IberSPEECH 2021
24-25 March 2021, Valladolid, Spain

New tools for the differential evaluation of Parkinson's disease using voice and speech processing

Laureano Moro-Velazquez[1,2]*, Jorge A. Gomez-Garcia*[2]*, Najim Dehak*[1]*, Juan I. Godino-Llorente*[2]

[1]Center for Language and Speech Processing, Johns Hopkins University, Baltimore, USA.
[2]Bioengineering and Optoelectronics lab (ByO), Universidad Politécnica de Madrid, Madrid, Spain.

laureano@jhu.edu

Abstract

Parkinson's Disease (PD) is a neurodegenerative condition that affects the motor capabilities of individuals. Early detection can potentially contribute to slow its progression in a near future. Therefore, new objective and reliable tools are needed to support its diagnosis. Literature suggests that the patients' speech can provide relevant information about the presence of the disease. In this study, five sets of experiments were carried out, each containing new approaches to detect the presence of the disease in the speech of idiopathic PD patients and control speakers from three parkinsonian corpora, two of them in the Spanish language. Different speech frame selection techniques are proposed, such as phonemic and acoustic landmark distillation, providing certain specific speech segments of interest to this work's purposes. Multiple cepstral and spectral features were employed, along with several classification techniques based on Gaussian models and speaker embeddings. The best accuracy results in detecting PD with the proposed methodologies reached values ranging from 85% to 94% with Area Under the Curve between 0.91 and 0.99, depending on the corpus. Results suggest that PD affects the movements related to all of the studied articulatory segmental groups but has a more evident influence in the consonants with a greater narrowing of the vocal tract, mainly plosives and fricatives. The new proposed methodologies demonstrate their ability to support PD's diagnosis during a patient's clinical assessment and are a step forward in PD's speech-based diagnosis systems.

Index Terms: Parkinson's Disease, speaker recognition, GMM-UBM, phonemic distillation, acoustic landmarks

1. Introduction

In this document we present an overview of the thesis [1] of the first author titled "Towards the differential evaluation of Parkinson's Disease by means of voice and speech processing"[1] to the Iberspeech 2020 *Ph. D. Thesis Special Session*.

Parkinson's Disease (PD) affects to 1% of the population over the age of 60 in industrialized countries and, with increasing life expectancy, it is expected to affect to more than 9 million people in industrialized nations by 2030 [2]. A slower advance of the disease in patients will diminish its impact on their daily activities and increase their quality of life. It will also greatly reduce the economic burden of PD of health-care systems. Unfortunately, the diagnosis is based on the assessment of a patient's signs and symptoms over an observation period which can last from months to years. Since PD is a motor system disorder, the analysis of a patient's movements while performing a complex motor task can lead to the identification of potential biomarkers. But, which type of tasks can be used within the analysis? Multiple studies have found enough evidences to propose biomarkers or automatic diagnosis schemes based on a patient's eye movements [3], or handwriting [4], among others. Speech can also be used to evaluate PD because it involves coordination and precision of movements in mainly the laryngeal and articulatory muscles [5].

Multiple studies have analyzed the phonatory, articulatory, prosodic and linguistic aspects of parkinsonian speech, and of them have proposed the use of artificial intelligence to provide diagnostic and assessment tools. Typical approaches include voice quality measurements (noise, frequency and amplitude perturbations, etc.) and classifiers such as Support Vector Machine (SVM) or Random Forest to identify patients with PD employing sustained phonations [6, 7]. Others propose the use of speech characterizations such as filterbank features or Mel-Frequency Cepstral Coefficients (MFCC) as input to SVM, Gaussian Mixture Model (GMM) and, in some cases, neural networks [8, 9]. An extensive review of these approaches and a list of common features and corpora can be found in [10, 11].

The purpose of this study is to propose new approaches to support the clinical diagnosis of PD by employing speech as the object of observation and to obtain new information about the influence of PD in different articulatory movements.

