation, and other symptoms. There are several studies addressing the problem of the automatic detection of PD from speech signals in order to develop computer aided tools for the assessment and monitoring of the patients. Recent works have shown that it is possible to detect PD from speech with accuracies above 90%; however, it is still unclear whether it is possible to make the detection independent of the spoken language. This paper addresses the automatic detection of PD considering speech recordings of three languages: German, Spanish and Czech. According to the results it is possible to classify between speech of people with PD and healthy controls (HC) with accuracies ranging from 84% to 99%, depending on the utterance.
Index Terms: Parkinson's disease, dysarthria, speech disorder, language, acoustic measures, voiced/unvoiced segments.

1. Introduction

PD is a neurological disorder characterized by the progressive loss of dopaminergic neurons in the substantia nigra pars compacta [1]. This disease affects over 1% of the people older than 65 [2], and there are more than one million of people with PD only in North America [3]. According to previous studies, about 90% of the people with PD develop vocal impairments [4] such as reduced loudness, monopitch, monoloudness, reduced stress, breathy, hoarse voice quality, imprecise articulation, among others [5]. The research community has been interested in the development of computer aided systems to perform automatic diagnosis and assessment of speech of people with PD [6, 7, 8].

In [9] the authors perform the automatic classification of 23 people with PD and 8 HC considering recordings of the English vowel /ah/ uttered in a sustained manner. The used features include different noise measures and several nonlinear dynamics features; the reported accuracy is 91%. In [10] an analysis of speech of 38 patients with PD and 17 HC is presented. All of the participants were English native speakers and they were asked to repeat three sentences several times per day. The English vowels /ah/, /i/ and /u/ were extracted from the recordings in order to measure different spectral features such as Vowel Space Area (VSA), Formant Centralization Ratio (FCR), the natural logarithm of VSA and the quotient $\frac{F2_i}{F2_u}$, where $F2_i$ and $F2_u$ are the values of second formant for the vowels /i/ and /u/, respectively. According to their results, FCR and the quotient can effectively and robustly differentiate between the speech of patients with PD and HC. Additionally, considering that the impact of PD in speech not only affects phonation, but also other characteristics of speech [11], in the recent years, the number of publications considering features related with phonation, articulation or prosody has increased. In [12] the authors measure several phonation and prosody features on four sentences uttered by 138 patients with PD and 50 age-matched HC (German native speakers). The results indicate that there is a correlation between PD symptoms and prosody variables such as the number of pauses in the speech. The phonation analysis is included in that paper considering pitch-related features, and the main finding is that the variation of the pitch is lower in the patients than in HC. Further, in [13] the authors evaluated a total of 46 Czech speakers (23 with PD and 23 HC). The work included features from phonation, articulation and prosody evaluated on 8 tasks including sustained phonations, diadochokinetic (DDK) evaluation (rapid repetition of the syllables /pa/-/ta/-/ka/), reading sentences, reading paragraph and a monologue. The authors concluded that 78% of the patients evidenced speech problems, and prosody was the most affected aspect even in the initial stage of the disease. They also found that the variation of the fundamental frequency contains very useful information for the separation of HC from PD speakers. Recently, in [7] the authors perform the automatic evaluation of speech of 168 patients with PD in different stages (English native speakers). The recordings include three different tasks: sustained phonation of the vowel /ah/, DDK evaluation and reading text. According to the results, the features extracted from the reading text and DDK tasks are most effective to perform the evaluation of the extent of the disease. The authors reported a mean absolute error of 5.5 explaining the 61% of the variance. In [14] considered a set of 176 German speakers (88 patients with PD and 88 HC) and performed the automatic detection of PD including acoustic, prosodic and voice related features. The accuracy reported in this work is 94%, considering a total of 1582 features.

Most of the recent works found in the literature consider speech recordings of different tasks such as sustained phonation of vowels, repetition of syllables, reading texts and monologues; however, there are other concerns that have not been addressed, i.e. (1) whether is possible to find features and/or

14 – 18 September 2014, Singapore

methods robust enough to perform automatic detection of PD in different languages or (2) whether those features and/or methods are robust enough to characterize speech samples recorded with different technical conditions.

