

# Voiced/unvoiced transitions in speech as a potential bio-marker to detect Parkinson's disease

J.R. Orozco-Arroyave<sup>1,2</sup>, F. Hönig<sup>2</sup>, J.D. Arias-Londoño<sup>1</sup>, J.F. Vargas-Bonilla<sup>1</sup>, S. Skodda<sup>3</sup> J. Rusz<sup>4,5</sup>, E. Nöth<sup>2</sup>

<sup>1</sup>Faculty of Engineering, Universidad de Antioquia, Medellín, Colombia
<sup>2</sup>Pattern Recognition Lab, Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany
<sup>3</sup>Department of Neurology, Knappschaftskrankenhaus, Ruhr-University Bochum, Germany
<sup>4</sup>Faculty of Electrical Engineering, Czech Technical University, Prague, Czech Republic
<sup>5</sup>First Faculty of Medicine, Charles University, Prague, Czech Republic

Corresponding author: rafael.orozco@i5.informatik.uni-erlangen.de

# Abstract

Several studies have addressed the automatic classification of speakers with Parkinson's disease (PD) and healthy controls (HC). Most of the studies are based on speech recordings of sustained vowels, isolated words, and single sentences. Only few investigations have considered read texts and/or spontaneous speech. This paper addresses two main questions still open regarding the automatic analysis speech in patients with PD, (a) "Is it possible to classify PD patients and HC through running speech signals in multiple languages?", and (b) "where is the information to discriminate between speech recordings of PD patients and HC?" In this paper speech recordings of read texts and monologues spoken in three different languages are considered. The energy content of the borders between voiced and unvoiced sounds is modeled. According to the results with read texts it is possible to achieve accuracies ranging from 91%to 98% depending on the language. With respect to the results on monologues, the accuracies are above 98% in all of the three languages. The presence of discriminant information in the voiced/unvoiced and unvoiced/voiced transitions is validated here, evidencing the problems of PD patients to stop/start the vocal folds movement during the production of running speech.

**Índex Terms**: Parkinson's disease, dysarthria, hesitation in speech, language and motor planning, energy content, voiced/unvoiced transitions.

## 1. Introduction

PD is a neurological disorder characterized by the progressive loss of dopaminergic neurons in the substantia nigra of the midbrain [1]. PD patients develop motor deficits such as bradykinesia, rigidity, postural instability, and resting tremor. Non-motor deficits can also be present including negative effects on sleep, cognition, and emotion [2]. PD affects about 2% of people older than 65 [3]. It is known that about 90% of people with PD develop different speech impairments including reduced loudness, a monopitch and monoloudness kind of speech, breathy voice, and imprecise articulation, among others. All of these symptoms can be grouped and called *hypokinetic dysarthria* [2]. Previous studies have shown that speech impairments are one of the earliest manifestations of the disease [4,5]. The most prevalent symptoms of Parkinsonian speech are reduced loudness, prosody impairments, and imprecise articulation [6].

