Automatic Detection of Parkinson's Disease with Connected Speech Acoustic Features: towards a Linguistically Interpretable Approach

Marta Maffia[1], Loredana Schettino[2] and Vincenzo Norman Vitale[2,3]

[1]Dept. of Literary, Linguistics and Comparative Studies, University of Naples L'Orientale, Italy
[2]Interdepartmental Research Center Urban/Eco, University of Naples Federico II, Italy
[3] DIETI - University of Naples Federico II, Italy

Abstract
Alterations in speech and voice are among the earliest symptoms of Parkinson's Disease (PD). Nevertheless, the rich information carried by patients' speech and voice is only partially used for diagnosis and clinical decision-making that is currently based on holistic ratings of speech intelligibility. An accurate diagnosis could be supported by the application of fully automated analytic methods and machine learning techniques on speech recordings. However, most of the proposed procedures were designed for highly functional but "artificial" vocal paradigms such as sustained phonation and consider all the considerable amount of features that can be extracted using automatic systems. In this work, we perform PD detection trials using features extracted from connected speech rather than isolated speech units. Moreover, we support the adopted machine learning-based methods with linguistic considerations so as to reduce the number of features to some meaningful ones. The main findings highlight that this procedure allows more accurate, economical and, most importantly, interpretable discrimination.

1. Introduction

[1] Parkinson's Disease (PD) is the most common movement disorder and the second most common neurodegenerative disorder worldwide after Alzheimer disease. It affects more than 2-3% of the population aged 65 and over [1, 2].

Caused by the deterioration or loss of dopaminergic neurons in the *substantia nigra* of basal ganglia, PD is generally diagnosed based on clinical criteria, by using a medical individual's history and a physical/neurological exam. The loss of dopamine in the central nervous system, along with the anatomical and physiological changes related to the disease, has an impact on laryngeal, respiratory and articulatory functions of Persons with PD (PwPD). Alterations in speech and voice are in fact among the earliest symptoms of PD, which results in a motor speech disorder called hypokinetic dysarthria [3, 4]. Nevertheless, the rich information carried by patients' speech and voice is only partially used for diagnosis and clinical decision-making, since the Unified Parkinson's Disease Rating Scale (UPDRS), a standardized rating tool used to assess the severity and progression of the pathology, only presents one item (item 3.1) that concerns the evaluation of speech [5]. This item is based on the clinician's perception and mostly considers speech in terms of intelligibility. A deeper understanding of speech and voice phenomena by advanced data analytics methods could be therefore very useful in both the diagnostic phase and in the monitoring of therapy response in PwPD.

2. Speech in Parkinson's Disease

PD-related dysarthria, caused by poor activation and coordination of the muscles involved in speech production, includes a range of alterations, extensively described in experimental studies on different languages [6].

As for the voice quality, a breathy, husky-semiwhisper and hoarse voice is often reported in PwPD, accompanied by vocal tremor, an increase in nasality, reduced voice intensity and constant loudness [7]. Voice quality spectrum was also studied using a deep learning approach applied to differential phonological posterior features for the characterization of pathological PD speech, collected through different tasks and compared to healthy non-modal phonation. [8].

At the segmental level, the decreased amplitude of motility of lips, tongue, and jaw provokes imprecision in the production of consonantal sounds, with the so-called spirantization phenomenon or occlusive weakening [9, 10]. A reduction in the vowel space area and an im-

CLiC-it 2023: 9th Italian Conference on Computational Linguistics, Nov 30 — Dec 02, 2023, Venice, Italy
maffia@unior.it (M. Maffia); loredana.schettino@unina.it (L. Schettino); vincenzonorman.vitale@unina.it (V. N. Vitale)
0000-0002-4913-374X (M. Maffia); 0000-0002-3788-3754 (L. Schettino); 0000-0002-0365-8575 (V. N. Vitale)

[1]This article is the result of the collaboration among the authors. However, for academic purposes, Marta Maffia is responsible for sections 1 and 2, Loredana Schettino for section 3 and Norman Vincenzo Vitale for section 4. All the authors are responsible for section 5.

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

paired and less distinctive formant generation in speech of PwPD have also been described, both in sustained prolongation of single vowels [11] and in continuous speech, such as sentence repetition [12] or reading passage [13]. The centralization of formant values, measured by the Vowel Articulation Index (VAI), was also proposed as a potential early marker of PD, especially when observed in spontaneous speech [14].

