Hallmarks of Parkinson’s disease progression determined by temporal evolution of speech attractors in the reconstructed phase-space

(Article begins on next page)

Hallmarks of Parkinson's disease progression determined by temporal evolution of speech attractors in the reconstructed phase-space

1st Federica Amato
Poliytechnic University of Turin
Turin, Italy
federica.amato@polito.it

2nd Valerio Cesarini
University of Rome Tor Vergata
Rome, Italy
valerio.cesarini@uniroma2.eu

3rd Luca Pietrosanti
University of Rome Tor Vergata
Rome, Italy
pietrosanti@ing.uniroma2.it

4th Giovanni Costantini
University of Rome Tor Vergata
Rome, Italy
costantini@uniroma2.it

5th Gabriella Olmo
Poliytechnic University of Turin
Turin, Italy
gabriella.olmo@polito.it

6th Giovani Saggio
University of Rome Tor Vergata
Rome, Italy
saggio@uniroma2.it

***Abstract*— Parkinson's disease (PD) is one of the most widespread neurodegenerative diseases worldwide, affected by a number of alterations, among which speech impairments that, interestingly, manifests up to 10 years before other major evidences (e.g. motor impairments). In this regard, we investigated the feasibility of a model based on the temporal evolution of speech attractors in the reconstructed phase space to identify hallmarks of PD identification and progression. To this end, the adopted dataset was made of vocal emissions of 46 de-novo and 54 mid-advanced People with PD, plus 113 healthy counterpart. A statistical analysis was applied to test the identified hallmarks effectiveness for diagnostic support, monitoring, and staging of the disease. According to the obtained results, the adopted approach of considering the temporal evolution of speech attractors in the reconstructed phase-space results effective to discriminate among the three groups of pathological or healthy voices.**

Keywords*— *Parkinson's Disease, Speech Analysis, Speech Attractors, Automatic assessment

I. INTRODUCTION

Parkinson's Disease (PD) is a long-term second most common (after Alzheimer) neurodegenerative disorder, expected to affect 9 million people by 2030 [1]. Motor symptoms (e.g., bradykinesia, tremor, rigidity...) are among the most common evidences of the disease [2] but approximately 75-90% of People with PD evidence abnormal speech too [3]. The measure and evaluation of such alterations can support early diagnosis and monitoring of PD, as a broad body of literature witnesses by implementing different metrics involving temporal, frequency, and mel-cepstrum analysis [4], [5], [6], [7] and/or accounting of non-linearities of vocal signals [8], [9]. Both normophonic and non-normophonic speakers indeed present non-linear phenomena occurring during voice production, mainly due to the pressure flow in the glottis, stress-strain curves of vocal fold tissues, and vocal fold collision [5], even further worsened by compensatory movements that People with PD tend to perform for lowering motor disfunctions [5]. To inspect the non-linear phenomena, speech signal is commonly represented in the state space according to embedded procedure [5] (Fig.1) from which several features can be estimated and employed to quantify the impairment. Among the most common features are: (i) Correlation dimension (i.e., a measure of the space filled by the points in the state space); (ii) Largest Lyapunov exponent (i.e., a representation of the average divergence rate of neighbor trajectories in the state space); (iii) Hurst exponent (i.e., a measure of the long term dependencies in a time series); (iv) Recurrence Period Density Entropy (i.e., a measure of the periodicity of the given signal) [10], [5]. According to a recent work [10] in which we a aimed to present a through literature review, especially towards feature-based methodology, non-linear measures proved their pivotal importance in evaluating PD-related alterations. Moreover, despite being less frequent than common acoustic features such as Fundamental Frequency (F0), Jitter, and Shimmer, DFA and RPDE mesures proved to among the most effective features in discriminating between Healthy Controls (HC) and People with PD. Despite these features are strongly effective, they also rely on very complex and computationally expensive algorithms. Moreover, the time evolution of speech signals is a very important aspect to differentiate between PD and HC subjects but, to the best of our knowledge, no study specifically investigated the temporal evolution of speech trajectories in the reconstructed phase space as a measure for PD hallmarks identification. In this context, we propose a model based on a 3-D geometry and its time evolution aiming at extracting information related to the speaker's health status. Through a detailed analysis of the volume, its variations between adjacent time windows, as well as the speed at which these changes occur, we aim to obtain a detailed description of the subject's vocal alterations, and to model the difficulty in performing fine and rapid movements, which it is well know to be a pivotal aspect of PD alterations [11]. Moreover, by using a dataset that includes both early-drug naïve (said de-novo) and mid-advanced People with PD, we aimed at identifying relevant correlations between the extracted features and the patient disease stage.