This paper addresses the automatic detection of PD considering speech recordings of three languages: German, Spanish and Czech. The speech samples recorded in German are the same used in [14], Czech data is the same as in [13] and for Spanish, the recordings were collected in 2012 in the hospital "Clínica Noel", in Medellín, Colombia (Further details in [15]). This work includes the evaluation of isolated words uttered in Spanish and German and the rapid repetition of the syllables /pa/-/ta/-/ka/ uttered in the three languages. The method presented in this paper is based on the systematic separation of voiced and unvoiced segments before addressing the characterization and classification processes as in [16].

The results of the experiments presented here suggest that it is possible to classify between speech of people with PD and HC with an accuracy of up to 99%, depending on the evaluated utterance. To the best of our knowledge, this is the first work devoted to the automatic detection of PD considering speech recordings of different languages.

The rest of the paper is organized as follows: Section 2 provides details of the three databases considered in the work. Section 3 includes details of the experiments. Section 4 shows the results obtained in the different experiments and finally, in Section 5 the conclusions derived from this work are provided.

2. Data

2.1. Spanish

This database contains speech recordings of 50 patients with PD and 50 HC sampled at 44.1 KHz with a resolution of 16 bits. All of the speakers are balanced by gender and age i.e. the age of the 25 male patients ranges from 33 to 77 (mean 62.2 ± 11.2) and the age of the 25 female patients ranges from 44 to 75 (mean 60.1 ± 7.8). For the case of HC, the age of the 25 men ranges from 31 to 86 (mean 61.2 ± 11.3) and the age of the 25 women ranges from 43 to 76 (mean 60.7 ± 7.7). All of the patients were diagnosed by neurologist experts; the values of their evaluation according to the UPDRS-III [17] and Hoehn & Yahr [18] scales are 36.7 ± 18.7 and 2.3 ± 0.8, respectively. The speech samples were recorded with the patients in ON-state, i.e. no more than 3 hours after the morning medication. None of the people in the HC group has a history of symptoms related to Parkinson's disease or any other kind of neurological disorder. This database was collected by *Universidad de Antioquia* in Medellín, Colombia. Further details of this database in [15].

2.2. German

This corpus consists of 178 German native speakers. The set of patients includes 88 persons (44 men and 44 women) with age ranging from 42 to 84 (mean 66.5 ± 8.9). The HC group contains 88 speakers (45 men, 43 women) and their age ranges from 26 to 85 (mean 63.2 ± 13.9). The values of their neurological evaluation according to the UPDRS-III and Hoehn & Yahr scales are 22.7 ± 10.9 and 2.4 ± 0.6, respectively. In this case the speech samples were also recorded with the patients in ON-state. The speech data were recorded in the hospital of Bochum, Germany. The voice signals were sampled at 16 KHz with a resolution of 16 bits. Further details of this database can be found in [12].

2.3. Czech

A total of 42 Czech native speakers were recorded, 21 of the participants were diagnosed with idiopathic PD and their age ranges from 37 to 83 (mean 62.2 ± 11.00). The remaining 21 participants were healthy and their age ranges from 36 to 80 (mean 57.2 ± 13.01). The mean values of their evaluation according to the UPDRS-III and Hoehn & Yahr scales are 17.9 $\pm$ 7.4 and 2.2 ± 0.5, respectively. None of the patients had been medicated at the recording session. The speech data was recorded in the General University Hospital in Prague, Czech Republic. The voice signals were sampled at 48 KHz with a resolution of 16 bits. Further details of this database in [13].

3. Methods

The databases considered in this paper contain recordings of sustained phonation of vowels, DDK evaluation, sentences and monologues. Only the Spanish and German data include also isolated words, thus the experiments performed here include the evaluation of isolated words uttered in these two languages. The DDK evaluation include utterances of the three languages and it is performed with the rapid repetition of the syllables /pa/-/ta/-/ka/. The process will be described in detail below.

3.1. Preprocessing

The recordings are segmented into *voiced* and *unvoiced* frames using the software Praat [19]. Voiced frames are the portion of the recording where Praat detected pitch and *unvoiced* frames are those where the software does not detect pitch. Silences in the borders of the utterances are previously removed manually.

The voiced and unvoiced segments are grouped separately and a set of measures is applied on each group. The first step is to verify if the length of each frame is enough to form at least one window of 40ms. If the segment is not long enough it is discarded, otherwise the center of the segment is located and the windowing process continues from the center to the boarders (to both sides) using windows of 40ms with 20ms of overlapping. Once the windowing process is done, different features are calculated on each group of segments. This characterization is described in the next subsections.

