# Detection of Parkinson's Disease through Acoustic Parameters

[1]Imen Daly, [1]Zied Hajaiej, [2]Ali Gharsallah

[1]Laboratory of Systems and signal Processing (LRSITI) National Engineering School of Tunis (ENIT)
[2]Research Units (CSEHF), Department of Physics, Faculty of Sciences of Tunis, Tunis El Manar University

*Abstract*—**Parkinson's disease is a neurological disorder. It is the second most common disease after Alzheimer's. Incidence rates for this disease are increasing rapidly with increasing life expectancy. The search for measures to diagnose the disease and monitor symptoms is an important step, despite the fact that it presents a number of challenges. Among the symptoms related to this disease is the disturbances of the voice which particularly occur in a remarkable way called hypokinetic dysarthria which is presented by the poverty of the gesture in all the characteristics of the speech (phonatory, prosodic, articulatory and rhythm). Our goal is to do a study based on voice analysis at the level of the glottis to examine some early parameters measured using the LF model and clinical manifestations to help diagnosis of the disease.**

*Keywords—Parkinson's disease; LF model*

## I. INTRODUCTION

The work presented in this paper focuses on variations and analysis of the parameters of the glottal source in Parkinson's speech. This research can be relevant for different speech signal domains; it announces the characteristics of the speech source of normal speakers and those with Parkinson's disease (PD). It is a chronic neurodegenerative disease with a variety of motor and non-motor symptoms. The second most commonly diagnosed after Alzheimer's disease [4]. It results in the slow and progressive death of a set of neurons of the human brain that have an important role in the control of movement [1]. It is a loss of muscular control and cognitive impairments, known by the trembling of the body and the voice called dysarthria [2]. Age is the simplest factor in this disease, so that the prevalence rate is 0.5% to 1% for the age of 65 to 69 years and 3% for the oldest 80 years [5], [3]. It is therefore preferable to detect it in advance before it is well developed. Although there are other symptoms of this disease, no final biomarker exists or specific measures for diagnosis [10]. The latest research findings indicate that approximately 70-90% of patients with this disease have a form of vocal disability [6] and this index may be one of the best symptoms of this disease. The idea of studying the speech signal and its tremors that either occur in instantaneous frequencies or in a variation in amplitudes is a good index for the diagnosis and monitoring of the progression of this disease. Ideally, such a measure would be non-invasive and could be acquired outside the clinical setting without the need for expert assistance. In this context, several studies have been carried out on acoustic measurements based on jitter and Shimmer and NHR ratios, fundamental frequency [7], intensity and formants [8] or joint model [9]. In our work, we are interested in measuring the

spectrum of the glottal source, estimated from the speech signal to identify parameters that behave differently in Parkinson's and Healthy speech. This study was motivated by all the research that used glottal behavior and characteristics [10]-[13].

The organization of this document and as follows. In the first section, we present the glottal model also the voice data used in the analysis of the sources. The next section describes the experience and the results obtained and finally a conclusion.

## II. METHODOLOGY

### A. Acoustic Modeling of Speech Production

The acoustic theory of speech production presented by [15] is based on the analysis of glottal sources. It is a theory that allows the functional separation of speech production in two parts (source and filter) to improve the understanding of this phenomenon. The filter is assumed linear time invariant (LTI), which means that each short-term speech segment contains constant parameters without any interaction with the glottal source. These hypotheses are clearly simplifications, since the exact physical description of the generation and propagation of sound in the vocal organs leads to a complex set of differential equations [16] in order to facilitate and separate the components of the model and understand their functioning. The acoustic speech signal S (n) is the result of the convolution of the waveform of the glottis source g (n), which presents the volume of the airflow exiting the lung to the vocal cords with an impulse response of the filter leaving the voice path representing the frequencies of resonant formants and the radiation of the lip r (n).

$$S(n) = g(n).v(n).r(n) \tag{1}$$

In the z domain, the model can be defined as:

$$S(Z) = G(Z).V(Z).R(Z) \tag{2}$$

Where R(Z) represents the discrete time radiation impedance and V(Z) the discrete time speech signal of the volume velocity at the glottis and E(Z) the discrete signal enters the time domain. Any loss in the system occurs by radiation on the lips through which has a high-pass filtering effect modeled with a single Zero such that R(Z) = 1-$Z^{-1}$

*1) Glottal source*: The production of the speech signal requires the use of several organs, a flow of air coming out of the lungs passes through the larynx where the vibration of the

area is produced by the voice piles. The opening between these two stacks is called the Glottis slit or glottis [13], see Fig. 1.