II. MATERIALS

We enrolled 100 People with PD (54 mid-advanced and 46 de-novo) and 113 age- and gender-matched healthy controls (HCs, age 70.3±10.3). Participants were recruited ar the IRCCS Neuromed Instituite and at the Department of System Medicine, Tor Vergata University of Rome, Italy.The patients were selected according to standardized diagnostic criteria by experienced neurologists. Patients with medium-advanced PD (age 72.1±8.1, UPDRS 3.1 2.7±0.6) were recorded in *off L-Dopa* state (at least 12 hours after the last drug intake), while *de-novo* patients (age 64.2±8.6, UPDRS 3.1 0.9±0.7) were drug-naive subjects (i.e., never underwent L-DOPA medication). Selection criteria for HCs and PDs included: (i) Italian native speakers; (ii) 18+ years; (iii) no previous history of smoking; (iv) no respiratory, gastro-esophageal, auditory, or vocal fold disease. The vocal samples were recorded in a silent and echo-free room using a dynamic WH20 microphone (Shure, USA) connected to a high-quality, uncompressed H4n (Zoom, Tokyo, Japan) voice recorder (.wav, 16-bit, 44.1 kHz). Data analysis was performed in Python: Topological Signal Processing (*Teaspoon*) library was applied to identify the most suitable parameters for the voice embedding procedure, the *Alphashape* and *Trimesh* libraries were used to calculate and parameterize the shape alpha-geometries derived from the reconstructed speech attractors.

The entire procedure was conducted following the Helsinki Declaration and approved by the institutional ethics committee (approval number 0026508/2019). Written informed consent was obtained; the demographic and clinical data were noted anonymously.

III. METHODS

A. Data Collection

One participant at a time was instructed to sit with the back and arms adhering to a chair [12] and to sustain the vowel /e/ as long as possible without efforts and at a comfortable volume, the microphone 5 cm from the mouth, according to an effective validated protocol [13], [14], [12] Conveniently, the sustained vowel task is non-influenced by linguistic confounding factors so to benefit a worldwide procedure.

B. Speech signal embedding

According to the embedding theorem originally proposed in [15], the set of diffeomorphic attractors generated in the state space as a solution of a system of nonlinear differential equations can be represented as

$$X(k) = \{x(k), x(k+\tau), \dots, x(k+(\theta-1)\tau)\} \quad (1)$$

where $X = \{X(k)\}$ is the set of points of the attractors, $x(k)$ is the original time signal, τ is the time delay estimated to assure minimum correlation among state variables, and θ is the dimension of the embedding space [16].

To ensure a proper reconstruction of vocal signal dynamics whilst minimizing the redundancy of information, a method based on the auto-mutual information is usually employed to set the value of τ [5]. According to this procedure, the most suitable time delay for a given signal can be computed as the first minima in the auto-mutual information function [5], [16]. As for the embedding dimension, we decided to set θ=3 to maintain a low complexity of the system and investigate the type and the quality of information retrieved from a 3-D representation of the vocal signal. It is worth noting that, to reduce the presence of noise in the reconstructed trajectories, for each k in the set of points of the attractor we computed $X(k)$ applying 50-samples moving average.