3.2. Characterization of Voiced Frames

Three measurements of the noise content of the voiced frames in the phonation process are taken: Harmonics to Noise Ratio (HNR) [20], Normalized Noise Energy (NNE) [21] and Glottal to Noise Excitation Ratio (GNE) [22]. Additionally, the first and the second formant in Hertz (F1 and F2) and twelve Mel-Frequency Cepstral Coefficients (MFCC) are extracted.

The features of each voiced frame are grouped into one feature vector and four statistics of these vectors are calculated: mean value (m), standard deviation (std), kurtosis (k) and skewness (sk). As in total 17 measures are taken on the voiced segments, the feature set contains 68 measures per recording.

3.3. Characterization of Unvoiced Frames

The set of features calculated for unvoiced windows includes twelve MFCC and the energy of the signal distributed in 25 Bark bands. The energy in Bark bands is not calculated for the voiced frames because in previous experiments we have noticed that the energy calculated in critical bands do not provide information additional to which is provided by the MFCC to discriminate between PD and HC speakers [23].

The process to calculate the energy in these critical bands is introduced in [24]; the frequencies of the spectrum are distributed according to the Bark scale and the energy on each band is calculated to be included in the feature set. The feature vector is formed with four statistics of each measure for a total of $(12\text{MFCC}+25\text{Bark}) \cdot 4 = 148$ features. The segmentation and characterization processes described above are depicted in figure 1.

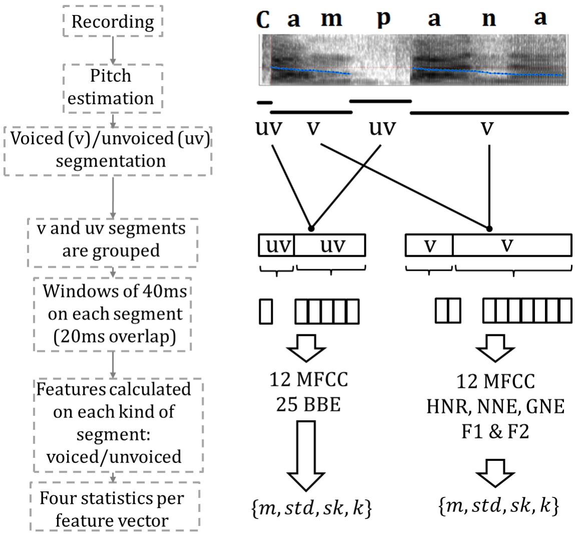

Figure 1: Segmentation and characterization processes

3.4. Classification

Classification process is performed with a soft margin support vector machine (SVM), with margin parameter C and a Gaussian kernel with parameter γ. The parameters C and γ are optimized up to powers of ten (grid-search with $1 < C < 10^4$ and $1 < \gamma < 10^3$, selection criteria: accuracy on test data). This will lead to slightly optimistic accuracy estimates but as there are only two parameters in the optimization process, the bias effect is minimal. The SVM is tested following a $10-$fold cross-validation strategy for the cases of the German and Spanish databases. In the case of the Czech database, it is tested following a leave-one-speaker-out strategy, as in [25].

It is important to note that even though the folds are taken randomly for German and Spanish data, on each fold the balance of age and gender of the speakers is guaranteed during the processes of training and testing. An SVM is used here due to its validated success in similar works related to automatic detection of pathological speech signals [26, 27, 28].

4. Results and Discussion

A total of 13 Spanish words were evaluated in this work. The two criteria for selecting those words were (1) they were uttered by all of the speakers and (2) the proportion of voiced and unvoiced frames was not less than 10%. For the case of German speakers, a total of six words were included. This work did not consider the evaluation of isolated words uttered in Czech.

The results obtained in the classification between PD and HC speakers considering the set of Spanish words are shown in Table 1 and the results with the German words are in Table 2. The rows indicate the results obtained with each word, separately for voiced (v) and unvoiced (uv) frames. The results are presented in terms of accuracy (Acc), specificity (Spec), sensitivity (Sens) and the area under the receiver operating curve (AUC). AUC is included since it is commonly used to evaluate the performance of different medical decision systems [29].