2. Hypothesis and goals

The phonatory and articulatory aspects of the patients with idiopathic PD are affected by the motor dysfunctions associated with this neurodegenerative disease [5]. Our hypothesis is that the resulting speech impairments, inaccuracy of articulatory movements and modified patterns of acceleration and velocity of the articulators can be characterized using signal processing techniques. This characterization, combined with machine learning classification schemes employed in speaker and speech recognition, can yield new tools for the diagnosis of PD.

The main goal of this study is to analyze distinct advanced speech processing technologies and machine learning techniques for the detection and assessment of PD from the speech's articulatory and phonatory aspects. Some specific objectives derived from the main goal are: 1) the analysis of supervised schemes to detect PD from phonatory aspects of speech; 2) the analysis of several supervised and unsupervised new schemes to detect PD from articulatory aspects of speech; 3) the analysis of the role of the distinct phones and articulatory manners (plosives, fricatives, nasals and liquids) in the automatic detection of PD; 4) the analysis of the distinct transitions between phones and their influence in the detection of PD; 5) the identification of the advantages and disadvantages of the different speech tasks in the proposed schemes to detect PD; 6) the analysis of the combination of a phonatory and articulatory subsystems to au-

[1]Available at:
http://oa.upm.es/51278/1/LAUREANO_MORO_VELAZQUEZ.pdf

10.21437/IberSPEECH.2021-36

tomatically detect PD; 7) the study of articulators' kinematics and its relevance in the automatic detection of PD; 8) The study of the generalization properties of the proposed systems.

3. Methodology

In this study we propose new approaches for the automatic detection of PD employing speech technologies which had not been exploited before for this task. Additionally, a combination of the articulatory and phonatory aspects is explored with the aim of using the potential complementary capabilities of the correspondent used methodologies for the automatic detection of the disease. Within each experiment, different approaches or variations of the same basis scheme were proposed in order to find the optimal solution. Six speech corpora, described in Section 4, were employed to carry out the experiments.

3.1. Experimental set 1: Speaker recognition technologies

In this experiment [12], several speaker recognition techniques were applied and adapted for the automatic detection of PD using the patient's speech. Three families of features were considered, MFCC, Rasta-Perceptual Linear Prediction (PLP) and Linear Predictive Coding (LPC), along with their respective derivatives, utilizing multiple configurations. Equally, two classification techniques, namely Gaussian Mixture Model-Universal Background Model (GMM-UBM) and i-Vectors with Gaussian Probability Linear Discriminant Analysis (GPLDA), were used to train and test the automatic detectors. The objective of this study was twofold: firstly to evaluate the application of these techniques to a new scenario, analyzing their different degrees of freedom to establish a baseline to compare results with further studies; and secondly, to evaluate the influence of kinetic changes of instantaneous coefficients and the importance of the time window used to estimate the derivatives for the detection of PD. A large amount of trials were performed using a single parkinsonian and a single auxiliary corpus. The configuration leading to the best results was tested again with the other two remaining parkinsonian corpora. Finally, cross-corpora trials were performed to validate the methodology at the optimum configuration.

3.2. Experimental set 2: Forced Gaussians

In this second experiment [13], several approaches using GMM are proposed. In this experiment, a Forced Alignment Model (FAM) was built to automatically segment and align all the frames on the speech in the parkinsonian and Universal Background Model (UBM) corpora. The resulting phonetic labels were employed to build GMM-UBM models containing Gaussians which were specific for each phonetic label (forced-GMM), yielding models able to compare all the phonetic units of the parkinsonian and control speakers more precisely. The different trials were performed in the three parkinsonian corpora. Finally, a group of cross-corpora trials was performed to validate the optimal configurations. These experiments allowed us to obtain more precise PD detectors and to evaluate which phones were more relevant in the automatic detection of PD.

