Towards an Objective Criteria for the Diagnosis of Parkinson Disease Based on Speech Assessment

J. Teixeira[1], L. Soares[1], M. Miguel[1], P. Martins[1], S. Reis[1], L. Coelho[1,2] and C. Lopes[3]

[1] Department of Physics, Instituto Superior Engenharia do Porto, Portugal, {1160(321/296/299/308), ssr, lfc}@isep.ipp.pt

[2] INEGI – Laboratory of Optics and Experimental Mechanics, Portugal

[3] LEMA and CEOS.PP, ISCAP – Politécnico do Porto, Portugal, cristinalopes@iscap.ipp.pt

Abstract

Parkinson's disease is a progressive neurodegenerative disease whose symptoms increase in severity with age. Loss of muscle control, with implications for speech, is one of the main complications associated with this disease. In this article we present an analysis of continuous speech signals in which the effects of this pathology are identified and characterized in the phonation process. The preliminary analysis is based on a stochastic and comparative baseline methodology, where records of informants with Parkinson's and asymptomatic informants are explored. The obtained results showed that in individuals with Parkinson's, the parameters Shimmer local and Shimmer apq11, suffer the most significant changes when compared with identical parameters in healthy individuals. The identification of these characteristics in Parkinson's patients, as well as the observed discriminatory capacity, may be the basis for a new non-invasive and low-cost methodology for the early diagnosis of this pathology.

1. Introduction

Neurodegenerative diseases, without a known cure, are characterized by the progressive dysfunction and atrophy of central nervous system structures, leading to loss of movement and, in other cases, to dementia [1]. Parkinson's disease (PD), the focus of your study, is the second most common neurodegenerative disease in the world and it is estimated that the number of patients continues to increase until year 2030 [2]. Incoordination in muscle control is one of the major problems arising from Parkinson's and therefore changes in speech fluency due to degenerative processes in associated neurons are often evident. The fact that degenerative processes are progressive and increase in severity with age, makes patients increasingly weak and in greater need of support for the performance of their daily tasks. Late detection of this disease limits the effectiveness of treatment options. The diagnoses, at the neurological level, are expensive, uncomfortable and often used in a stage where clinical symptoms are already visible. Thus, it is paramount to develop new practical low-cost techniques that allow the detection of symptoms in a pre-clinical stage so that the necessary therapy processes can be initiated, maximizing their effectiveness and providing a longer quality of life for the patient. Speech signals can be of particular interest for our purposes because they are very easy to acquire and they present highly regular acoustic-phonetic patterns over time [3,4].

It is well documented that the natural human aging process may cause acoustic modifications on the voice, however, PD and its early appearance accentuate these changes, having a very pronounced and negative repercussion in the affected individuals [5]. Hypokinetic is the most described symptom, however, others may be less perceptible, but they have a relevant importance in the detection since they are present at a very early stage of the disease, specifically dysphagia, dysarthria or dysarthrophony [6]. According to [7], patients with PD may present lower vocal tessitura and slower melodic variation, comparing with healthy individuals.

In [8], 26 adults with PD were submitted to laryngeal electromyography exams and vocal acoustic analysis. After collection of action potentials, both in vocal rest as well as in phonation, distinct results were found. A hypercontractibility electromyographic pattern was detected during vocal rest without evidences of vocal tremors. However, the spectrogram analysis of the acoustic signal showed several vocal tremors, in fact, this was the predominant characteristic of the target group, without correlation with the electromyographic data [8]. This brings concerns about the fallibility of some screening tests and introduce doubts on possible diagnosis of PD.

In the study of [9], a statistical difference in Pitch amplitude was determined between two groups. In the PD patients group, 80.77% had a value lower than 0.3, while in the control group only 12.28% presented values below 0.3. However, according to [10], some voice features are not a reliable indicator to determine PD, as they have many similarities with healthy individuals.

The i-Prognosis project aims to develop an early detection methodology for PD based in a mobile application that will monitor the interaction of users with technology and will track several parameters that may indicate the onset of the disease. In a possible pathology detection, the user is advised to seek a doctor [11].