Different studies in the literature have addressed the problem of the automatic detection of PD through speech signals. Phonation characteristics of PD patients are studied in [7]. The authors consider a total of 43 speakers, 33 with PD and 10 HC. The sustained phonation of the vowel /a/ was repeated about six times per participant, comprising a total of 263 speech samples. Different voice characteristics were measured, including jitter, shimmer, harmonics-to-noise ratio (HNR), Glottal to noise excitation ratio (GNE), mel-Frequency Cepstral Coefficients (MFCCs), and nonlinear measures, among others. The authors apply feature selection techniques and two different classification strategies, random forest and support vector machine with Gaussian kernel. Tests are performed through a 10fold cross-validation (CV) strategy. Each fold was formed randomly, which means that the speaker-independence was not fulfilled in the experiments. The authors report accuracies ranging from 94.4% to 98.6%. In [8] articulation features of speech in patients with PD are analyzed. The authors intended to reveal correlations among vowel articulation, global motor performance, and the stage of the disease. A total of 100 native German speakers (68 PD patients and 32 HC) were asked to read a text. The first two formants (F1 and F2) were extracted from the vowels /a/, /i/, and /u/ in the text. The articulation analysis was based on the triangular vowel space area (tVSA) and the vowel articulation index (VAI). The authors conclude that VAI is lower for PD patients than for HC, while tVSA was significantly reduced only in PD female patients. No correlations were found between vowel articulation and the severity of motor symptoms of PD. Prosodic information is analyzed in [9], where several intonation-related features are measured on four sentences uttered by 138 PD patients and 50 HC (native German speakers). The estimated features are based on the fundamental frequency of speech  $(F_0)$ . The speech rate is also analyzed through the length of the spoken syllables and pauses. Net speech rate (NSR) was measured in syllables per second related to the total speech duration in milliseconds. Further, the authors introduced the concept of articulatory acceleration as the difference between the NSR of the first and the fourth sentences. According to the results, the variation of F<sub>0</sub> is lower in PD patients than in HC. The authors also observed correlations among several symptoms of PD and prosody variables such as the number of pauses in speech. In [10] the vowel articulation of a set with 35 native Czech speakers (20 with PD and 15 HC) is considered. A total of four speech tasks are evaluated including sustained vowels, sentences repetition, read text, and monologue. The set of measures includes  $F_1$ ,  $F_2$ , VAI, tVSA, and the quotient  $\frac{F_{2i}}{F_{2u}}$ , where  $F_{2i}$  and  $F_{2u}$  are the second formant measured from the the vowels /i/ and /u/, respectively. The authors report classification scores of about 80% separating speech recordings of PD patients and HC. In [11] different articulatory deficits in the speech of PD patients are modeled. The study comprises a set with six articulatory aspects including vowel quality, coordination of laryngeal and supra-laryngeal activity, precision of consonant articulation, tongue movement, occlusion weakening, and speech timing. The authors report that, in accordance with previous studies in the literature, imprecise articulation is one of the most predominant characteristic of PD speech [12]. The accuracy obtained in separating speech signals of PD patients and HC using the aforementioned features is 88%.

In a previous work [13] we performed the automatic classification speech signals of people with PD and HC. Speech recordings of isolated words and the rapid repetition of the syllables /pa-ta-ka/ were evaluated in three different languages, Spanish, German, and Czech. The results indicated that there is discriminant information in the unvoiced sounds, with accuracies ranging from 84% to 99%. These results are in accordance with previous observations performed by neurologists and clinicians. Speakers with PD have been found to have problems to pronounce consonants that require pressure build-up in the mouth and lingual movements, respectively [14]. The most serious pronunciation problems have been observed in plosive sounds such as /p/, /t/, /k/, /b/, and /d/ which require the movement of several articulatory organs like the lips, tongue, jaw, epiglottis, and larynx [15]. The production of plosive sounds and unvoiced consonants require to stop the vibration of vocal folds if the previous sound is voiced. In the same way, if there is a voiced sound after an unvoiced consonant, the vocal folds vibration has to start. This pattern has been observed in the gait of PD patients because the disease generates difficulties starting and stopping the movements of the patients [16]. As speech production is a motor task that involves the movement of several muscles and limbs, such difficulties should be also observed in speech recordings. In this paper we address different experiments to model the transitions between voiced and unvoiced frames extracted from running speech signals. Recordings of monologues and read texts spoken in three different languages are considered.

The rest of the paper is organized as follows: Section 2 describes the three databases considered in this study. Section 3 illustrates the stages of the proposed methodology and the addressed experiments. Section 4 shows the results obtained and provides a brief discussion. Finally, Section 5 includes the conclusions derived from this work.

# 2. Corpus of speakers

#### 2.1. Spanish recordings

The PC-GITA database [17] is considered here. It contains speech samples of 100 speakers (50 PD patients and 50 HC balanced in age and gender). All of the participants were recorded in a sound-proof booth at Clínica Noel of Medellín in Colombia, using a dynamic omni-directional microphone and professional audio card. The sampling frequency of the recordings was 44.1 KHz with 16 bits of resolution. The speakers were asked to read a text with 36 words. For the monologues, the speakers were asked to talk about their daily activities. The average duration of the monologues is  $48 \pm 29$  seconds and

 $45 \pm 24$  seconds for PD and HC groups, respectively. All of the patients were recorded in ON-state, i.e. no more than three hours after their morning medication. The patients were diagnosed by a neurologist expert. The mean values of their evaluation according to the MDS-UPDRS-III and Hoehn & Yahr [18] scales are  $36.7 \pm 18.7$  and  $2.3 \pm 0.8$ , respectively. Further details of this database can be found in [17].