3.3. Experimental set 3: Phonemic distillation

As several works in the literature point out, the consonants produced after a strong constriction of the articulators are usually more affected by PD [14, 15, 16]. The idea behind the third experiment [17, 18] was to use only specific segments of the speech, depending on the manner of articulation and the narrowing of the vocal tract to analyze their influence in the detection of PD. The speech signals from the different corpora were segmented and labeled (when possible) by using forced alignment techniques. The segments containing a single phone category depending on the manner of articulation (fricative, liquid, nasal, plosive and vowels) were used to train and test separately GMM-UBM classifiers (one classifier per manner) employing PLP as features. Thus, the obtained models to detect PD were focused on these different phonetic groups. The different trials were performed in three parkinsonian corpora. In this case, a fusion of scores coming from the models trained using different specific segments was also carried out. Finally, a group of cross-corpora trials was performed to validate the optimal configurations.

3.4. Experimental set 4: Acoustic landmark distillation

In this fourth articulatory experiment [19], different speech segments related to relevant articulatory moments, such as bursts, transitions between vowels and consonants or the beginning and end of glottal activity, were used to identify relevant transition segments which were employed to detect PD, using an acoustic landmark distillation. This experiment is similar to Experiment 3 but in this case, instead of using whole phonetic segments, only the transitions are used. To identify the acoustic landmarks and, thus, the transition segments, the methodology proposed by [20] was followed. In this case, new GMM-UBM classifiers were trained and tested using *burst*, *sonorant*, and *glottal* transition segments separately. Then, a fusion of scores from

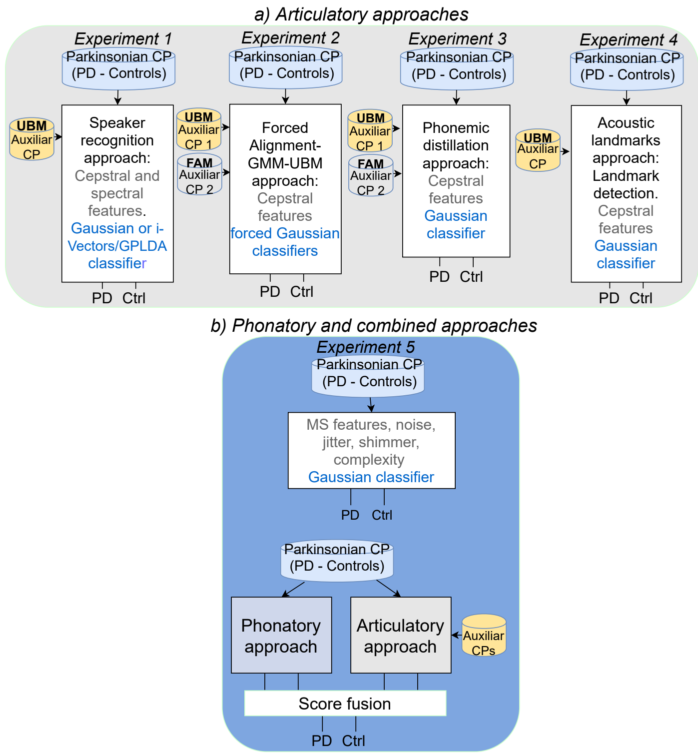

Figure 1: *Overview of the methodology employed in the study, including the four experiments on the articulatory approaches (a), the phonatory approaches, and the combination of the phonatory approach with one articulatory approach (b), where the articulatory part corresponds to the methodology providing the best results among the four articulatory experiments. CP stands for 'corpus'.*

the different models (one per segment type) was performed to combine systems.

3.5. Experimental set 5: Phonatory analysis and score fusion

A large portion of the approaches found in literature use jitter, shimmer and noise to detect PD [11] but not many of them propose new features for this task. In the experimental set 5 [21], a new group of features based on the use of Modulation Spectrum (MS) [22, 23][2] were used in a GMM classification scheme to automatically discriminate patients from controls in three different corpora. These features, containing information about noise, voice frequency and amplitude perturbations, and tremor, were considered as appropriate for the proposed task. To finish the experiments, a fusion of scores between the phonatory approach and the articulatory approach yielding the best results in terms of accuracy and Area Under the Curve (AUC) was performed in order to combine the two aspects into a single approach.

