Journal of Pharmaceutical Research International

23(3): 1-7, 2018; Article no.JPRI.42403
ISSN: 2456-9119
(Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759)

Physiology of Speech/Voice Production

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

[1]Signal, Image and Information Technology Laboratory, National School of Engineers of Tunis, Tunisia.
[2]Type Units of Research in High Frequency Electronic Circuits and Systems, Faculty of Mathematical, Physical and Natural Sciences of Tunis, Tunisia.

Authors' contributions

This work was carried out in collaboration between all authors. Author ID designed the study, performed the statistical analysis, wrote the protocol and wrote the manuscript. Authors ZH and AG managed the analyses of the study. All authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2018/42403
Editor(s):
(1) Vasudevan Mani, Universiti Teknologi MARA (UiTM), Selangor, Malaysia.
Reviewers:
(1) Wen-Dien Chang, China Medical University, China.
(2) Yueh-Feng Sung, Tri-Service General Hospital, National Defense Medical Center, Taiwan.
(3) Md. Noorain Alam, Postgraduate Institute of Medical Education and Research, Chandigarh, India.
(4) Jose M. Lasso, Universidad Complutense Madrid, Spain.
Complete Peer review History: http://www.sciencedomain.org/review-history/25827

Original Research Article

Received 11th May 2018
Accepted 21st July 2018
Published 8th August 2018

ABSTRACT

Aims: This study looks at the pathological disorders analysis from speech signals. She presents the effectiveness of the few parameters of the glottal source and these moments of closure for the diagnosis of the Parkinson's disease.

Study Design: This analysis is performed around the Czech database that contains many recordings were divided between PD and healthy control (HC). More than 68 subjects are used where each speaker pronounce the sustained vowel [a] in normal intonation.

Methodology: Several researchers are trying to produce objective measures to diagnose Parkinson's disease or to act as an objective tool to assist in its decision-making by measuring some vocal parameters. In this study, we studied the hyper nasality of PD and HC using acoustic analyzes.

Results: The results presented in this study indicate that speech related symptoms of PD are evident in the glottal flow signal. Particularly notable results were recorded for the glottal parameters, with the correlations found between the measurements and the deviations of the

**Corresponding author: E-mail: i_daly@yahoo.fr;*

speech during the pronunciation of the vowel. These results can be used as positive indicators for the diagnosis of the disease

Conclusion: This study presents a glottal source analysis with its different parameters in the frequency domain and the detection of the closure time by using two algorithms. This study suggests that these findings will be useful as objective indices that can either diagnose Parkinson's disease or act as an objective decision support tool.

Keywords: Parkinson's disease; healthy control; glottal source.

1. INTRODUCTION

In Speech processing domain, voice quality analysis methods, and closure time detection (GCI) have been used in several research [1]. GCIs are referred to cases of significant excitation of the vocal tract during each glottal cycle .These instants are presented of high energy in the glottal signal [2].

The detection of these instants has been used in several applications such as optimization through speech synthesis [3], modeling of the glottal source through inverse filtering [4] and speech modifications [5]. In recent years, several researches have been based on speech technology to facilitate the modification of the characteristics of the human voice for the evaluation of dysphonia. This study is based on the analysis of laryngeal pathologies related to the vocal cords which cause disorders during the production of speech. According to the model source filter these disorders are concentrated in the source of excitation named glottis source [6] where can detect several diseases such as Parkinson's disease. It is a disease is known by a neurological disorder characterized by the loss of dopaminergic neurons in the substance nigra pars compacta, it is associated with various motor and non-motor deficits [7]. Disorders associated with this condition often occur in the 1% to 2% of people over 60 [8] because of remarkable symptoms such as muscle rigidity, tremors, and postural instability. A previous study has shown that 70-90% of these people have speech impairment characterized by hypokinetic dysarthria [9].

It affects many aspects of the vocal signal, such as noisy breathing, voice quality, reduced height and volume variability, and reduced stress [10] [11]. This is why dysarthria is a prodromal manifestation that may be present several years before diagnosis with its various indices that can give important information in the treatment of the voice in the patients themselves [12,13].

In particular, several studies have been carried out on the behavior of the glottal waveform estimated from the vocal signal in order to identify these indices, which behave differently in parkinsonian speech, contrary to healthy speech.

This research was motivated by several studies using the potential of orthopedic measures to diagnose disease [14] and studies describing glottal behavior with GCI analysis to obtain an accurate estimate using the SEDREAMS [15] and SE-VQ methods [16].

The organization of this document and as follows. In the first section, we present the glottal model as well as the vocal data used in the source analysis. The following section describes the experience, the results obtained, and finally a conclusion contains a summary and future work.