#### 2.2. German recordings

This corpus consists of 170 native German speakers (85 PD patients and 85 HC). The participants were recorded in a quiet room at the Knappschaftskrankenhaus of Bochum in Germany. The sampling frequency of the recordings was 16 KHz with a resolution of 16 bits. The speakers read a text with 81 words and talked about their daily activities. The average duration of the monologues is  $33 \pm 8$  seconds and  $28 \pm 6$  seconds for PD and HC groups, respectively. The speech samples were also recorded with the patients in ON-state. The mean values of their neurological evaluation according to the UPDRS-III and Hoehn & Yahr scales are  $22.7 \pm 10.9$  and  $2.4 \pm 0.6$ , respectively. Further details of this database can be found in [9].

### 2.3. Czech

A total of 35 native Czech speakers were recorded (20 PD patients and 15 HC). All of the patients were newly diagnosed with PD, and none of them had been medicated before or during the recording session. The participants were recorded in the General University Hospital in Prague, Czech Republic. The signals were sampled at 48 KHz with a resolution of 16 bits. The participants read a text with 80 words and talked about their daily activities. The average duration of the monologues is  $115 \pm 56$  seconds and  $130 \pm 51$  seconds for PD and HC groups, respectively. The mean values of their evaluation according to the UPDRS-III and Hoehn & Yahr scales are  $17.9 \pm 7.4$  and  $2.2 \pm 0.5$ , respectively. Further details of this database can be found in [10].

# 3. Methods

**The hypothesis:** PD patients have difficulty to begin and to stop the vocal fold vibration, and such difficulty can be observed on speech signals by modeling the frequency content of the transitions between voiced and unvoiced sounds.

In order to validate this hypothesis, recordings of read texts and monologues spoken in three different languages are considered here. The proposed methodology comprises three main steps. (1) Preprocessing and automatic detection of the transitions between voiced and unvoiced frames, (2) characterization, and (3) automatic classification of PD and HC speech signals. The methodology is summarized in Figure 1. Each step is described with details in the following subsections.

#### 3.1. Preprocessing and segmentation

The recordings of the three databases are re-sampled to 16 KHz. Mean cepstral subtraction is performed to eliminate possible bias introduced by the channel i.e. microphone and sound card. Pitch values are estimated in Praat [19]. Then the transitions from voiced to unvoiced (v-uv), i.e. offset, and from unvoiced to voiced (uv-v), i.e. onset, are detected. 40 ms of the signal are taken to the left and to the right of each border, forming "chunks" of signals with 80 ms length. Frames shorter than 40 ms are excluded from the analysis. Hamming windows with

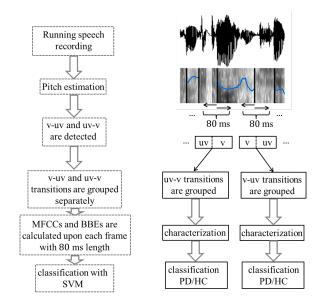


Figure 1: Methodology

20 ms length and time shift of 10 ms are applied to those chunks.

Table 1 indicates the percentage of onsets and offsets detected on the recordings of the three languages. These percentages are calculated as the ratio between the duration of the detected transitions and the total duration of the recordings. The frequency of occurrence (Freq.) of such onsets and offsets is given in "transitions per second". Note that the percentages and the frequency of the transition occurrences are in the same range per language; however, the values obtained with the recordings of the PD patients are slightly lower than those obtained with the healthy speakers. This behavior can be observed in both speech tasks per language.

Table 1: % and frequency of occurrence of the transitions.

	Read texts		Monologues	
	HC	PD	HC	PD
	Spanish			
onsets %	$13 \pm 3$	$11 \pm 4$	$13 \pm 3$	$12 \pm 4$
offsets %	$14 \pm 3$	$11 \pm 4$	$13 \pm 3$	$11 \pm 4$
Freq.	$1.7 \pm 0.4$	$1.4 \pm 0.5$	$1.6 \pm 0.4$	$1.4 \pm 0.5$
	German			
onsets %	$24 \pm 6$	$21 \pm 4$	$21 \pm 6$	$18 \pm 5$
offsets %	$24 \pm 6$	$21 \pm 4$	$21 \pm 6$	$18 \pm 5$
Freq.	$2.9 \pm 0.7$	$2.6 \pm 0.5$	$2.6 \pm 0.7$	$2.2 \pm 0.6$
	Czech			
onsets %	$21 \pm 3$	$20 \pm 3$	$20 \pm 3$	$18 \pm 3$
offsets %	$21 \pm 3$	$20 \pm 3$	$20 \pm 3$	$18 \pm 3$
Freq.	$2.6 \pm 0.4$	$2.5 \pm 0.4$	$2.4 \pm 0.3$	$2.3 \pm 0.4$