The air pressure exerted on the cells is sufficiently high to separate them and allows the flow of air to flow so that the glottis is released to close. Several factors contribute to movement. The first is the nature of the elastic tissue which forces the vocal folds to return to its initial position, the second is described by the Bernoulli effect [18] (the retro-aspiration effect of the cordial mucosa) and the third by the action of the Decreased Pressure [14].

This cycle is repeated several times periodically, its duration is called the fundamental period noted T0 and its rhythm determines the fundamental frequency of the voice (number of vibrations per second denoted F0 expressed in Hz).

The flow waveform is the result of vibration of the vocal chords from an open position and a closed position that form the two most important phases of the glottal flow impulse during each cycle.
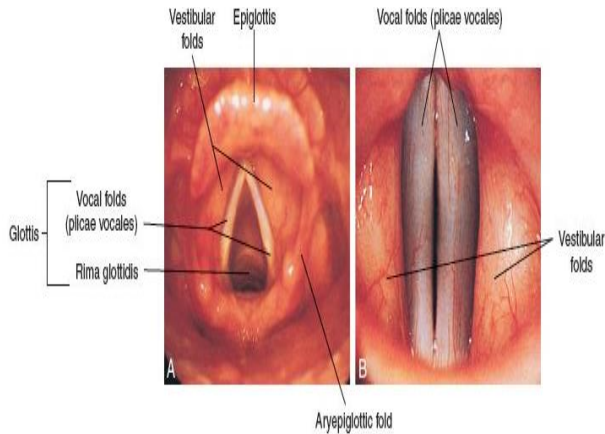


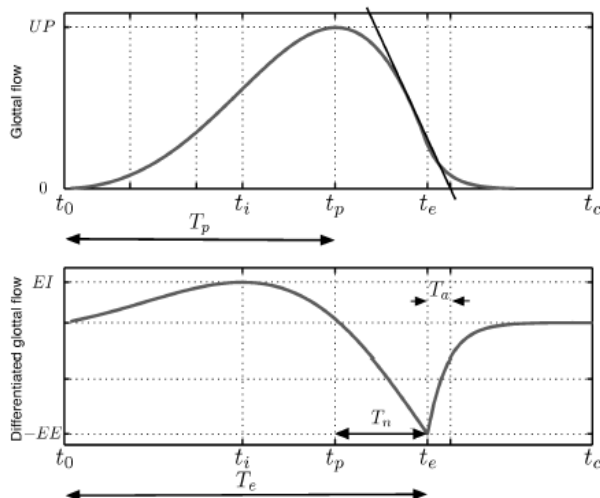Fig. 1. The Positions of the Open Vocal Cords (A) and Closed (B).



Fig. 2. The different instants and periods of the LF Model.

### a) LF Model

Several models exist to determine the characteristics of the glottal flow. We chose to use the Liljencrants fant (LF) model [19] in this document because it gives a good fit in the waveform.

This model can be presented with parameters such as tp, te and ta (see Fig. 2). It is composed in two parts opening and closing and calculated as follows:

$$E(t) = \begin{cases} E_0\, e^{\alpha t}\, \sin w_g t & 0 \leq t \leq t_e \\ -\frac{EE}{\varepsilon t_a}\left[e^{-\varepsilon(t-t_e)} - e^{-\varepsilon(t_c-t_e)}\right], & t_e < t \leq t_c \end{cases} \tag{3}$$

Where ta is related to the return phase of the model and tp is the positive peak of the flux also defined by the zero point before the derivative of the flux and α, ε the parameters controlling the shape of the model [20].

The pulse form of this model can also be presented with other parameters.

- The R parameters

In the section, a set of parameters was determined and are expressed in quotients for to describe the shape of the glottal source signal E (t) and characterize the shape of the pulse of the model LF.

The Rk parameter is defined by (4), it is the measure of the asymmetry in the cycle, Rg represents the normalizer of the fundamental frequency presented in (5) and Rd in (6) to capture all the variation of the model LF It is related to the fundamental frequency and the other two parameters Rg and Rk.

$$R_k = \frac{t_e - t_p}{t_p} \tag{4}$$

$$R_g = \frac{T}{2T_p} \tag{5}$$

$$R_d = 1000\left(\frac{A}{E_e}\frac{f0}{110}\right) \tag{6}$$

The interval between the beginning of the glottal impulse and the instant where tp reaches the maximum is called the opening phase, see (2). At this moment, the vocal folds begin to close and the amplitude of the flow begins to decrease until the sudden closure of the globe. The time that corresponds to the duration of passage of the area flux through the globe when the folds are open is known by the open phase is measured by the following formula te −Ta. With Ta the point where the tangent to the exponential in t = te, it also serves to determine the phase of transition between the open phase and the closed phase known by the return phase is calculated by the following formula [21]:

$$T_a = t_a - t_e \tag{7}$$

Finally, the last phase of the cycle when the folds are completely closed at time tc. The model can also be presented by other parameters, which are correlated with the quality of the voice and its spectral properties as in [22].