There is an agreement on the importance of evaluation of speech as a differential PD diagnostic tool, however, there is still a gap in relation to the objective definition of acoustic patterns that characterize the disease [6]. It is essential to determine, in the totality of the acoustic features, the most relevant ones.

In this paper, we propose a new method for the detection of Parkinson's disease in pre-clinical conditions based on the analysis of speech characteristics, using a stochastic approach that allowed to identify the best feature set for maximizing the discrimination between pathological and non-pathological classes.

2. Methodology

2.1. Pipeline

The different steps of the proposed method are outlined in Figure 1 and its main functions are: 1) Use audio recordings (speech, short phrases), previously stored in a database (DB), to identify the typical speech pattern in people with a medically confirmed PD diagnose; 2) Audio recording (speech, short phrases) of healthy people for later comparison; 3) Use of Praat software [12] for extracting features from the audio signal; and 4) Statistical processing of the parameters registered in the features, identifying which are the indicators of propensity for PD.

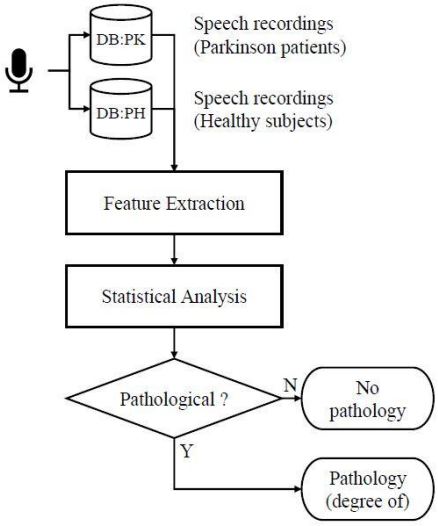

Figure 1. *Pipeline for the detection of PD through speech.*

2.2. Materials

For the development of this method, a speech database of PD patients was kindly provided by Professor Perdigão and his team [4]. The database was composed by several speech recordings from 22 patients (10 males and 12 females) aged between 44 and 79 years. Subsequently, to build a reference group, a similar acquisition was made in 22 healthy individuals (11 males and 11 females) aged between 20 and 73 years. Everyone contributed with different speech sentences, very short and objective.

2.3. Feature Set

For pitch estimation and feature extraction we have used the Praat software [12] that have allowed to obtain eighteen features grouped into four classes: a) Pitch, which is defined as the percussive sensation of the fundamental frequency. The software returns four features related to this class: PitchMed, PitchSDev, PitchMin and PitchMax [12,13]; b) Jitter and c) Shimmer, and derived indicators, are the most commonly used parameters for acoustic analysis [12]. These are indicative of vocal regularity on short pre-defined time intervals. Jitter indicates a disturbance in the fundamental frequency, whereas Shimmer indicates a disturbance in the amplitude. The software returns five values of Jitter (Jitter local, Jitter local absolute, Jitter rap, Jitter ppq5 and Jitter ddp) and six Shimmer values (Shimmer local, Shimmer local in dB, Shimmer apq3, Shimmer apq5, Shimmer apq11 and Shimmer ddp) [12-14]; d) Harmonic-to-Noise Ratio (HNR), also referred to as harmonic, represents the degree of acoustic periodicity of the signal and it is used to evaluate signal-to-noise ratio and voice quality. The respective unit of this parameter is displayed in dB. In the software, three types of HNR are analyzed: Hnr mean autocorrelation, Noise-to-harmonics ratio mean (Nhr) and Harmonics-to-noise ratio mean (Hnr). All the details for parameter calculation can be found in the Praat documentation [12].

2.4. Statistical analysis methods

After processing and extraction of the features, the data was organized into two populations: 1) Healthy Population – PH, and 2) Population with Parkinson's disease – PK.

The objective of this study was to test the following hypothesis formulated: The mean of each feature of the PH is different from the mean of each feature of the PK. For this propose, a feature analysis of each population was done with statistical method ANOVA.