2. MATERIALS AND METHODS

2.1 Database

For the purposes of the current study recordings from a total number of 68 subjects was used. All the participants were equally distributed between PD and HC groups. The PD group consisted of 20 males and 14 females.

All the participants met the criteria for the diagnosis of Parkinson's disease and were on the stable dopaminergic medication for at least 4 weeks before the date of examination, which was performed in the on medication state.

Before the experiment, each patient underwent a neurological examination, was scored according to the Hoehn and Yahr (H&Y, ranging from 1 to 5, where 1 indicates mild unilateral motor disorder and 5 indicates confinement to wheelchair or bed) [17] and the motor section of Unified Parkinson’s Disease Rating Scale (UPDRS III, ranging from 0 to 108, with 0 for no motor manifestation and 108 representing severe motor distortion).

The healthy control group (HC) was composed of 20 males and 14 females. None of the participants had a history of neurological.

The study was approved by the Committee of the General University Hospital in Prague, Czech Republic. The description of some records used in this document is displayed in the following Table 1.

Table 1. Statistics of patients

	General age		**Female group age**		**Male group age**	
F/M	**mean**	**SD**	**mean**	**SD**	**mean**	**SD**
PD						
14/20	64.21	9.46	60.14	8.90	67.05	8.97
HC						
14/20	64.21	9.22	60.29	8.65	66.95	8.79

2.2 Experimental Settings

The recordings were made in a quiet room with the presence of a specialist scientist the signal was digitized at a sampling rate of 48 KHz and a quantization of 16 bits. He used a Beyer-Dynamics Opus 55 microphone, Heilbronn, Germany) placed 5 cm from the patient's mouth and asked the participants to breathe and pronounce the A vowel as long as possible.

3. PHYSIOLOGY OF SPEECH PRODUCTION

The acoustic theory of speech production presented by Fant [18] 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. Some diseases can cause changes during speech production. These main symptoms result in the amplitude variation and pitch perturbation called by the aperiodicity in the vocal signal also the acceleration decrease as it is displayed in the following Table:

Table 2. The most deviant criteria in Parkinson's disease

Monotony
Decrease in accentuation
Mono-intensity
Uncertainty of the consonants
Inappropriate breaks
Brief acceleration (paroxysmal)
Hoarsely
Blown voices
Average height (lowering)
Variable flow (accelerated)

3.1 Proposed Method

In this section, we present our methods and its different steps in the following Fig. 1.

<u>3.1.1 The moments of glottal (opening and closure)</u>

The periodic movement of the vocal cords characterizes the laryngeal cycle, with these opening moments and closure. The excitation of vocal leads is made at the instant of closing of the glottis, it represents the most significant moment and its determination is an important step especially in the case of pathological. The detection of his instants was made through two well-known algorithms the first base on the mean (SEDREAMS) and the second SE- VQ.

3.1.1.1 Method residual excitation and a mean based signal (SEDREAMS)

It is a method based on residual excitation presented by Drugman [19]. It is based on the signal mean in the same way as the ZFF method except that it is calculated directly from signal rather than the 0 Hz resonator output.

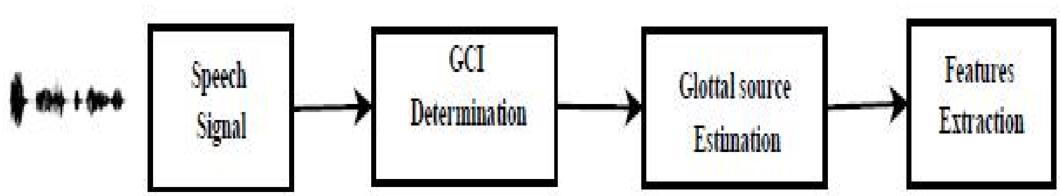

Fig. 1. Different steps for features extraction

3.1.1.2 Method (SE-VQ)

This method is based on the SEREAMS algorithm with some modification to better manage voice quality (VQ). These changes make it possible to select the optimal path at the location of GCI by removing the false positives found in the squeaky signal [20].

3.1.2 Glottal source estimation

The estimation has been the subject of several works but it is very well known that the voice is produced by an excitation signal (source) convoluted by an impulse response at the vocal tract (filter)

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

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

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

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}

The majorities of the estimation methods are based on inverse filtering. They make it possible to estimate the vocal tract and dismount its glottis flow by eliminating its contribution produced by the vocal tract.

The most used is the iterative adaptive inverse filtering method (IAIF) [23] that is based on iterative filtering as well as the glottis components. The latter also had an improvement through the replacement of the LPC analysis by the technique DAP (Discret all pole) [22] which gave more precise results for the acute voices [21].

3.1.3 Glottal features

3.1.3.1 Frequency domain features

Parameterization of the frequency of the glottal source has been used in several studies to evaluate the spectral slope of the glottis source using the fundamental frequency and its harmonics.