- The parameters NAQ and QOQ

Several temporal characteristics can be determined as a function of time derived from the glottal form [23], using the model LF and its parameters such as the open quotient Qq, the asymmetry coefficient αm and the speed quotient Sq. These moments are always difficult to detect correctly. To solve this problem a new parameter has been proposed by Hacki [24] describing the open time of the glottis named quasi-open quotient (QOQ). It is defined by the ratio between the time of opening and closing of the glottis and corresponds to the period during which the flow of the glottis is greater than 50% of the difference of two maximum and minimum flows. Laukkanen [25] has to experience that this parameter can be used to study the variation of the glottal source in case of stress and emotion as well as Airas and alku which have announce that this quotient and the best parameter to use to test the degrees of reflection of the changes of phonation [26].

The temporal characteristics of the glottal source can also be presented by measuring the amplitude of the glottis and its derivative.

$$NAQ = \frac{f_{ac}}{d_{peak}.T0} \qquad (8)$$

It is the normalized amplitude quotient (NAQ) proposed by Alku et al [27] using two amplitudes f_ac the amplitude of flow and d_peak the negative peak amplitude (see Fig. 3).
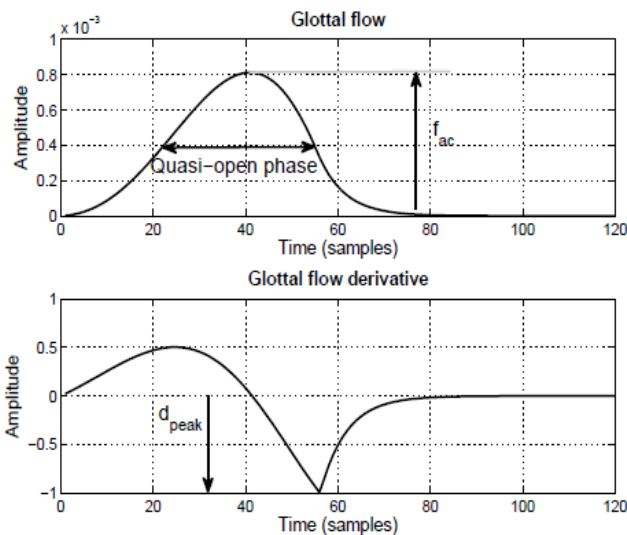


Fig. 3. Parameters needed for NAQ and QOQ measurement.

## III. EXPERIMENTAL DETAILS

The aim of this article is to determine whether the parameters of the glottal signal have remarkable characteristics between the two databases PD and HC. To evaluate this, we used Matlab to measure its various parameters and anova for statistical measurements.

### A. Database

In this study we used, recording of 36 native Czech speakers, the PD age group 64.21 ± 9.46 SD (extreme 41-82). Before the experience each patient did a neurological examination to rank them according to the scale of which varies from 1 to 5 according to the degree of unilateral motor disorder as well as the classification scale (UPDRS III) also The H&Y score was 2.1 ± 0.4(1–3),. The healthy control group (HC) with a mean age of 64.21 ± 9.22 (42-80) years. None of the HC group participants had a history of neurological disorders or speech. The study was approved by the Ethics Committee of the General University Hospital in Prague, Czech Republic, and all participants provided informed and written consent.

The description of some records used in this document is displayed in the following Table I:

TABLE I. SOME CHARACTERISTIC OF PARKINSON'S PATIENTS

| Participant code | Gender | Age (years) | Age of disease onset (years) | Duration of disease from first symptoms (years) | UPDRS III total | UPDRS III 18. Speech |
|---|---|---|---|---|---|---|
| PN301 | M | 74 | 64 | 10 | 22 | 2 |
| PN302 | M | 77 | 70 | 7 | 14 | 1 |
| PN304 | M | 76 | 70 | 6 | 23 | 1 |
| PN307 | M | 66 | 61 | 5 | - | - |
| PN314 | F | 62 | 57 | 5 | 32 | 0 |
| PN316 | F | 61 | 57 | 4 | 38 | 2 |
| PN401 | M | 56 | 55 | 1 | 8 | 0 |
| PN403 | M | 68 | 59 | 9 | 12 | 0 |
| PN405 | M | 76 | 65 | 11 | 27 | 2 |
| PN407 | M | 55 | 44 | 11 | 16 | 0 |
| PN409 | M | 79 | 70 | 9 | 9 | 0 |

### B. Recording

The recordings of the speakers are recorded in a quiet room with a professional microphone (Beyer-Dynamics Opus 55, Heilbronn, Germany) placed 5 cm from the mouth of each patient. All of his recordings were digitized at a sampling rate of 48 Khz with 16-bit quantization. All speakers pronounced the vowel a under the control of a speech specialist without any time limit was imposed during the recording.