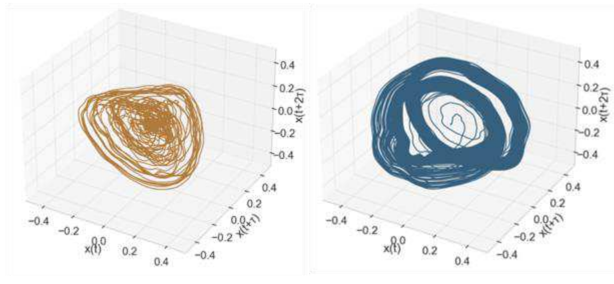


Fig. 1. Example of reconstructed attractor for a PD patient (a) and a HC (b)

C. $\alpha - Shape$ computation

α-Shapes have been employed in several fields to generalize bounding polytopes containing a set of points [17], [18]. According to the original definition proposed in [19] by Edelsbrunner et al., the α-shape of a given set of points S can be seen as the straight line graph whose vertices are the α-extreme points and whose edges connect the respective α-neighbors. A point p in S is defined as an α-extreme if there exists a closed disc of radius 1/α, such that this latter contains all the points in S. Similarly, two α-extreme points are considered α-neighbors if there exists a closed disc of radius 1/α such that they lay on its boundary and the geometry contains all the points in S.

On these basis, it is possible to define an α-Shape such that its boundaries, containing all the points of the reconstructed attractors, describe the smallest volume filled by the set of trajectories in the phase space. It is worth noting that α-Shape geometries are usually reconstructed as an ensemble of various triangles, whose overall geometry can be described by means of their edges and vertices. Taking advantage of this representation of the reconstructed attractors, namely triangular meshes, it is possible to study the speech signal in the phase space through lightweight and efficient algorithms. Here, we evaluated an α-Shape solid for each reconstructed attractor by empirically setting α=30. Fig. 2 shows an example of an α-Shape triangular mesh computed from the reconstructed attractor of an HC subject (the same as in Fig.1.b)

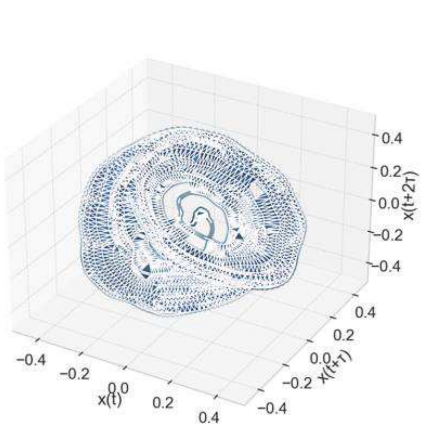


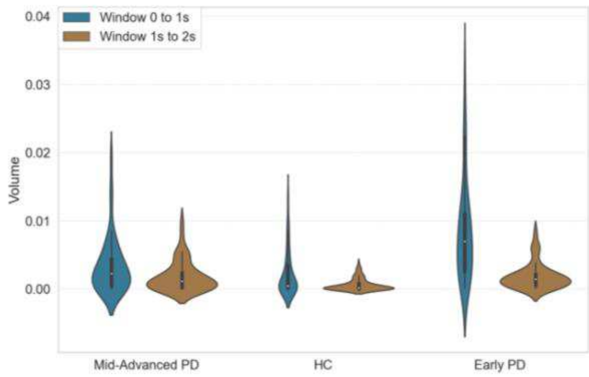
Fig. 2. $\alpha - Shape$ for a HC subject (as in Fig1.b)