As for the suprasegmental aspects, PwPD often report a significantly narrower tonal range (monopitch) or an abnormal pitch variability, along with a compromised ability to consciously manipulate intonation [15, 4]. Articulation and speech rate are also altered in PD, although previous findings do not highlight a uniform pattern of variation in the speech of PwPD: in some studies a reduction in speech rate was observed in PD patients [16], while some reported the opposite effect [17, 18] and other found no intergroup differences between pathological and healthy speech [19]. Furthermore, different rhythmic metrics were used to describe the alteration of rhythm in PD speech, as part of a more “general dysrhythmia” [20]. In recent studies on Italian PD patients, the percentage of vocalic intervals (%V) was found to be effective in characterizing pathological speech, when compared to that of healthy individuals, both in read and spontaneous conditions and even at a very early stage of the disease [21, 22].

In the last decades, in line with the growing interest and efforts in the identification of reliable linguistic and acoustic biomarkers of PD, some studies demonstrated that an accurate diagnosis could be supported by the application of fully automated analytic methods and machine learning techniques on speech recordings [23]. However, most of the proposed procedures were designed for highly functional (but “artificial”) vocal paradigms such as sustained phonation, diadochokinetic tasks, syllable repetition, short sentences [24, 25, 26, 27, 28, 29]. These kinds of elicitation techniques indeed provide highly controlled signals, but such control affects phonation and may even mask features that may emerge in less controlled semi(spontaneous) connected speech. In addition, previous studies often achieve high levels of accuracy in the detection of PD speech by taking into account a very large number of features, and the classification focuses on computational aspects rather than linguistic ones [30].

In this contribution, we address the following issues:

- investigate the role of acoustic features, usually overlooked or, however, not always or directly taken into account by specialists for PD diagnosis;
- consider patterns that emerge from connected (read) speech rather than isolated speech units (phones, syllables, words) productions;
- support machine learning-based methods with linguistic considerations so as to reduce the size of the big sets of features automatically extracted to some meaningful ones and provide an effective linguistic interpretation of the results.

3. Method

3.1. Data and Annotation

The study has been conducted on data from the *Italian Parkinson’s Voice and Speech* corpus [31, 32], which consists of speech data collected through different speech production tasks from three groups of Italian (Apulian) speakers: PD patients, age-matched healthy control (HC) speakers and younger HC speakers.

In particular, we considered a subset of this corpus, consisting in 25 speech samples elicited through a reading task[2] from 15 PD patients and 10 age-matched healthy speakers. Subjects in the PD group are classified by the specialists as <4 on the modified Hoehn and Yahr scale, which stands for a non-severe stage of the severity of their disease. The patients’ speech ability is evaluated following the tips provided in section 3.1 (eloquence) of the Unified Parkinson’s Disease Rating Scale (UPDRS) as minimally/slightly impaired (maximum score is 4 = severe impairment). Demographic and clinical features of patients with PD and HC speakers are resumed in Table 1.

	HC (n=10)	**PD (n=15)**
Age (m±SD)	68±6	64±9
Sex (M/F)	4/6	11/4
H&Y	-	<4
UPDRS (Item 3.1)	-	1.07±1.18

Table 1
Biographical (Sex and Age) characteristics of the PD and HC speakers and clinical data (H&Y: Hoehn & Yahr scale; UPDRS: Unified Parkinson’s Disease Rating Scale) of PD speakers [32].

The considered dataset had already been the object of a spectroacoustic analysis in a previous study [22] and the acoustic signal had been therefore manually segmented and annotated into vowel (V) and consonantal (C) intervals (see Figure 1). Main descriptive statistics of the dataset are reported in Table 2.

3.2. Analysis

In this study, we intend to use the described continuous speech data for PD detection based on a reduced set of interpretable features of the acoustic signal. To this aim,

[2] The reading task was based on a phonemically balanced text [31].

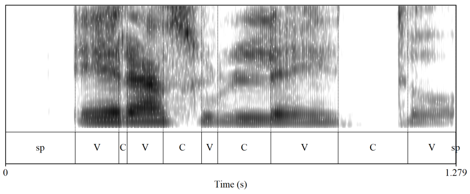

Figure 1: Spectrogram and annotation of the utterance "*era sul letto*", "(Dad) was on the bed". C: consonantal interval, V: vowel interval, sp: silent pause.