A scheme of all of the experiments studied in this study and the final combination of aspects can be found in Figure 1.

4. Corpora

In this study, six corpora were employed: Neurovoz [13], GITA [24], CzechPD [25], Albayzin [26], Vystadial-Czech [27] and FisherSP [28]. The first three corpora contain different speech tasks from parkinsonian and control speakers and were used to train and test the models proposed in the methodology. The three latter corpora are considered auxiliary and were employed to train UBMs and FAMs. The mother tongue of the subjects in GITA, Neurovoz, Albayzin and FisherSp is the Spanish language, whereas CzechPD and Vystadial-Czech contain the speech from Czech speakers.

The Neurovoz corpus[3] is a new speech dataset recorded in this study. It contains 47 parkinsonian and 32 control speakers whose mother tongue is Spanish Castillian. This corpus was recorded in collaboration with the otorhinolaryngology and neurology departments of the Gregorio Marañón hospital in Madrid, Spain.

5. Results and discussion

Table 1 contains the best results of the five experimental sets for the parkinsonian corpora employing different speech tasks: Text-Dependent Utterance (TDU), Diadochokinetic (DDK), monologues and sustained vowel /a:/. Table 2 includes the cross-corpora results for experimental sets 1, 2, 3, and 4. In this last case, each parkinsonian corpus was used to test the models trained with the other two parkinsonian datasets. Best results per corpus are shaded and best overall results are in bold. The first experimental set is considered the baseline of this study and provides maximum accuracy results between 81% and 91%, with AUC between 0.88 and 0.94 depending on the employed corpus and speech task.

The forced-GMM techniques proposed in the experimental set 2 can be considered novel on PD detection, and these allow us to observe the influence of PD in each of the individual phonetic units. The results suggest that phonetic units requiring a higher narrowing of the vocal tract but without a burst tend to be more influential in the detection of PD, while nasal consonants are the less influential. The results suggest that the phones /C/, /D/, /g/ and /R/ tend to produce better accuracy, and /m/ and /B/ are less relevant in the detection of PD.

Additionally, other two new techniques were proposed in this study: phonemic distillation (experimental set 3) and acoustic landmark distillation (experimental set 4). While both techniques are based on frame selection, the first one uses specific speech segments related to the manner of articulation, such as plosives or fricatives, while the second one is focused in the selection of certain types of transitions (from silence to plosive phone or from vowel to consonants). The use of phonemic or landmark distillation in the UBM produces improvements respect to the baseline (Experimental set 1) in the three corpora and with almost any speech task. Attending to Tables 1 and 2, the results of Experimental set 3 (phonemic distillation) tend to outperform the rest of the approaches in the k-folds cross-validations and in the cross-corpora validations. Cross-validation results are reported at speaker level. In the cross-validation process, the speakers considered in the training folds were not included in the testing folds.

Concerning the cross-corpora tests (carried out using only

Table 1: *Best cross-validation results for each experiment as a function of the employed corpus and speech task. Sust. v. stands for sustained vowel, and Exp. for Experimental set.*