D. Speech features evaluation

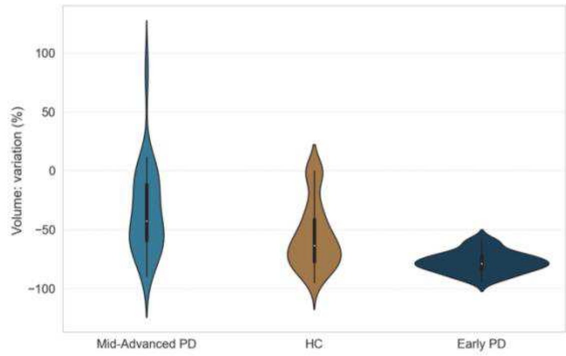
We propose a set of both novel and previously validated features to describe the speech attractors in the reconstructed phase space. As previously demonstrated [5], [10], the volume occupied by the points in the reconstructed phase space retains crucial information on the original signal. Higher points densities are associated to higher chaotic systems: the more regular the voice production system is, the more the trajectories of attractors tend to overlap and converge toward a predefined pattern. In this work, we measured the volume occupied by the set of points as the volume of the associated α-Shape mesh. Given our aim (evidence temporal evolution of speech attractors to derive information about vocal impairment), volume measurements in adjacent speech frames are included too. We selected two equal length windows: from 0 to 1 second ($\text{Volume}_{0\text{-}1s}$) and from 1 second to 2 seconds ($\text{Volume}_{1s\text{-}2s}$). This split enabled, on one hand, to retain for each window enough samples to properly reconstruct an accurate α-Shape mesh, and on the other hand, to study the different phenomena that occur during the voice production process. Since signals in the employed dataset presented various durations, we decided to use for subsequent analysis only recordings longer than 3s, to avoid the final decay of the phonation that could bias the results. It is worth noting that the first window includes the voice signal attack phase, which should exhibit more chaotic behavior even in normo-phonic speakers. Once the initial transient phase is exhausted, the signal is expected to exhibit a more predictable behavior and the reconstructed trajectories evolve to a predefined pattern. Therefore, we propose two measures to model the temporal evolution of speech attractors: the volume variation between adjacent windows (ΔVolume) and the distance between the two α-Shapes triangular meshes. To compute the latter, we aligned the two consecutive tri-meshes using the principal axes of inertia as a starting point and measured the distance between the two geometries as the average square distance per point included on the surface of the 3-D object. In this way, we are able not only to quantify the variation of volumes but also the changes in the overall geometry. Finally, to further describe the 3D geometry of the reconstructed attractor, we evaluated whether the model is watertight (WT) (i.e., a closed surface that does not contain a hole). In fact, in the presence of more chaotic structures or lower recurrence periods, the points of the attractors tend to accumulate at the center of the 3-D geometry, resulting in a convex solid.

E. Feature importance analysis

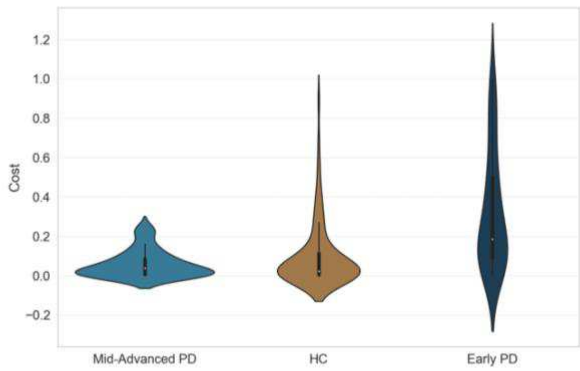
The effectiveness of the employed features was studied through two subsequent steps. As for volume and distance measures, we first analyzed feature distributions and trends using violin plots. Thereafter, we applied Kruskal-Wallis statistical test to investigate whether the features can highlight significant differences between the classes. To further dive into the physical meaning of each feature, we performed additional statistical tests between paired groups: (i) HC vs People with PD (both early and mid-advanced) to assess the effectiveness of the approach in modeling speech alterations; (ii) HC vs early PD to test the capability of the model to identify early markers for neurodegeneration; (iii) early vs mid-advanced PD to evaluate the capability of the approach in differentiating different stages of the disease. As



(a)



(b)



(c)

Fig. 3. Violin plots represent the feature distribution for each of the investigated features. (a): $\text{Volume}_{0\text{-}1}$ and $\text{Volume}_{0\text{-}2;}$ (b): ΔVolume; (c) Distance between α-Shapes in the two temporal windows analyzed

for WT, due to the categorical and binary nature of this feature, we counted the number of occurrences of watertight solids for each window and each class and investigated the presence of recurring 3-D geometries associated with the nature of the original signal. It is worth nothing that, prior to the feature importance analysis, we applied an outlier removal step and selected all those instances included within the 20^{th} and the 80^{th} percentile.