In Table 1 it can be noticed that in 8 of the 13 words the accuracy obtained with the unvoiced frames is greater than 95%. The highest accuracy with unvoiced frames is obtained with the word *atleta*. For the case of voiced frames, the highest accuracy is obtained with *petaka*. The percentage of unvoiced frames on every word was also calculated (not included in this paper) and we found that it is larger for healthy speakers than for PD patients. The exception is *campana* and this is the only word among the set of Spanish words included here that contains nasal sounds. It suggests that further research should be made in this regard. In Table 2 it can be observed that in 4 of the 7 evaluated words, the accuracy is higher than 90% when the features calculated on the unvoiced frames are considered. For the case of the voiced frames, the highest accuracy is 73% which is obtained with the word *Perlenkettenschachtel*.

Table 1: Results with Spanish words

		Acc (%)	**Sens (%)**	**Spec (%)**	**AUC**
atleta	uv	99	98	100	0.99
	v	82	86	78	0.79
campana	uv	99	98	100	0.99
	v	73	86	60	0.76
gato	uv	98	98	98	0.98
	v	76	84	68	0.76
petaka	uv	97	96	98	0.98
	v	84	88	80	0.82
braso	uv	96	100	92	0.98
	v	75	86	64	0.74
caucho	uv	96	92	100	0.95
	v	80	86	74	0.83
presa	uv	95	92	98	0.94
	v	81	80	82	0.81
apto	uv	95	98	92	0.95
	v	78	80	76	0.78
flecha	uv	94	98	90	0.93
	v	76	76	76	0.85
trato	uv	94	100	88	0.95
	v	77	90	64	0.83
coco	uv	93	98	88	0.94
	v	76	74	78	0.68
plato	uv	88	92	84	0.92
	v	69	74	64	0.64
pato	uv	84	90	78	0.83
	v	76	86	66	0.75

Figure 2 shows the ROC curves associated with the words *atleta* and *Bahnhofsvorsteher* which were those with the highest accuracies. On each word the difference of the performance obtained with unvoiced and voiced frames can be observed .

The results obtained with the recordings of the DDK evaluation performed in Spanish, German and Czech are presented in Table 3. Note that the accuracy is higher than 97% in all of the cases when the unvoiced segments are evaluated. For the case of the voiced segments, the highest accuracy is 90% which is obtained with Czech data. The good results obtained for three different databases with utterances of three different languages, recorded in different technical conditions, demonstrate the va-

Table 2: Results with German words

		Acc (%)	Sens (%)	Spec (%)	AUC
Bahnhofsvorsteher	uv	96	95	97	0.97
	v	72	66	77	0.72
Rettungsschwimmer	uv	95	93	98	0.96
	v	68	58	78	0.66
Toilettenpapier	uv	94	97	91	0.95
	v	70	69	72	0.70
Bundesgerichtshof	uv	93	94	93	0.95
	v	61	33	90	0.65
Bedienungsanleitung	uv	89	80	98	0.92
	v	67	65	69	0.61
Perlenkettenschachtel	uv	84	89	78	0.85
	v	73	66	79	0.70

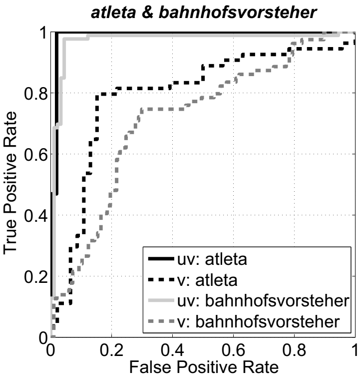

Figure 2: ROC curves of the best results obtained with German and Spanish words

lidity of the method proposed in this paper.

Table 3: Results with the rapid repetition of /pa/-/ta/-/ka/

		Acc (%)	Sens (%)	Spec (%)	AUC
Spanish	uv	99	100	98	0.99
	v	80	90	70	0.82
German	uv	97	98	96	0.98
	v	69	61	77	0.68
Czech	uv	97	100	95	0.99
	v	90	100	80	0.94

The difference in the performance obtained with the voiced and unvoiced frames of each language can be observed in Figures 4 and 3, respectively. Note that the performance obtained with the voiced frames of the Czech data is the highest among all of the results with voiced segments.

5. Conclusions

A method for the automatic classification of speech of people with PD and HC is presented in this paper. The method is based on the systematic separation of voiced and unvoiced segments before addressing the characterization and classification processes. The method is evaluated on three different databases, with speech samples of three different languages recorded in different technical conditions.