## 3.2. Characterization

The frequency content of each windowed segment is modeled with 12 MFCCs and 25 Bark band energies (BBEs) [20]. MFCCs are used because they provide a coarse representation of the short-time spectrum and have shown to be appropriate to model irregular movements in the vocal tract [21]. The features are grouped into one feature vector per speaker and four functionals are calculated: mean value, standard deviation, kurtosis, and skewness, forming a 148-dimensional feature vector  $(37 \times 4 = 148)$  per recording.

## 3.3. Classification

The classification process is performed using a radial basis SVM as in [13], with margin parameter C and Gaussian kernel with parameter  $\gamma$ . C and  $\gamma$  are optimized in a grid-search up to powers of ten with  $10^{-1} < C < 10^4$  and  $10^{-1} < \gamma < 10^3$ . The selection criterion was based on the accuracy on the test data, which can lead to slightly optimistic estimates but bias should be small due to the low number of parameters.

The validation strategy addressed here is the same as in [13]. For the Spanish and German datasets consisted on 10-fold cross validation. Each fold was randomly formed but assuring the balance in age and gender. Due to the low number of recordings, the validation for the Czech data was addressed with a leave-one-speaker-out cross-validation strategy. Results are presented in terms of accuracy (Acc.), specificity (Spec.), sensitivity (Sens.), and area under the receiver operating characteristic curve (AUC). These measures are commonly used to evaluate the performance of medical systems [22].

## 4. Results and Discussion

Tables 2 and 3 show the results obtained on read texts and monologues, respectively considering the uv-v transitions (onset). Note that with such transitions the difficulty of the PD patients to start moving the vocal folds (to start the vibration) is modeled. The results are higher on monologues than on read texts. It can be likely explained either because the richness of the monologues in terms of speech information content e.g. syllables, words, articulatory movements, and pronunciation or because the impaired motor-planning process of PD patients [23]. This impairment affects language production due to hesitations to initiate speech [24, 25]. Thus, it seems like the motor-planning process is more complex in the production of monologues than reading texts.

Table 2: Results with the onset transitions on read texts				
	Acc (%)	<b>Sens</b> (%)	<b>Spec</b> (%)	AUC
Spanish	$98 \pm 6$	$98 \pm 6$	$98 \pm 6$	0.98
German	$95\pm7$	$92 \pm 12$	$98 \pm 5$	0.96
Czech	$91\pm20$	$85\pm37$	$97\pm13$	0.91

Table 3: Results with the onset transitions on monologues				
	Acc (%)	<b>Sens</b> (%)	<b>Spec</b> (%)	AUC
Spanish	$98 \pm 4$	$96 \pm 8$	$100 \pm 0$	0.97
German	$99 \pm 2$	$99 \pm 1$	$99 \pm 4$	0.99
Czech	$99 \pm 5$	$100 \pm 0$	$97 \pm 11$	0.99

Tables 4 and 5 show results obtained with the v-uv transitions (offset) on read texts and monologues, respectively. The analysis of v-uv transitions allow to model difficulties of the patients to stop the vocal folds vibration. Although such difficulties to stop movements are less typical, they are also documented [16]. In this case the results on monologues are also higher than in read texts, which is consistent with the results presented in Tables 2, and 3.

Figures 2 and 3 show the ROC curves obtained with the uv-v and v-uv borders, respectively. ROC curves illustrate the results more compactly, and give a visual idea of the performance of the system [22]. In Figure 2, the low true positive rate of the uv-v characterization on texts read in German and Czech

Table 4: Results with the offset transitions on read texts

	Acc (%)	<b>Sens</b> (%)	<b>Spec</b> (%)	AUC
Spanish	$95 \pm 5$	$92 \pm 10$	$98 \pm 6$	0.97
German	$96 \pm 5$	$96 \pm 8$	$97 \pm 6$	0.97
Czech	$94\pm17$	$97\pm10$	$90 \pm 31$	0.94

Table 5: Results with the offset transitions on monologu	es
----------------------------------------------------------	----

	<b>Acc</b> (%)	<b>Sens</b> (%)	<b>Spec</b> (%)	AUC
Spanish	$99 \pm 3$	$98 \pm 6$	$100 \pm 0$	0.99
German	$99 \pm 2$	$99 \pm 4$	$100 \pm 0$	0.99
Czech	$97 \pm 12$	$100 \pm 0$	$94 \pm 24$	0.97

can be observed (Sens = 92% and 85%, respectively). On the other hand, Figure 3 shows the low false positive rate of the system when tested on Czech recordings with features of the v-uv borders (Spec = 90%).