ANOVA, or analysis of variance, is used to test hypotheses about population means. The main purpose of this method is to identify important independent variables and determine how they affect the response [15]. To apply this test, it is necessary to have the following assumptions: a) Samples are random and independent of each other; b) Underlying populations are normally distributed; c) Underlying population variances are equal.

The one-way ANOVA method was applied in this work to check if there are significant differences between the features of PH and the features of PK, and which features present greater variability between one population and another. For this propose, two hypotheses are considered: 1) Null hypothesis (H_0) – All population means are equal; and 2) Alternative hypothesis (H_1) – There are significant differences in the population means.
Considering the formulated hypothesis, ANOVA will indicate features that present different means between the PH and the PK if the following conditions are satisfied:

$$P < 0.05 \Leftrightarrow F > Fc$$

where F is the test statistic, Fc is the critical value of the Fisher distribution for 5% significance (which in this case is Fc=3.9258), and P is p-value of each test.

3. Results

The analysis of Table 1 shows that the features that verify $P < 0.05$ (and $F > Fc$) are: PitchSDev, PitchMax, Jitter local absolute, Shimmer local, Shimmer local in dB, Shimmer apq3 and Shimmer apq11. Of these, the most relevant features are Shimmer local, Shimmer local in dB

and Shimmer apq11, due to their higher values in the ANOVA test statistic, indicating a very significant difference between the populations means'. The Shimmer local is the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude [12], as shown in the equation below:

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

The Shimmer apq11 is the 11-point amplitude perturbation quotient, the average absolute difference between the amplitude of a period and the average of the amplitudes of it and its ten closest neighbours, divided by the average amplitude [12], as below:

$$Shim_{apq11} = \frac{\frac{1}{N-1}\sum_{i=5}^{N-5}\left|A_i - (\frac{1}{11}\sum_{n=i-5}^{i+5} A_n)\right|}{\frac{1}{N}\sum_{i=1}^{N} A_i} \times 100 \quad (2)$$

The remaining features, PitchSDev, PitchMax, Jitter local absolute and Shimmer apq3, do not present such evident values.

Features	*F (p-value)*	*Features*	*F (p-value)*
PitchMed	2.9297 (0.0897)	Shimmer local	20.2207 ($1.69E^{-05}$)
PitchSDev	9.5780 (0.0025)	Shimmer local in dB	29.4934 ($3.30E^{-07}$)
PitchMin	1.3144 (0.2540)	Shimmer apq3	5.2617 (0.0237)
PitchMax	5.6926 (0.0187)	Shimmer apq5	2.1855 (0.1421)
Jitter local	3.8279 (0.0529)	Shimmer apq11	21.9284 ($8.00E^{-06}$)
Jitter local absolute	5.5804 (0.0199)	Shimmer ddp	1.6123 (0.2068)
Jitter rap	2.1696 (0.1436)	Hnr mean autocorr.	0.0203 (0.8869)
Jitter ppq5	0.6394 (0.4256)	Nhr mean	0.0483 (0.8265)
Jitter ddp	2.1758 (0.1430)	Hnr mean	0.2175 (0.6418)

Table 1. *ANOVA test statistics for each feature (Fc=3.9258).*

After the identification of these relevant features, considering that the hypothesis formulated above was that the mean values of the features of the PH were different from the means of the PK, the means of each feature were analyzed (Table 2).

Features / Groups	*Mean*	*Variance*
Shimmer local_PH	0.156087	0.000525
Shimmer local_PK	0.126316	0.001974
Shimmer apq11_PH	0.183586	0.003280
Shimmer apq11_PK	0.139381	0.001799

Table 2. *Characterization of the populations according to the average of the features Shimmer local and Shimmer apq11.*

While analyzing the values of Shimmer local and Shimmer apq11 from each population (PH and PK), it was verified that the hypothesis formulated is verified and in addition: a) Mean PH ≠ Mean PK and b) Mean PH > Mean PK.

From these evidences, the following selection criterion was formulated: any individual with relevant features (Shimmer local and Shimmer apq11) with numerical values lower than those observed in a healthy individual could present PD.