## IV. RESULTS

Glottic pulse parameters related to time and frequency provides quantitative information for the examiner about their importance in biomedical applications. To prove it, we used both types of recording of PD and HC speakers, regardless of age or gender. The results for each glottal parameter are shown in the following Table II. This table presents the mean value and standard deviation calculated for all patients in two PD and HC groups. A remarkable difference can be observed in the different parameters.

The analysis of group differences between PD and controls performed using a rank sum of Wilcoxon. It is a non-parametric statistical test that uses the data distribution assumption the same with the Spearman correlation reliability test that is used if the static correlation variables do not have affine relationships, see Table III.

These parameters were also tested following the clinical manifestations with a significant correlation between the NAQ parameter and the UPDRS discourse 18 (Z = 0.36, p <0.04) (Table IV).

TABLE II.    THE RESULTS OF MEASURING THE MEAN AND THE STANDARD DEVIATION OF GLOTTAL PARAMETERS FOR PD AND HEALTHY SPEECH

| | Results | | | |
|---|---|---|---|---|
| | PD speech | | HC speech | |
| | Average | Standard Deviation | Average | Standard Deviation |
| Ee | 0,0566 | 0,0926 | -0,1447 | 0,1393 |
| Te | 0,0027 | 0,0029 | 0,0043 | 0,0149 |
| Tp | 0,0025 | 0,0027 | -0,0256 | 0,0999 |
| Rd | 0,0140 | 0,0179 | 0,0391 | 0,0667 |
| Rg | 0,1730 | 0,1383 | 0,2971 | 0,3193 |
| Rk | 1,1767 | 1,3183 | 1,2827 | 3,3212 |

TABLE III.    THE RESULTS STATISTICAL DIFFERENCES BETWEEN ALL GROUPS

| Variable | Spearman correlation | | Wilcoxon rank sum test | |
|---|---|---|---|---|
| | PD vs HC | | PD vs HC | |
| | Z | p | Z | p |
| NAQ | 0,81 | p< 0.001 | -2,19 | p < 0,05 |
| QOQ | 0,78 | p< 0.001 | -1,84 | p = 0,06 |

TABLE IV.    THE CLINICAL MANIFESTATIONS WITH NAQ ET QOQ

| Variable | UPDRS | | UPDRS speech | | UPDRS rigidity | |
|---|---|---|---|---|---|---|
| | Z | p | Z | p | Z | p |
| NAQ | 0,24 | 0,17 | 0;36 | 0,04 | -2,19 | 0,96 |
| QOQ | -0,17 | 0,93 | 0,96 | 0,98 | 0,078 | 0,66 |

## V.  DISCUSSION AND CONCLUSION

This work, we presented a glottal analysis based on the estimation of flux and the extraction of some parameters in the time domain in order to examine their relations in the two bases PD and HC. The results found are based on several direct correlations to demonstrate voice quality and glottal velocity. A significant effect was found with the NAQ parameter through a positive correlation describing the amplitude variation during pronunciation and reflecting the distinctive effect of Bradykinesia in Parkinson's disease.

Another presented part of our paper is the analysis of glottal source local relations and the clinical manifestations described by the UPDRS speech scale. Our results show that the NAQ parameter confirms the intelligibility of the manifestation.

There is a limit in this analysis, in particular that our results are based on a relatively small number of PD patients who are heterogeneous in terms of age, duration of illness and gender of the patient.

REFERENCES

[1]  Kostikis, N., et al. "A smartphone-based tool for assessing parkinsonian hand tremor." IEEE journal of biomedical and health informatics 19.6 (2015): 1835-1842.

[2]  Rhouma, A. Ben, and S. Ben Jebara. "Features based on quasi-sinudoidal modeling for tremor detection in Parkinsonian voice." Advanced Technologies for Signal and Image Processing (ATSIP), 2014 1st International Conference on. IEEE, 2014.

[3]  Nussbaum, Robert L., and Christopher E. Ellis. "Alzheimer's disease and Parkinson's disease." New england journal of medicine 348.14 (2003): 1356-1364.