According to the results obtained with isolated words, in Spanish it is possible to achieve accuracies from 84% to 99%, while for German words the accuracy ranged from 84% to 96%.

The results obtained with /pa/-/ta/-/ka/ show that it is possible to achieve accuracies ranging from 97.6% to 99% with recordings of three different languages.

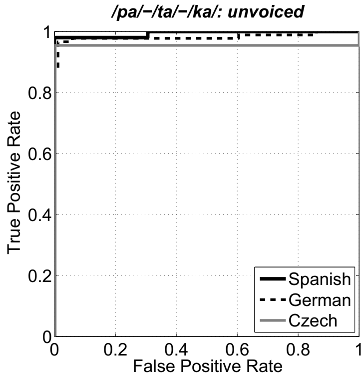

Figure 3: ROC curves obtained with the unvoiced frames of /pa/-/ta/-/ka/ uttered in Spanish, German and Czech

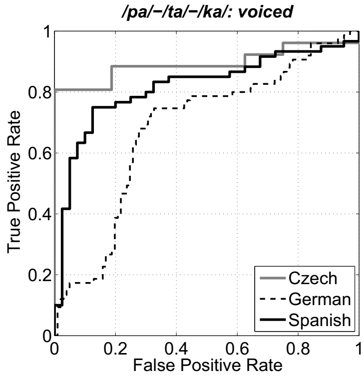

Figure 4: ROC curves obtained with the voiced frames of /pa/-/ta/-/ka/ uttered in Spanish, German and Czech

The results obtained both in isolated words and in the DDK evaluation demonstrate the validity of the method proposed in this paper. To the best of our knowledge, this is the first work that addresses the problem of automatic classification of speech of people with PD and HC in different languages. The results obtained suggest that it is possible to detect PD from speech using the same method in different languages, they also motivate us to perform cross-language experiments, i.e. to train and test the system with utterances of different languages. Future works with different databases might consider that from the neurologist's point of view, dysarthria in PD has very characteristic patterns (monopitch, monoloudness, etc.). However, since PD is a chronic progressive disorder, the degree and individual appearance of voice and speech impairment can differ considerably between patients. Additionally, further study is required considering the variability among the patients of different databases.

6. Acknowledgements

Juan Rafael Orozco Arroyave is under grants of COLCIENCIAS through the call Nº 528"generación del bicentenario 2011". Jan Rusz is supported by the Czech Science Foundation (GACR 102/12/2230). This work was also financed by COLCIENCIAS through the project Nº 111556933858.

7. References

[1] O. Hornykiewicz, "Biochemical aspects of parkinson's disease," *Neurology*, vol. 51, no. 2, pp. S2–S9, 1998.

[2] M. de Rijk, "Prevalence of parkinson's disease in europe: A collaborative study of population-based cohorts," *Neurology*, vol. 54, pp. 21–23, 2000.

[3] A. Lang and A. Lozano, "Parkinson's disease: First of two parts," *New England Journal of Medicine*, vol. 339, pp. 1044–1053, 1998.

[4] A. Ho, R. Jansek, C. Marigliani, J. Bradshaw, and S. Gates, "Speech impairment in a large sample of people with parkinson's disease," *Behavioral Neurology*, vol. 11, pp. 131–137, 1998.

[5] L. Ramig, C. Fox, and S. Sapir, "Speech treatment for parkinson's disease," *Expert Review Neurotherapeutics*, vol. 8, no. 2, pp. 297–309, 2008.

[6] A. Tsanas, M. Little, C. Fox, and L. Ramig, "Objective Automatic Assessment of Rehabilitative Speech Treatment in Parkinson's Disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 22, no. 1, pp. 181–190, 2014.

[7] A. Bayestehtashk, M. Asgari, I. Shafran, and J. Mcnames, "Fully automated assessment of the severity of parkinson's disease from speech," *Computer Speech & Language, 2014*, pp. 1–14, to appear.

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

[9] M. A. Little, P. McSharry, E. Hunter, J. Spielman, and L. Ramig, "Suitability of dysphonia measurements for telemonitoring of parkinson's disease," *IEEE transactions on bio-medical engineering*, vol. 56, no. 4, pp. 1015–1022, 2009.

[10] S. Sapir, , L. Raming, J. Spielman, and C. Fox, "Formant centralization ratio (FCR): A proposal for a new acoustic measure of dysarthric speech," *Journal of Speech Language and Hearing Research*, vol. 53, no. 1, pp. 1–20, 2010.