IV. RESULTS

Fig. 3 shows the violin plots associated to the features investigated in this study. Fig.3a highlights the difference between $\text{Volume}_{0\text{-}1s}$ and $\text{Volume}_{1s\text{-}2s}$, the comparison is

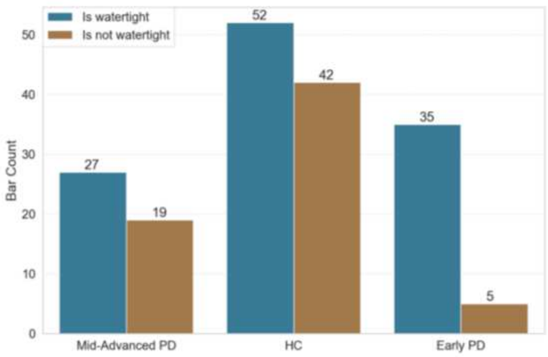


Fig.4 Bar plots representing the proportion between watertight and non-watertight solids for each class.

reported for each class. Fig.3b evidences the percentage of variation of the α-Shape volumes. Finally, Fig.3c shows the distance between the α-Shape geometry in window1 and window2.

Fig.4 reports the bar plot representing the number of α-Shape geometries presenting a watertight structure. For the sake of conciseness, we reported the results only for the second investigated window, given that no significant variation related to this feature was observed when changing the time window analyzed.

A Kruskal-Wallis test was calculated to evidence whether the continuous features employed can highlight statistically significant differences between HC, early PD, and mid-advanced PD. Table I shows the results of the statistical analysis performed.

V. DISCUSSION

Our results confirmed the feasibility of an analysis based on the reconstructed speech attractors to differentiate between HC and People with PD. The feature distribution analysis reported in Fig.3a revealed increased attractor volumes in People with PD. The evidence is in accordance with previous similar works and confirms that an impaired voice production process leads to highly chaotic attractors which tend not to converge to predictable patterns. In addition, altered attractors' temporal evolution are demonstrated to be associated with People with PD' vocal samples. However, this phenomenon seems to diversely affect early and mid-advanced subjects. In the former group, we observed higher volumes in the first window, which drastically reduced in the adjacent segment. On the other hand, mid-advanced patients presented with a lower volume in the first windows that remains almost stable for the entire signal duration. As for the overall temporal evolution, Fig. 3a, Fig. 3b, and Fig. 3c reveal different trends for HC, early PD, and mid-advanced PD. HC subjects show a negative volume variation and a small but non-negligible cost measure between the two windows.