[4]  Liu, Yao, et al. "The analysis of regularity and synchrony of Parkinsonian tremor using Approximate Entropy and Cross-Approximate Entropy." Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), International Congress on. IEEE, 2016.

[5]  Liu, Chen, et al. "Closed-Loop Control of Tremor-Predominant Parkinsonian State Based on Parameter Estimation." IEEE Transactions on Neural Systems and Rehabilitation Engineering 24.10 (2016): 1109-1121.

[6]  Ho, Aileen K., et al. "Speech impairment in a large sample of patients with Parkinson's disease." Behavioural neurology 11.3 (1999): 131-137.

[7]  Jones, Harrison N., et al. "Speech motor program maintenance, but not switching, is enhanced by left-hemispheric deep brain stimulation in Parkinson's disease." International journal of speech-language pathology 12.5 (2010): 385-398.

[8]  Rusz, J., et al. "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease." The journal of the Acoustical Society of America 129.1 (2011): 350-367.

[9]  Bocklet, Tobias, et al. "Detection of persons with Parkinson's disease by acoustic, vocal, and prosodic analysis." Automatic Speech Recognition and Understanding (ASRU), 2011 IEEE Workshop on. IEEE, 2011.

[10] Hanson, Helen M., and Erika S. Chuang. "Glottal characteristics of male speakers: Acoustic correlates and comparison with female data." The Journal of the Acoustical Society of America 106.2 (1999): 1064-1077.

[11] Kreiman, Jody, Bruce R. Gerratt, and Norma Antonanzas-Barroso. "Measures of the glottal source spectrum." Journal of speech, language, and hearing research 50.3 (2007): 595-610.

[12] Li, Haoxuan, Ronan Scaife, and Darragh O'Brien. "LF model based glottal source parameter estimation by extended Kalman filtering." Proceedings of the 22nd IET Irish Signals and Systems Conference. 2011.

[13] Drugman, Thomas. "Advances in glottal analysis and its applications." University of Mons, Belgium (2011).

[14] Drugman, Thomas, et al. "Glottal source processing: From analysis to applications." Computer Speech & Language 28.5 (2014): 1117-1138.

[15] Fant, G. (1970). Accoustic theory of speech production: with calculations based on X-ray studies of Russian articulations. Mouton & Company.

[16] Rabiner, L. R., & Schafer, R. W. (1978). Digital processing of speech signals (Vol. 100, p. 17). Englewood Cliffs, NJ: Prentice-hall.

[17] Hamlet, S. M. (2012). Comparing Acoustic Glottal Feature Extraction Methods with Simultaneously Recorded High-speed Video Features for Clinically Obtained Data.

[18] Pelorson, X., Hirschberg, A., Van Hassel, R. R., Wijnands, A. P. J., & Auregan, Y. (1994). Theoretical and experimental study of quasi steady flow separation within the glottis during phonation. Application to a modified two mass model. The Journal of the Acoustical Society of America, 96(6), 3416-3431.

[19] Fant, G., Liljencrants, J., & Lin, Q. G. (1985). A four-parameter model of glottal flow. STL-QPSR, 4(1985), 1-13.

[20] Doval, Boris, Christophe d'Alessandro, and Nathalie Henrich. "The spectrum of glottal flow models." Acta acustica united with acustica 92.6 (2006): 1026-1046.

[21] Childers, Donald G. "Glottal source modeling for voice conversion." Speech communication 16.2 (1995): 127-138.

[22] Alku, Paavo. "An automatic method to estimate the time-based parameters of the glottal pulseform." Acoustics, Speech, and Signal Processing, 1992. ICASSP-92., 1992 IEEE International Conference on. Vol. 2. IEEE, 1992.

[23] Alku, P. (1992). Glottal wave analysis with pitch synchronous iterative adaptive inverse filtering. Speech communication, 11(2-3), 109-118.

[24] Hacki, T. (1989). Klassifizierung von glottisdysfunktionen mit hilfe der elektroglottographie. Folia Phoniatrica et Logopaedica, 41(1), 43-48.

[25] Laukkanen, A. M., Vilkman, E., Alku, P., & Oksanen, H. (1996). Physical variations related to stress and emotional state: a preliminary study. Journal of Phonetics, 24(3), 313-335.

[26] Airas, M., & Alku, P. (2007). Comparison of multiple voice source parameters in different phonation types. In Eighth Annual Conference of the International Speech Communication Association.

[27] Alku, P., Bäckström, T., & Vilkman, E. (2002). Normalized amplitude quotient for parametrization of the glottal flow. the Journal of the Acoustical Society of America, 112(2), 701-710.