Corpus	Speech task	Exp.	Accuracy ± CI	AUC	Sens.	Spec.
GITA	TDU	1	80 ± 8	0.85	0.82	0.78
		2	81 ± 8	0.88	0.84	0.78
		3	$\mathbf{85 \pm 7}$	**0.91**	0.82	0.88
		4	85 ± 7	0.89	0.84	0.86
	DDK	1	81 ± 8	0.88	0.82	0.8
		2	79 ± 8	0.86	0.86	0.72
		3	83 ± 7	0.89	0.86	0.8
		4	83 ± 7	0.88	0.86	0.8
	Monol.	1	80 ± 8	0.88	0.76	0.84
		2	78 ± 8	0.84	0.73	0.82
		3	82 ± 8	0.89	0.8	0.84
		4	80 ± 8	0.86	0.76	0.84
	Sust. v.	5	71 ± 9	0.8	0.72	0.7
	TDU + Sust. v.	5	85 ± 7	0.91	0.82	0.88
Neurovoz	TDU	1	86 ± 8	0.93	0.87	0.84
		2	81 ± 9	0.87	0.83	0.78
		3	$\mathbf{89 \pm 7}$	**0.93**	0.87	0.91
		4	$\mathbf{89 \pm 7}$	**0.93**	0.91	0.84
	DDK	1	79 ± 9	0.85	0.87	0.65
		2	79 ± 8	0.86	0.86	0.72
		3	86 ± 8	0.88	0.89	0.81
		4	83 ± 7	0.88	0.86	0.8
	Monol.	1	79 ± 12	0.81	0.59	0.9
		2	66 ± 14	0.67	0.35	0.83
		3	77 ± 12	0.79	0.53	0.9
		4	79 ± 12	0.9	0.47	0.97
	Sust. v.	5	64 ± 10	0.68	0.75	0.48
	TDU + Sust. v.	5	87 ± 7	0.94	0.85	0.91
CzechPD	DDK	1	88 ± 1	0.94	0.85	0.93
		2	94 ± 1	0.97	0.9	1
		3	94 ± 1	0.98	0.9	1
		4	$\mathbf{94 \pm 1}$	**0.99**	0.9	1

[2]MS features code available at: https://github.com/jorgomezga/AVCA-ByO

[3]Data and more information about this corpus are available at https://zenodo.org/record/3557758

Table 2: *Best cross-corpora results for all the experiments. In all cases, speech task is DDK.*

Test corpus	Exp.	Accuracy ± CI	AUC	Sens.	Spec.
GITA	1	73 ± 9	0.82	0.84	0.62
	2	66 ± 9	0.76	0.9	0.42
	3	75 ± 8	0.84	0.86	0.64
	4	73 ± 9	0.81	0.8	0.66
Neurovoz	1	75 ± 10	0.82	0.8	0.65
	2	74 ± 10	0.78	0.87	0.5
	3	81 ± 9	0.83	0.91	0.62
	4	76 ± 10	0.83	0.78	0.73
CzechPD	1	79 ± 14	0.91	1	0.5
	2	76 ± 14	0.87	0.95	0.5
	3	82 ± 13	0.95	0.85	0.79
	4	82 ± 13	0.95	1	0.57

DDK task as it is the only common task in the three corpora) the best AUC is never reduced more than 0.05 absolute points respect to the k-folds trials, as it is observed in Table 2. The use of cross-corpora trials –i. e. classifiers trained with a corpus and evaluated with a different one– for the automatic detection of PD is a novelty by itself since, to the authors of this study's knowledge, no published work had succesfully performed this type of trials before.

Globally, the highest accuracy obtained in the trials is 94% with AUC of 0.99, sensitivity of 0.90 and specificity of 1.00, in the CzechPD corpus using acoustic landmark distillation on the UBM, Experimental set 4. It is important to consider that this corpus mainly contains newly diagnosed patients. Therefore, results suggest that the new proposed approaches can be valid detecting PD in early stages. The best results obtained in the k-folds cross-validation in the three corpora are always over 85%, as shown in Table 1.

In this study four types of speech tasks were considered: TDU, DDK tasks, monologues and sustained vowel /a/. TDU provided the best results in almost every trial, as it can be inferred from Table 1. In this respect, a detailed analysis of the results reveals that the impact of PD on speech is not limited to a few articulatory movements, phones or transitions, but influences them all in a higher or lower degree. TDU contain more variety of phones and articulatory movements than DDK tasks, justifying the differences between the results provided by both tasks. On the other hand, although monologues include a larger variety of articulatory movements than DDK, these are different for each utterance in the training and testing sets and the resulting models are text-independent. Text-dependent models demonstrate to outperform text-independent approaches as the first ones allow to compare more precisely the same specific segments (transitions, phones, acceleration or velocity of articulatory movements, etc).