Such a measure, called the Harmonic Richness Factor (HRF), has been demonstrated by [24]. This parameter quantifies the amount of harmonics in the spectrum of the glottal source. It is defined as the ratio between the sum of the amplitudes of the harmonics and that of the fundamental frequency. Another measure has also been used is the parameter H1-H2 [25].

It is known by the ratio between the amplitudes of the source and the second harmonic and it is used to characterize the quality and the variation of the human voice [26] also the PSP parabolic spectral parameter, which makes it possible to quantify the volume of the glottal wave at the limit of the maximum spectral decay.

4. RESULTS AND DISCUSSION

We did a comparative study between the two methods in order to detect the closing moments as it is presented in the following Fig. 2. A remarkable difference is displayed because of the disruption of the glottis during a cycle. It is clear that the GCI obtained by the SE-VQ algorithm is more accurate since it is able to select the optimal closure time path through post-processing in order to remove the false positives that occur in the acute regions of the speech. Which explains our choice to use it in our analysis for the measurements of the frequency parameters.

In this context, we measured the hypernasality of PD and HC using acoustic analyzes. With these measures, we found positive direct correlations between the two bases with the wilcoxon rank sum test as shown in Table 1. The analysis revealed a significant difference in the H1_H2 (Z = -2.63, n1 = n2 = 34, p < 0 01) and PSP (Z = -2.34, n1 = n2 = 34, p < 0 05).

Table 3. Acoustic analyzes between all groups

Particulars	Z	P
HRF	1.26	**0.20**
H1H2	-2.63	**0.008**
PSP	-2.34	**0.01**

For analysis of the relationship between measures and clinical manifestations no statistically significant correlation was found (see Table 4).

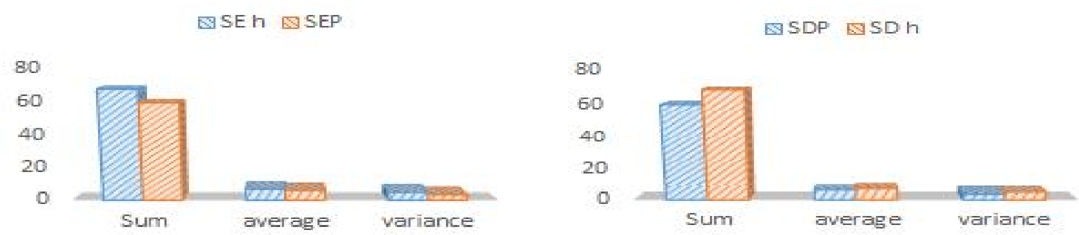

Fig. 2. GCI determination from HC and PD obtained by SE-VQ and SEDREAMS

Table 4. Results of correlations between acoustical and perceptual measures of hyper nasality and clinical manifestations of PD groups

Particulars	UPDRS		UPDRS speech		UPDRS rigidity	
	P	r	P	r	P	r
H1H2	0.44	0.13	0.18	0.23	0.34	0.17
HRF	0.25	-0.20	0.70	0.06	0.62	-0.08
PSP	0.42	0.14	0.40	0.14	0.96	0.007

The results obtained in this paper on the glottal flow estimation are encouraging, The effects of the disease were also demonstrated in the study [27] with laryngeal abnormalities that gave a high degree of sex differences. The same effects were found in [28] through aerodynamic differences in mean subglotal pressure and laryngeal resistance such that the majority of patients were unable to produce phonation at comfortable intensity levels for normal subjects. This proves that voice production in Parkinson's disease is a complicated process and abnormal stiffness of the larynx is not the only cause, but it has been shown to be a significant cause of pathological vibratory characteristics. The same also with kinematic, pressure and airflow parameters were used to study the nature of voice breathing [29].

In addition, some recent studies have shown that the vocal tract is also important for voice production based on a nonlinear vocal model with high coefficients and that has captured many vocal features and processing effects the voice in Parkinson's disease.

5. CONCLUSION

This study focus on the use of the parameters describing the glottal source to characterize the laryngeal pathologies. Using measurements of the glottal source, our study minimizes the effect of other speech system deficits such as articulatory interruptions and reveals a significant effect of PD on the vocal cords. The results found showed a good robustness with respect to the reduced quality of the speech signal caused by the variation of the speech because of the disease.

In future work, we are firstly, motivated to extract the same features from different vowels with intonations (low high and low-high-low). Where the main objective is to find a best separation in particular with pathological voice.

CONSENT

As per international standard or university standard written participant consent has been collected and preserved by the authors.