[11] J. Rusz, R. Cmejla, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, and E. Ruzicka, "Acoustic assessment of voice and speech disorders in parkinson's disease through quick vocal test," *Movement Disorders*, vol. 26, no. 10, pp. 1951–1952, 2011.

[12] S. Skodda, W. Grnheit, and U. Schlegel, "Intonation and speech rate in parkinson's disease: General and dynamic aspects and responsiveness to levodopa admission," *Journal of Voice*, vol. 25, no. 4, pp. 199–205, 2011.

[13] J. Rusz, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated parkinson's disease," *The Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.

[14] T. Bocklet, S. Steidl, E. Nöth, and S. Skodda, "Automatic evaluation of parkinson's speech - acoustic, prosodic and voice related cues," in *Proceedings of the 14th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2013, pp. 1149–1153.

[15] J. Orozco-Arroyave, J. Arias-Londoño, J. Vargas-Bonilla, M. González-Rátiva, and E. Nöth, "New spanish speech corpus database for the analysis of people suffering from parkinson's disease," in *Proceedings of the 9th Language Resources and Evaluation Conference (LREC), 2014*, to appear.

[16] C. Clavel, I. Vasilescu, L. Devillers, G. Richard, and T. Ehrette, "Fear-type emotion recognition for future audio-based surveillance systems," *Speech Communication*, vol. 50, no. 6, pp. 487–503, 2008.

[17] G. Stebbing and C. Goetz, "Factor structure of the unified parkinsons disease rating scale: Motor examination section," *Movement Disorders*, vol. 13, pp. 633–636, 1998.

[18] M. M. Hoehn and M. D. Yahr, "Parkinsonism: onset, progression, and mortality," *Neurology*, vol. 17, pp. 427–442, 1967.

[19] P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer," *Glot Ing.*, vol. 5, pp. 341–345, 2001.

[20] E. Yumoto, W. Gould, and T. Baer, "Harmonics-to-noise ratio as an index of the degree of hoarseness," *Journal of the Acoustical Society of America*, vol. 71, no. 6, pp. 1544–1550, 1982.

[21] H. Kasuya, S. Ogawa, K. Mashima, and S. Ebihara, "Normalized noise energy as an acoustic measure to evaluate pathologic voice," *Journal of the Acoustical Society of America*, vol. 80, no. 5, pp. 1329–1334, 1986.

[22] D. Michaelis, T. Gramss, and H. Strube, "Glottal-to-noise excitation ratio-a new measure for describing pathological voices," *Acustica/ActaAcustica*, vol. 83, pp. 700–706, 1997.

[23] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, and E. Nöth, "Perceptual analysis of speech signals from people with parkinson's disease," *Lecture Notes in Computer Science*, vol. 7930, pp. 201–211, 2013.

[24] T. Jehan, "Creating music by listening," Ph.D. dissertation, Massachusets Institute of Technology, September 2005.

[25] T. Bocklet, E. Nöth, G. Stemmer, H. Ruzickova, and J. Rusz, "Detection of persons with parkinson's disease by acoustic, vocal and prosodic analysis," in *Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, 2011, pp. 478–483.

[26] A. Marier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster, and E. Nöth, "PEAKS - A system for the automatic evaluation of voice and speech disorders," *Speech Communications*, vol. 51, no. 5, pp. 425–437, 2009.

[27] J. Arias-Londoño, J. Godino-Llorente, N. Sáenz-Lechón, V. Osma-Ruiz, and G. Castellanos-Domínguez, "Automatic detection of pathological voices using complexity measures, noise parameters, and mel-cepstral coefficients." *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 2, pp. 370–9, 2011.

[28] J. R. Orozco-Arroyave, J. F. Vargas-Bonilla, J. D. Arias-Londoño, S. Murillo-Rendón, G. Castellanos-Domínguez, and J. Garcés, "Nonlinear dynamics for hypernasality detection in spanish vowels and words," *Cognitive Computation*, vol. 5, no. 4, pp. 448–457, 2013.

[29] N. Sáenz-Lechón, J. Godino-Llorente, V. Osma-Ruiz, and P. Gómez-Vilda, "Methodological issues in the development of automatic systems for voice pathology detection," *Biomedical Signal Processing and Control*, vol. 1, pp. 120–128, 2006.