These experiments were carried out employing only idiopathic PD patients and controls, and no other parkinsonism or neurodegenerative disease such as Friedrich's ataxia or Huntington's Disease (HD) has been considered. Consequently, it is not possible to evaluate if the proposed approaches provide discrimination between idiopathic PD and other neurodegenerative diseases, important for early diagnosis, which will be carried out in future studies. The analysis of results from male and female subjects separately, creating their respective specific models would have been desirable. However, larger corpora are advisable for these purposes.

Finally, the analysis performed in this work has been complemented in further studies employing other state-of-the-art speech and speaker recognition methodologies such as end-to-end Automatic Speech Recognition (ASR) models [29], x-vectors [30] or Long Short-Term Memory (LSTM) networks [31], providing new approaches and insights about how PD affects the speech of patients.

6. Conclusions

In this study, new approaches to support the differential evaluation of PD by means of voice and speech processing have been proposed and analyzed.

One of the first conclusions obtained from the analysis of results is that kinematic changes, characterized by means of the derivatives of PLP coefficients of speech are crucial in the detection of PD using speech. These provide information about acceleration and velocity of the articulators during speech, two characteristics that are highly influenced by the disease in the forms of hypokinesia or other types of kinetic impairments such akinesia or bradykinesia.

Also, since the motor dysfunctions caused by PD produce misarticulation, results suggest that these deficits can be found in different types of segments of speech (i.e, distinct groups of phones and transitions). The articulatory movements requiring a higher narrowing of the vocal tract, that normally occur during the pronunciation of fricatives and plosives, are the most influenced by the disease. However, these are not the only segments of speech containing information about PD since schemes based only on vowels provide significant results too. Results suggest that PD influences all the studied phonetic groups, affecting to the articulatory sequence as a whole. For this reason, the analysis of phonetically balanced speech tasks allows to evaluate the presence of PD from speech by using automatic detectors. Additionally, when employing TDU as speech tasks, the obtained classification models are text-dependent and allow to compare more precisely the articulation of patients and controls since all the speakers repeat the same sequence of phonemes.

Regarding the two studied speech aspects, the discriminatory properties of the proposed phonatory approaches are quite limited in comparison with the articulatory approaches. The combination of these approaches yields better results in some cases but the advantages of this combination are unclear.

Finally, the best results obtained with the proposed methodologies reach accuracy values ranging from 85% to 94% with AUC between 0.91 and 0.99 and sensitivity between 0.82 and 0.91 depending on the parkinsonian corpus considered. These results are obtained employing the proposed phonemic distillation technique in the UBM corpus in a GMM-UBM classification scheme. These values can be considered close to the maximum feasible accuracy which is estimated to be around 95% due to possible misdiagnosis in the corpora [13]. These true limits have not been considered before in other works performing automatic detection of PD. The proposed approaches exhibit a lower accuracy in the cross-corpora trials with respect to the cross-validation (k-folds) trials but still between 75% and 82% with AUC between 0.84 and 0.95 and sensitivity between 0.85 and 1, depending on the testing corpus, indicating that these approaches generalize in scenarios with different recording conditions.

7. Acknowledgements

This work was supported by the Government of Spain under grants TEC2012-38630-C04-01 and DPI2017-83405-R1 and by the MISTI Global Seed Funds Award ("Objective Evaluation of PD and other parkinsonisms from the speech" project).

8. References

[1] L. Moro-Velázquez, "Towards the differential evaluation of parkinson's disease by means of voice and speech processing," Ph.D. dissertation, 2018. [Online]. Available: http://oa.upm.es/51278/1/LAUREANO_MORO_VELAZQUEZ.pdf

[2] E. R. Dorsey *et al.*, "Projected number of people with parkinson disease in the most populous nations, 2005 through 2030," *Neurology*, vol. 68, no. 5, pp. 384–386, 2007.