To proceed with the validation and visual confirmation of this criterion, several samples of both populations were chosen randomly. Within these samples, only the relevant features were considered: Shimmer local and Shimmer apq11. Subsequently, for each healthy individual, a comparison was made of the value of each of their relevant features, with the value of each relevant feature of all individuals with Parkinson's.

Analyzing the results, it was observed that the criterion was verified and validated, except for sample number sixty (individual with PD), only for feature Shimmer local. Considering that the acquisition of a signal is influenced by several factors, for example, noise, it is thought that this sample could represent an outlier.

Continuing a deeper analysis of the results, a graph (Figure 2) was obtained with the data for Shimmer local and Shimmer apq11 in two groups by calculating the 95% confidence intervals. For each feature, the values obtained in PH and PK were overlapped, and lower and upper numerical limits (threshold values) that represent the possibility of being or not being diagnosed can be observed.

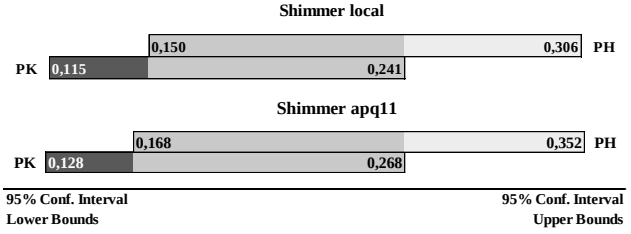

Figure 2. *Representation of threshold values for possible identification of PD.*

Evaluating the graph, it is predicted that from Shimmer local less than 0.150 the individual could present the disease, just as in the case of presenting a value of Shimmer apq11 lower than 0.168. On the other hand, in the situation where values higher than 0.241 and 0.268, for Shimmer local and Shimmer apq11, respectively, it may be admitted that a favorable diagnosis is being made without the presence of the disease.

With this approach we cannot evaluate the probability and respective percentage of Parkinson's positive check. The analysis of the confidence intervals only indicates that the may or may not have the disease, when the values obtained fall within the lower or upper bands, respectively. For example, for Shimmer local, it can be said that in 95% of the samples of a healthy population that can be collected in an analogous way, that is, under the same conditions in which the audio was recorded and with the use of same software, there would be a value above 0.150. If there are values that fit the grey zone

(intermediate zone), that is, for Shimmer local values between 0.150 and 0.241 and, Shimmer apq11 between 0.168 and 0.268, other features or conditions of the patient should be considered for a better diagnosis.

In [12], threshold values are presented for pathology with reference to values obtained by the Multi-Dimensional Voice Program (MDVP) software. Thresholds values for Shimmer local and apq11 are identified as 0.038 and 0.031, respectively [12]. However, it is not possible to compare with the values that are obtained by Praat, since each software executes its method, thus being able to have a divergence in the values, but always confirms the excellence character of the general feature Shimmer to determine several pathologies inherent to speech [16]. According to [16], in tests performed with male and female volunteers, both computer programs identified significant differences in the values obtained in the Jitter and Shimmer features. Thus, for an efficient diagnosis, the threshold values for pathology should be considered in a different way through the respective software.

4. Conclusions

In vocal acoustic analysis, interesting results were obtained regarding some relevant features, Shimmer local and Shimmer apq11, accessed through the Praat software and from the audio signal of the vocal tract, which can be considered as important factors in the detection of Parkinson's. These present significant differences when compared and analyzed in healthy individuals and in individuals with Parkinson's disease. These results lead us to consider and conclude that: 1) Shimmer local and Shimmer apq11 are the most prominent features in the detection of individuals with Parkinson's; 2) Individuals with Parkinson's presented numerical values for these features lower than the values obtained in healthy individuals; and 3) Analysis of signals of the vocal tract and detection of these discrepancies may contribute to the early detection of this pathology.

Therefore, all individuals with these features considered relevant for this evaluation, with Shimmer local and apq11 values lower than 0.150 and 0.168, respectively, could present the disease.

With the development of this method, it is expected to obtain early diagnosis, avoiding the use of expensive and invasive tests, so that physicians can prescribe the drugs available in advance and still at an early stage of the disease, exponentially retarding the effects inherent to this condition.