This latter can be associated with the tendency of the attractor points to rapidly accumulate on the same trajectories with no significant differences in the overall 3D geometry. On the contrary, early and mid-advanced PD subjects present increased volumes for each time window analyzed. Going into more detail about the subtle difference between early and mid-advanced PD, we observed comparable volumes of the attractors in the second window. However, our results revealed a higher capability of early patients to rapidly exhaust the effects of the attack phase (similarly to HCs) which is not present in advanced patients, which exhibit almost stable chaotic behavior. The observed differences between early and mid-advanced PD may be due to different reasons. Primarily, whilst mid-advanced People with PD are recorded in OFF-condition, it is known that the effect of L-Dopa on PD symptoms does not completely vanish after 12 hours. This residual medication could lead to a partially reduced chaoticity during the initial transient. On the other hand, the capability of early People with PD to perform changes after the initial transient phase, may be due to a preserved capability to perform fine and rapid movements (similarly to the HC group), which is well known to be strongly affected after the arising of PD. As for the 3-D geometry of the $\alpha - Shape$, our results (Fig. 4) demonstrated that early People with PD have a strong tendency to produce non-watertight solids. From a physical point of view, this result points to wider trajectories being described by the set of points of the attractors associated with the speech of early drug-naive patients. The results of the Kruskal-Wallis test reported in Table I revealed that both the punctual and the differential features associated to the volume, as well the cost measure, present significant differences between HC, early PD, and mid-advanced PD. To further dive into this evidence, we performed additional statistical tests between paired groups. According to our results punctual volume measures are generally effective and including the initial attack phase can be more helpful when the aim is to differentiate between disease stages. As for differential volume measures, they generally resulted in statistically significant results ($p<0.001$) and outperformed punctual features especially in the comparison between HC and mid-advanced PDs. As for the cost function, the statistical test revealed significant differences between the compared populations ($p<0.001$) except when considering HC and mid-advanced PD ($p = 0.86$).

TABLE I. KRUSKAL-WALLIS TEST RESULTS

Kruskal-Wallis test results	**Feature name**	**p-value**
HC vs Early PD vs Mid-Advanced PD	$Volume_{0\text{-}1s}$	**< 0.001**
	$Volume_{1s\text{-}2s}$	**< 0.001**
	ΔVolume	**< 0.001**
	Cost	**< 0.001**
HC vs PD (Early PD and Mid-Advanced PD)	$Volume_{0\text{-}1s}$	**< 0.001**
	$Volume_{1s\text{-}2s}$	**< 0.001**
	ΔVolume	0.63
	Cost	0.0018
HC vs Early PD	$Volume_{0\text{-}1s}$	**< 0.001**
	$Volume_{1s\text{-}2s}$	**< 0.001**
	ΔVolume	**< 0.001**
	Cost	**< 0.001**
HC vs Mid-Advanced PD	$Volume_{0\text{-}1s}$	0.064
	$Volume_{1s\text{-}2s}$	0.0012
	ΔVolume	**< 0.001**
	Cost	0.86
Early PD vs Mid-Advanced PD	$Volume_{0\text{-}1s}$	**< 0.001**
	$Volume_{1s\text{-}2s}$	0.29
	ΔVolume	**< 0.001**
	Cost	**< 0.001**

VI. CONCLUSIONS AND FUTURE WORK

We investigated the feasibility of a tool for evaluating vocal impairment in People with PD through the analysis of attractors' temporal evolution. To ease the parametrization of the reconstructed trajectories in the phase-space, a method based on the α-Shapes was proposed. The derived features resulted to be effective in discriminating healthy and pathological voices, confirming the ability of the reconstructed attractor to describe the presence of vocal alterations. Moreover, our results pointed out the importance of studying attractor's temporal evolution to increase the performance of the algorithms, especially in presence of finer tasks, such as disease staging. Given the promising results of this preliminary work, we intend to further deepen this approach and address the current limitations. First, the current evaluation of the $\alpha - Shape$ is based on an empirically derived value of α. In future work, we plan to automatize the parameter choice and optimization. If properly validated, this approach can be embedded into Machine Learning models improve the study of speech nonlinearities with applications not only to PD but also to other speech affecting diseases, such as dysphonia [20], [21], [22], essential tremor [23], dysphagia [24], and COVID-19 [25], [26]. Moreover, the current results are obtained on a statistically relevant but limited dataset. In future studies we plan to increase the dataset numerosity and perform a more adequate stratification among classes to test the robustness of the proposed results. Finally, in the present work only binary or multiclass analyses were performed. In future studies we plan to introduce the information on the disease staging (e.g., UPDRS) and study the capability of the proposed features to perform a finer classification between people with PD.