[3] F. Chan, I. T. Armstrong, G. Pari, R. J. Riopelle, and D. P. Munoz, "Deficits in saccadic eye-movement control in parkinson's disease," *Neuropsychologia*, vol. 43, no. 5, pp. 784–796, 2005.

[4] P. Drotár, J. Mekyska, I. Rektorová, L. Masarová, Z. Smékal, and M. Faundez-Zanuy, "Decision support framework for parkinson's disease based on novel handwriting markers," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 3, pp. 508–516, 2015.

[5] J. R. Duffy, *Motor speech disorders: Substrates, differential diagnosis, and management*. Elsevier Health Sciences, 2013.

[6] J. Mekyska, Z. Galaz, Z. Mzourek, Z. Smekal, I. Rektorova, I. Eliasova, M. Kostalova, M. Mrackova, D. Berankova, M. Faundez-Zanuy *et al.*, "Assessing progress of parkinson's disease using acoustic analysis of phonation," in *Bioinspired Intelligence (IWOBI), 2015 4th International Work Conference on*. IEEE, 2015, pp. 111–118.

[7] J. Rusz, J. Hlavnička, T. Tykalová, M. Novotný, P. Dušek, K. Šonka, and E. Růžička, "Smartphone allows capture of speech abnormalities associated with high risk of developing parkinson's disease," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 8, pp. 1495–1507, 2018.

[8] M. Novotny, J. Rusz, R. mejla, and E. Rka, "Automatic evaluation of articulatory disorders in parkinson s disease," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 9, pp. 1366–1378, 2014.

[9] J. Vásquez-Correa, T. Arias-Vergara, J. Orozco-Arroyave, and E. Nöth, "A multitask learning approach to assess the dysarthria severity in patients with parkinson's disease," *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, pp. 456–460, 2018.

[10] L. Moro-Velazquez and N. Dehak, "A review of the use of prosodic aspects of speech for the automatic detection and assessment of parkinson's disease," *Automatic Assessment of Parkinosnian Speech. Communications in Computer and Information Science (CCIS) Series*, vol. 1295, pp. 53–73, 2020.

[11] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, and J. I. Godino-Llorente, "Advances in parkinson's disease detection and assessment using voice and speech: A review of the articulatory and phonatory aspects," *Biomedical Signal Processing and Control*, vol. 66, p. 102418, 2021.

[12] L. Moro-Velazquez, J. A. Gomez-Garcia, J. I. Godino-Llorente, J. Villalba, J. R. Orozco-Arroyave, and N. Dehak, "Analysis of speaker recognition methodologies and the influence of kinetic changes to automatically detect parkinsonś disease," *Applied Soft Computing*, vol. 62, pp. 649–666, 2018.

[13] L. Moro-Velazquez, J. A. Gomez-Garcia, J. I. Godino-Llorente, J. Villalba, J. Rusz, S. Shattuck-Hufnagel, and N. Dehak, "A forced gaussians based methodology for the differential evaluation of parkinson's disease by means of speech processing," *Biomedical Signal Processing and Control*, vol. 48, pp. 205–220, 2019.

[14] J. A. Logemann and H. B. Fisher, "Vocal Tract Control in Parkinson's Disease," *Journal of Speech and Hearing Disorders*, vol. 46, no. 4, p. 348, 1981.

[15] J. Kegl, H. Cohen, and H. Poizner, "Articulatory consequences of Parkinson's disease: perspectives from two modalities." *Brain and Cognition*, vol. 40, no. 2, pp. 355–86, 1999.

[16] J. Godino-Llorente, S. Shattuck-Hufnagel, J. Choi, L. Moro-Velázquez, and J. Gómez-García, "Towards the identification of idiopathic parkinson's disease from the speech. new articulatory kinetic biomarkers," *PloS one*, vol. 12, no. 12, p. e0189583, 2017.