As future works, the voice parameters can be used in logistic regression models to find the probability of occurrence of Parkinson disease in a patient. Another possible approach would be to use multivariate discriminant analysis to combine the several parameters and find the border of the region in which the disease is to be diagnosed. These methods could be implemented in an algorithm that makes a pathological evaluation and, in the positive case, determines its stage and its probability in relation to the disease.

Unlike other machine learning systems, where the decision model appears as a black box, our approach allows to have a real insight of the data and to identify what acoustic features contribute to specify objective criteria towards the diagnose of Parkinson.

References

[1] Crankshaw CL. Neurodegenerative Diseases: Alzheimer's Disease, Huntington's Disease, Parkinson's Disease. *Biofiles, Sigma-Aldrich*, vol.7, n.2, 2012, pp.3-20.

[2] Dorsey ER et al. Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030. *Neurology*, vol.68, 2007, pp.384-386.

[3] Website of the EPDA, European Parkinson's Disease Association. http://www.epda.eu.com/en/ (Consulted: March 2017).

[4] Proença J, Veiga A, Candeias S, Lemos J, Januário C, and Perdigão F. Characterizing Parkinson's Disease Speech by Acoustic and Phonetic Features. *Springer International Publishing Switzerland*, 2014, pp. 24-35.

[5] Ferreira FV, Cielo CA, Trevisan ME. Vocal acoustic measures in Parkinson disease: case study. *Journal CEFAC*, 2009.

[6] Dias AE, Limongi JC. Treatment of vocal symptoms in Parkinson's disease: The Lee Silverman method. *Arch Neuro-Psychiatry*, vol.61, n.1, 2003, pp.61-66.

[7] Azevedo LL, Cardoso F, Reis C. Acoustic analysis of prosody in females with Parkinson's disease: comparison with normal controls. *Archive Neuro-Psychiatry*, vol.61, n.4, 2003, pp.999-1003 (ISSN: 1678-4227).

[8] Zarzur AP, Duarte IS, Gonçalves GN, Martins MA. Laryngeal electromyography and acoustic voice analysis in Parkinson's disease: comparative study. *Brazilian J. of Otorhinolaryngology*, vol.76, n.1, 2010, pp.40-43 (ISSN: 1808-8694).

[9] Pereira JC et al. Residual signal auto-correlation to evaluate speech in Parkinson's disease patients. *Arch Neuro-Psychiatry*, 4-64, 2006, pp.912-915 (ISSN: 1678-4227).

[10] Benba A, Jilbab A, Hammouch A. Voice analysis for detecting patients with Parkinson's disease using hybridization of the best acoustic features. *Int. J. Electronic Eng and Informatics*, vol.8, n.1, 2016, pp.108-116.

[11] Website of the Project i-Prognosis. http://www.i-prognosis.eu (Consulted: March 2017).

[12] Boersma P, Weenink D. Praat: doing phonetics by computer (Computer program). Version 6.0.25, retrieved 1 March 2017 from http://www.praat.org/.

[13] Steffen N, Moschetti MB. 248 Children's Jitter and Shimmer Acoustic Parameters. *Brazilian J. of Otorhinolaryngology*, vol.63, n.4, 1997, pp.329-334.

[14] Teixeira JP, Ferreira DB, Carneiro SM. Vocal acoustic analysis - determination of Jitter and Shimmer for diagnosis of speech pathologies. IPB, ESTG, Bragança. *Luso-Mozambican Engineering Congress*, 2011.

[15] Wackerly D, Mendenhall W, Scheaffer R. Mathematical Statistics with Applications. 7th Ed., Duxbury, 2008 (ISBN: 978-0495110811).

[16] Lovato A et al. Multi-Dimensional Voice Program (MDVP) vs Praat for Assessing Euphonic Subjects: A Preliminary Study on the Gender-discriminating Power of Acoustic Analysis Software. *Journal of Voice*, vol.30, n.6, 2016, pp.765.e1-765.e5.