VII. REFERENCES

[1] E. R. Dorsey *et al.*, «Projected number of people with Parkinson disease in the most populous nations, 2005 through 2030», *Neurology*, vol. 68, fasc. 5, pp. 384–386, gen. 2007, doi: 10.1212/01.wnl.0000247740.47667.03.

[2] G. Di Lazzaro *et al.*, «Technology-Based Objective Measures Detect Subclinical Axial Signs in Untreated, de novo Parkinson's Disease», *JPD*, vol. 10, fasc. 1, pp. 113–122, gen. 2020, doi: 10.3233/JPD-191758.

[3] G. Defazio, M. Guerrieri, D. Liuzzi, A. F. Gigante, e V. di Nicola, «Assessment of voice and speech symptoms in early Parkinson's disease by the Robertson dysarthria profile», *Neurol Sci*, vol. 37, fasc. 3, pp. 443–449, mar. 2016, doi: 10.1007/s10072-015-2422-8.

[4] F. Asci, G. Costantini, G. Saggio, e A. Suppa, «Fostering Voice Objective Analysis in Patients with Movement Disorders», *Movement Disorders*, vol. 36, fasc. 4, pp. 1041–1041, apr. 2021, doi: 10.1002/mds.28537.

[5] J. R. Orozco-Arroyave, *Analysis of speech of people with Parkinson's disease*. in Studien zur Mustererkennung, no. Band 41. Berlin: Logos Verlag GmbH, 2016.

[6] G. Costantini *et al.*, «Artificial Intelligence-Based Voice Assessment of Patients with Parkinson's Disease Off and On Treatment: Machine vs. Deep-Learning Comparison», 2023.

[7] F. Asci *et al.*, «Machine-Learning Analysis of Voice Samples Recorded through Smartphones: The Combined Effect of Ageing and Gender», *Sensors*, vol. 20, fasc. 18, p. 5022, set. 2020, doi: 10.3390/s20185022.

[8] J. R. Orozco-Arroyave, J. D. Arias-Londoño, J. F. Vargas-Bonilla, e E. Nöth, «Analysis of Speech from People with Parkinson's Disease through Nonlinear Dynamics», in *Advances in Nonlinear Speech Processing*, T. Drugman e T. Dutoit, A c. di, in Lecture Notes in Computer Science, vol. 7911. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 112–119. doi: 10.1007/978-3-642-38847-7_15.

[9] P. Henriquez, J. B. Alonso, M. A. Ferrer, C. M. Travieso, J. I. Godino-Llorente, e F. Diaz-de-Maria, «Characterization of Healthy and Pathological Voice Through Measures Based on Nonlinear Dynamics», *IEEE Trans. Audio Speech Lang. Process.*, vol. 17, fasc. 6, pp. 1186–1195, ago. 2009, doi: 10.1109/TASL.2009.2016734.

[10] F. Amato, G. Saggio, V. Cesarini, G. Olmo, e G. Costantini, «Machine learning- and statistical-based voice analysis of Parkinson's disease patients: A survey», *Expert Systems with Applications*, vol. 219, p. 119651, giu. 2023, doi: 10.1016/j.eswa.2023.119651.

[11] F. Amato, L. Borzì, G. Olmo, e J. R. Orozco-Arroyave, «An algorithm for Parkinson's disease speech classification based on isolated words analysis», *Health Inf Sci Syst*, vol. 9, fasc. 1, p. 32, dic. 2021, doi: 10.1007/s13755-021-00162-8.

[12] G. Saggio e L. Sbernini, «New scenarios in human trunk posture measurements for clinical applications», in *2011 IEEE International Symposium on Medical Measurements and Applications*, Bari, Italy: IEEE, mag. 2011, pp. 13–17. doi: 10.1109/MeMeA.2011.5966651.

[13] A. Suppa *et al.*, «Voice in Parkinson's Disease: A Machine Learning Study», *Front. Neurol.*, vol. 13, p. 831428, feb. 2022, doi: 10.3389/fneur.2022.831428.