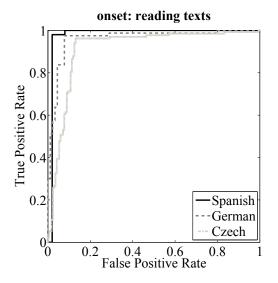


Figure 2: ROC curve obtained with the onset of read texts.

It is worthy to highlight that we made several experiments (not reported here due to space limitations) considering only the voiced segments of the recordings. Such frames were characterized using "classical" features including MFCCs, noise measures, jitter, and shimmer, among others. The results ranged from 75% to 90% depending on the speech task and the language [13]. These experiments confirm that the methodology presented in this paper is more accurate and robust than the classical approaches.

## 5. Conclusions

Two questions regarding the analysis of speech of people with Parkinson's disease are addressed in this paper. (a) "Is it possible to discriminate between PD patients and HC through running speech signals?", and (b) "where is the discriminant information to separate speech recordings of PD patients and HC?"

To address these two questions, read texts and monologues uttered by PD patients and their respective healthy controls are considered. The proposed method is tested on speech recordings of three languages, Spanish, German, and Czech. According to the results, the information to discriminate speech

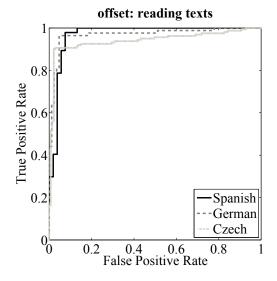


Figure 3: ROC curve obtained with the offset of read texts.

of people with PD and HC is located in the transition between voiced and unvoiced frames. Additionally, as there are two possible transitions, i.e. from unvoiced to voiced and from voiced to unvoiced, the difficulty of the patients to start and to stop the vibration of the vocal folds is modeled here. The presence of discriminant information in the v-uv/uv-v transitions has been shown in this paper. According to the results, with read texts it is possible achieve accuracies ranging from 91% to 98% depending on the language. With respect to the results with monologues, the accuracies are above 98% in all of the three languages. These results evidence the presence of highly discriminant information in the transitions. The high accuracies obtained with monologues can be likely explained due to two possible reasons (1) because the richness of monologues in terms of variety of words, syllables, and articulatory movements, among other or (2) because the motor-planning process to produce spontaneous speech is more complex and contributes to the appearance of articulation errors and disfluency in the speech of PD patients. Further experiments are required to validate the usefulness of such information to predict the neurological state of the patients, e.g. mapping the values of the UPDRS-III or Hoehn & Yahr scales. The results presented in this paper represent a step forward to a deeper understanding of the phenomena involved in the speech of people with PD. As the findings reported here show difficulties in the patients to start/stop the vocal fold movements, this result could be used to design new techniques for speech therapy, which focus on improving the ability to start/stop those movements. Finally, we are aware of the high accuracies reported in this paper. It can be explained not only because of the presence of discriminative information in the transitions, but also because for the datasets used here the classification task is "relatively easy".

# 6. Acknowledgments

Juan Rafael Orozco-Arroyave is under grants of COLCIEN-CIAS through the program "Convocatoria N<sup>o</sup> 528, generación del bicentenario 2011". This work was also financed by COL-CIENCIAS project N<sup>o</sup> 111556933858. The authors express thanks to CODI for its support through "estrategia de sostenibilidad 2014-2015 de la Universidad de Antioquia".