[17] L. Moro-Velazquez, J. Gomez-Garcia, J. I. Godino-Llorente, J. Rusz, S. Skodda, F. Grandas, J.-M. Velazquez, J. R. Orozco-Arroyave, E. Noth, and N. Dehak, "Study of the automatic detection of parkison's disease based on speaker recognition technologies and allophonic distillation," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 1404–1407.

[18] L. Moro-Velazquez, J. A. Gomez-Garcia, J. I. Godino-Llorente, F. Grandas-Perez, S. Shattuck-Hufnagel, V. Yagüe-Jimenez, and N. Dehak, "Phonetic relevance and phonemic grouping of speech in the automatic detection of parkinson's disease," *Scientific Reports*, vol. 9, no. 1, pp. 1–16, 2019.

[19] L. Moro-Velazquez, J. Godino-Llorente, J. Gómez-García, J. Villalba, S. Shattuck-Hufnagel, and N. Dehak, "Use of acoustic landmarks and gmm-ubm blend in the automatic detection of parkinson's disease," in *Models and Analysis of Vocal Emissions for Biomedical Applications: 10th International Workshop, december, 13-15, 2017*, vol. 117. Firenze University Press, 2017, p. 73.

[20] K. N. Stevens, "Toward a model for lexical access based on acoustic landmarks and distinctive features," *The Journal of the Acoustical Society of America*, vol. 111, no. 4, pp. 1872–1891, 2002.

[21] L. Moro-Velázquez, J. A. Gómez-García, N. Dehak, and J. I. Godino-Llorente, "Analysis of phonatory features for the automatic detection of parkinson's disease in two different corpora," *Models and Analysis of Vocal Emissions for Biomedical Applications (MAVEBA) 2019*, p. 33, 2019.

[22] L. Moro-Velázquez, J. A. Gómez-García, J. I. Godino-Llorente, and G. Andrade-Miranda, "Modulation spectra morphological parameters: A new method to assess voice pathologies according to the grbas scale," *BioMed research international*, vol. 2015, p. 259239, 2015.

[23] L. Moro-Velazquez, J. A. Gomez-Garcia, and J. I. Godino-Llorente, "Voice pathology detection using modulation spectrum-optimized metrics," *Frontiers in bioengineering and biotechnology*, vol. 4, p. 1, 2016.

[24] J. Orozco-Arroyave and J. Arias-Londoño, "New Spanish speech corpus database for the analysis of people suffering from Parkinson's disease." *Proceedings on the International Conference on Language Resources and Evaluation (LREC)*, pp. 342–347, 2014.

[25] J. Rusz *et al.*, "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.

[26] A. Moreno, D. Poch, A. Bonafonte, E. Lleida, J. Llisterri, J. B. Mariño, and C. Nadeu, "Albayzín speech database: Design of the phonetic corpus," in *Eurospeech 1993. Proceedings of the 3rd European Conference on Speech Communication and Technology*, vol. 1. ISCA, 1993, pp. 175–178.

[27] M. Korvas, O. Platek, O. Dusek, L. Zilka, and F. Jurcicek, "Free english and czech telephone speech corpus shared under the cc-by-sa 3.0 license," in *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014)*, 2014, pp. 4423—-4428.

[28] "Fisher spanish speech." [Online]. Available: https://catalog.ldc.upenn.edu/LDC2010S01

[29] L. Moro-Velazquez, J. J. Cho, S. Watanabe, M. A. Hasegawa-Johnson, O. Scharengorb, H. Kim, and N. Dehak, "Study of the performance of automatic speech recognition systems in speakers with parkinson's disease," *Interspeech 2019*, pp. 3875–3879, 2019.

[30] L. Moro-Velazquez, J. Villalba, and N. Dehak, "Using x-vectors to automatically detect parkinson's disease from speech," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1155–1159.

[31] S. Bhati, L. Moro-Velazquez, J. Villalba, and N. Dehak, "Lstm siamese network for parkinson's disease detection from speech," in *2019 IEEE Global Conference on Signal and Information Processing (GlobalSIP)*. IEEE, 2019, pp. 1–5.