[14] G. Saggio e G. Costantini, «Worldwide Healthy Adult Voice Baseline Parameters: A Comprehensive Review», *Journal of Voice*, p. S0892199720303283, ott. 2020, doi: 10.1016/j.jvoice.2020.08.028.

[15] F. Takens, «Detecting strange attractors in turbulence», in *Dynamical Systems and Turbulence, Warwick 1980*, D. Rand e L.-S. Young, A c. di, in Lecture Notes in Mathematics, vol. 898. Berlin, Heidelberg: Springer Berlin Heidelberg, 1981, pp. 366–381. doi: 10.1007/BFb0091924.

[16] A. C. Lindgren, M. T. Johnson, e R. J. Povinelli, «Speech recognition using reconstructed phase space features», in *2003 IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03).*, Hong Kong, China: IEEE, 2003, p. I-60–3. doi: 10.1109/ICASSP.2003.1198716.

[17] J. D. Gardiner, J. Behnsen, e C. A. Brassey, «Alpha shapes: determining 3D shape complexity across morphologically diverse structures», *BMC Evol Biol*, vol. 18, fasc. 1, p. 184, dic. 2018, doi: 10.1186/s12862-018-1305-z.

[18] W. Zhou e H. Yan, «Alpha shape and Delaunay triangulation in studies of protein-related interactions», *Briefings in Bioinformatics*, vol. 15, fasc. 1, pp. 54–64, gen. 2014, doi: 10.1093/bib/bbs077.

[19] H. Edelsbrunner, D. Kirkpatrick, e R. Seidel, «On the shape of a set of points in the plane», *IEEE Trans. Inform. Theory*, vol. 29, fasc. 4, pp. 551–559, lug. 1983, doi: 10.1109/TIT.1983.1056714.

[20] M. G. Tulics e K. Vicsi, «The automatic assessment of the severity of dysphonia», *Int J Speech Technol*, vol. 22, fasc. 2, pp. 341–350, giu. 2019, doi: 10.1007/s10772-019-09592-y.

[21] G. Costantini *et al.*, «Machine Learning based Voice Analysis in Spasmodic Dysphonia: An Investigation of Most Relevant Features from Specific Vocal Tasks»:, in *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies*, Online Streaming, --- Select a Country ---: SCITEPRESS - Science and Technology Publications, 2021, pp. 103–113. doi: 10.5220/0010344601030113.

[22] A. Suppa *et al.*, «Voice analysis in adductor spasmodic dysphonia: Objective diagnosis and response to botulinum toxin», *Parkinsonism & Related Disorders*, vol. 73, pp. 23–30, apr. 2020, doi: 10.1016/j.parkreldis.2020.03.012.

[23] A. Suppa *et al.*, «Voice Analysis with Machine Learning: One Step Closer to an Objective Diagnosis of Essential Tremor», *Movement Disorders*, vol. 36, fasc. 6, pp. 1401–1410, giu. 2021, doi: 10.1002/mds.28508.

[24] V. Cesarini, N. Casiddu, C. Porfirione, G. Massazza, G. Saggio, e G. Costantini, «A Machine Learning-Based Voice Analysis for the Detection of Dysphagia Biomarkers», in *2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4.0&IoT)*, Rome, Italy: IEEE, giu. 2021, pp. 407–411. doi: 10.1109/MetroInd4.0IoT51437.2021.9488503.

[25] G. Costantini *et al.*, «Deep learning and machine learning-based voice analysis for the detection of COVID-19: A proposal and comparison of architectures», *Knowledge-Based Systems*, vol. 253, p. 109539, ott. 2022, doi: 10.1016/j.knosys.2022.109539.

[26] C. Robotti *et al.*, «Machine Learning-based Voice Assessment for the Detection of Positive and Recovered COVID-19 Patients», *Journal of Voice*, p. S089219972100388X, nov. 2021, doi: 10.1016/j.jvoice.2021.11.004.