## 7. References

- O. Hornykiewicz, "Biochemical aspects of parkinson's disease," *Neurology*, vol. 51, no. 2, pp. S2–S9, 1998.
- [2] J. Logemann, H. Fisher, B. Boshes, and E. Blonsky, "Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of parkinson patients," *Journal of Speech and Hearing Disorders*, vol. 43, pp. 47–57, 1978.
- [3] M. de Rijk, "Prevalence of parkinson's disease in europe: A collaborative study of population-based cohorts," *Neurology*, vol. 54, pp. 21–23, 2000.
- [4] B. T. Harel, M. S. Cannizzaro, H. Cohen, N. Reilly, and P. J. Snyder, "Acoustic characteristics of Parkinsonian speech: a potential biomarker of early disease progression and treatment," *Journal of Neurolinguistics*, vol. 17, pp. 439–453, 2004.
- [5] 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," *Journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.
- [6] C. Stewart, L. Winfield, A. Hunt, S. B. Bressman, S. Fahn, A. Blitzer, and M. F. Brin, "Speech dysfunction in early Parkinson's disease," *Movement Disorders*, vol. 10, no. 5, pp. 562–565, 1995.
- [7] A. Tsanas, M. Little, P. Mcsharry, J. Spielman, and L. Ramig, "Novel Speech Signal Processing Algorithms for High-Accuracy Classification of Parkinson's Disease," *IEEE Transactions on Bio-Medical Engineering*, vol. 59, no. 5, pp. 1264–1271, 2012.
- [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] S. Skodda, W. Grönheit, 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.
- [10] 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," *The Journal of the Acoustical Society of America*, vol. 134, no. 3, pp. 2171–2181, 2013.
- [11] M. Novotný, J. Rusz, R. Cmejla, and E. Ruzicka, "Automatic Evaluation of Articulatory Disorders in Parkinsons Disease," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 9, pp. 1366–1378, 2014.
- [12] L. Ramig, C. Fox, and S. Sapir, "Speech treatment for parkinson's disease," *Expert Review Neurotherapeutics*, vol. 8, no. 2, pp. 297–309, 2008.
- [13] J. R. Orozco-Arroyave, F. Hönig, J. D. Arias-Londoño, J. F. Vargas-Bonilla, S. Skodda, J. Rusz, and E. Nöth, "Automatic detection of parkinson's disease from words uttered in three different languages," in *Proceedings of the 15th Annual Conference of the International Speech Communication Association (INTERSPEECH)*, 2014, pp. 1473–1577.

- [14] J. Duffy, "Motor Speech Disorders: Clues to Neurologic Diagnosis," in Parkinson's Disease and Movement Disorders: Diagnosis and Treatment Guidelines for the Practicing Physician, 2000, pp. 35–53.
- [15] J. Logemann and H. Fisher, "Vocal tract control in parkinson's disease: Phonetic feature analysis of misarticulations," *Journal of Speech and Hearing Disorders*, vol. 46, no. 4, pp. 348–452, 1981.
- [16] J. Wiese, Clinical Clerkships: The Answer Book, 1st ed. Baltimore, USA: Lippincott Williams & Wilkins, 2006.
- [17] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, M. C. 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, 2014, pp. 342–347.
- [18] C. Goetz, W. Poewe, O. Rascol, C. Sampaio, G. Stebbins, C. Counsell, N. Giladi, R. Holloway, C. Moore, G. Wenning, M. Yahr, and L. Seidl, "Movement Disorder Society Task Force report on the Hoehn and Yahr staging scale: status and recommendations," *Movement disorders : official journal of the Movement Disorder Society*, vol. 19, no. 9, pp. 1020–1028, 2004.
- [19] P. Boersma and D. Weenink, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, no. 9/10, pp. 341–345, 2001.
- [20] E. Zwicker and E. Terhardt, "Analytical expressions for critical-band rate and critical bandwidth as a function of frequency," *Journal of Acoustical Society of America*, vol. 68, no. 5, pp. 1523–1525, 1980.
- [21] J. Godino-Llorente, P. Gómez-Vilda, and M. Blanco-Velasco, "Dimensionality reduction of a pathological voice quality assessment system based on gaussian mixture models and short-term cepstral parameters," *IEEE Transactions on Bio-Medical Engineering*, vol. 53, no. 10, pp. 1943–1953, 2006.
- [22] 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.
- [23] M. W. Rogers and C. W. Chan, "Motor planning is impaired in Parkinson's disease," *Brain Research*, vol. 438, pp. 271–276, 1988.
- [24] J. Illes, "Neurolinguistic features of spontaneous language production dissociate three forms of neurodegenerative disease: Alzheimer's, Huntington's, and Parkinson's," *Brain and language*, vol. 37, pp. 628–642, 1989.
- [25] B. Walsh, A. Smith, and W. Lafayette, "Linguistic complexity, speech production, and comprehension in Parkinson's disease: behavioral and physiological indices," *Journal of Speech, Language, and Hearing Research*, vol. 54, no. June, pp. 787–803, 2011.