ETHICAL APPROVAL

It is not applicable.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. Naylor PA, Kounoudes A, Gudnason J, Brookes M. Estimation of glottal closure instants in voiced speech using the DYPSA algorithm. IEEE Transactions on Audio, Speech, and Language Processing. 2007;15(1):34-43.
2. Smits R, Yegnanarayana B. Determination of instants of significant excitation in speech using group delay function. IEEE

Transactions on Speech and Audio Processing. 1995;3(5):325-333.

3. Stylianou Y. Synchronization of speech frames based on phase data with application to concatenative speech synthesis. In Sixth European Conference on Speech Communication and Technology; 1999.
4. Drugman T, Dutoit T. Glottal closure and opening instant detection from speech signals. In Tenth Annual Conference of the International Speech Communication Association; 2009.
5. Rao KS, Yegnanarayana B. Prosody modification using instants of significant excitation. IEEE Transactions on Audio, Speech and Language Processing. 2006; 14(3):972-980.
6. Mezzalama M, Prinetto P, Morra B. Experiments in automatic classification of laryngeal pathology. Medical and Biological Engineering and Computing. 1983; 21(5):603-611.
7. Hornykiewicz O. Biochemical aspects of Parkinson's disease. Neurology. 1998; 51(Suppl 2):S2-S9.
8. Marsden CD. Parkinson's disease. J Neurol Neurosurg Psychiatry. 1994;57: 672–681.
9. Logemann JA, Fisher HB, Boshes B, Blonsky ER. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. Journal of Speech and Hearing Disorders. 1978;43(1):47-57.
10. Canter GJ. Speech characteristics of patients with Parkinson's disease: I. intensity, pitch, and duration. Journal of Speech & Hearing Disorders; 1963.
11. Darley FL, Aronson AE, Brown JR. Clusters of deviant speech dimensions in the dysarthrias. Journal of Speech, Language, and Hearing Research. 1969; 12(3):462-496.
12. Rusz J, Hlavnička J, Tykalová T, Bušková J, Ulmanová O, Růžička E, Šonka K. Quantitative assessment of motor speech abnormalities in idiopathic rapid eye movement sleep behaviour disorder. Sleep Medicine. 2016;19:141-147.
13. Postuma RB, Lang AE, Gagnon JF, Pelletier A, Montplaisir JY. How does parkinsonism start? Prodromal parkinsonism motor changes in idiopathic REM sleep behaviour disorder. Brain. 2012;135: 1860-1870.
14. QUATIERI, Thomas F. Speech signal processing. Prentice Hall Signal Processing Series; 2002.
15. Hanson DG, Gerratt BR, Ward PH. Cinegraphic observations of laryngeal function in Parkinson's disease. 1984; 94(3):34853.
16. Gath I, Yair E. Analysis of vocal tract parameters. Journal of the Acoustical Society of America. 1988;84:1628.
17. Marsden CD. Parkinson's disease. J Neurol Neurosurg Psychiatry. 1994;57: 672–681.
18. Fant G. Acoustic theory of speech perception. Mouton, The Hague; 1960.
19. Drugman T, Dutoit T. Glottal closure and opening instant detection from speech signals. In Tenth Annual Conference of the International Speech Communication Association; 2009.
20. Kane J, Gobl C. Evaluation of glottal closure instant detection in a range of voice qualities. Speech Communication. 2013;55(2):295-314.
21. Alku P, Vilkman E. Estimation of the glottal pulseform based on discrete all-pole modeling. In Third International Conference on Spoken Language Processing; 1994.
22. El-Jaroudi A, Makhoul J. Discrete all-pole modeling. IEEE Transactions on Signal Processing. 1991;39(2):411-423.
23. Alku P, Vilkman E, Laine UK. Analysis of glottal waveform in different phonation types using the new IAIF-method. In Proc. 12th Int. Congress Phonetic Sciences. 1991;4:362-365.
24. Childers DG, Lee CK. Vocal quality factors: Analysis, synthesis, and perception. the Journal of the Acoustical Society of America. 1991;90(5):2394-2410.
25. Hanson HM. Glottal characteristics of female speakers: Acoustic correlates. The Journal of the Acoustical Society of America. 1997;101(1):466-481.
26. Kreiman J, Gerratt BR, Antoñanzas-Barroso N. Measures of the glottal source spectrum. Journal of Speech, Language, and Hearing Research. 2007;50(3):595-610.
27. Stelzig Y, Hochhaus W, Gall V, Henneberg A. Laryngeal manifestations in patients

with Parkinson disease. Laryngo-rhino-otologie. 1999;78(10):544-551.

28. Jiang J, O'Mara T, Chen HJ, Stern JI, Vlagos D, Hanson D. Aerodynamic measurements of patients with Parkinson's disease. Journal of Voice. 1999;13(4):583-591

29. Solomon NP, Hixon TJ. Speech breathing in Parkinson's disease. Journal of Speech, Language and Hearing Research. 1993; 36(2):294-310.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history/25827