	Tot	HC	PD
Total duration (s)	1765	614	1151
Duration of speech portions (s)	1206	455	751
Duration of samples (s) (m±SD)	71±17	61±4	77±20
n. of V intervals	4664	1761	2903
n. of CV intervals	5260	2107	3153
n. of Phonetic Chains	910	312	598

Table 2
Descriptive statistics of the considered subset.

three trials of PD detection were conducted, each time considering a different basic unit, namely:

- Vowels (V) - in previous studies, the percentage of vocalic interval in the speech signal was demonstrated to be informative in PD detection. So we investigate whether vowels alone contain enough information for the detection task;
- Consonant and Vowels (CV) - we extend the context of vowels to the previous consonants, obtaining a wider feature extraction window to evaluate the influence of consonants preceding vowels on PD detection;
- Phonetic Chains (PC) - lastly we employ the phonetic chain, namely the sequence of vowels and consonants between two silent pauses. On the one hand, such units provide the most comprehensive automatically detectable window for feature extraction. On the other hand, being a larger unit of speech production, it should provide far enough features to discriminate speaker status.

Based on the OpenSmile toolkit [33], we selected the eGeMAPSv02 [34] as the basic feature set, and then investigated which features could be considered as the most relevant for discrimination considering previous literature [35] and inspection of the data with the Orange software [36].

Then, the impact of the selected features was evaluated by employing two unsupervised machine-learning techniques:

- The **K-Means**[2] [37] a vector-quantization method which divides n objects in k clusters based on their mean distance.
- **Hierarchical Agglomerative Clustering (HAC)**[2] [38] is a greedy technique that aims at grouping (or splitting) clusters based on a similarity measure. The final output is a clusters hierarchy which could be divided based on the number of desired clusters.

These simple yet efficient techniques were employed to obtain explainable and interpretable results.

The PD detection trials were conducted considering the following sets of features:

- a full feature set, i.e. the eGeMAPSv02 complete feature set (88 features) [34] plus the speakers' sex.
- a subset feature set, i.e. **18** features from the eGeMAPSv02 feature set, plus the sex (see Appendix A).

In both cases, features were normalized at zero mean and unitary variance.

4. Results

The inspection conducted with the Orange software highlighted that the most relevant features for discriminating between PwPD and HC speakers are those concerning the spectral distribution (i.e., slope, alpha ratio, Hammarberg index), followed by those concerning energy and amplitude (i.e. loudness, shimmer), and frequency (MFCC). The observed features were included in the subset employed for the discrimination trials (as reported in Appendix A. Also, the table in Appendix C shows the Mean values and Standard Deviation of these features in PC units per speaker).

Results show that classification based on the Phonetic Chain (see Figure 4) outperforms by far classifiers based on both V and CV. On the one hand, the HAC classifier with the full feature set reaches nearly 99% of true positive detection and 85% of true negative detection. On the other hand, the K-means performs at its best with the feature subset with an 89% of true positive and a 72% of true negative. This means that by reducing the number of features of 75% with respect to the original feature set, the K-means has a 10% reduction in true positive (i.e., PD) detection and a 13% reduction in true negative (i.e., HC) detection, with respect to HAC on the full feature set.

The vowels-based setting (see Figure 2) shows better performances with the feature subset with both K-means

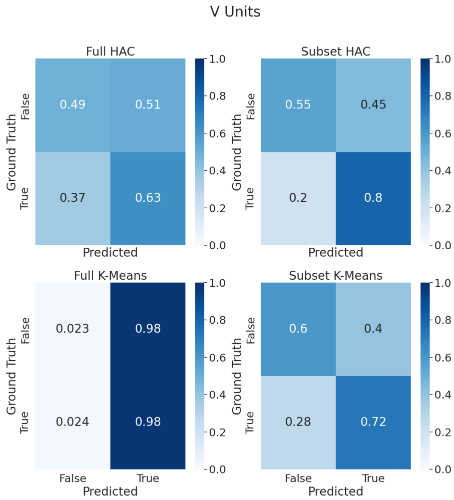

Figure 2: V Clustering.

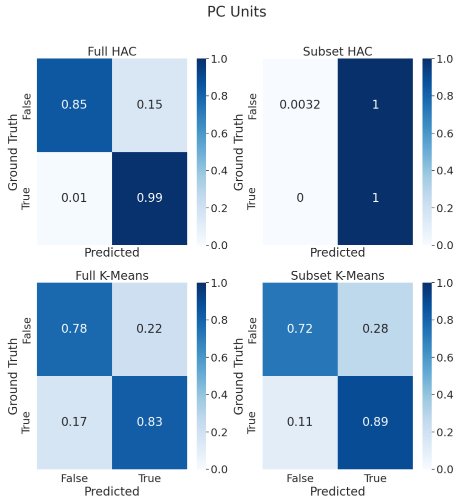

Figure 4: PC Clustering.

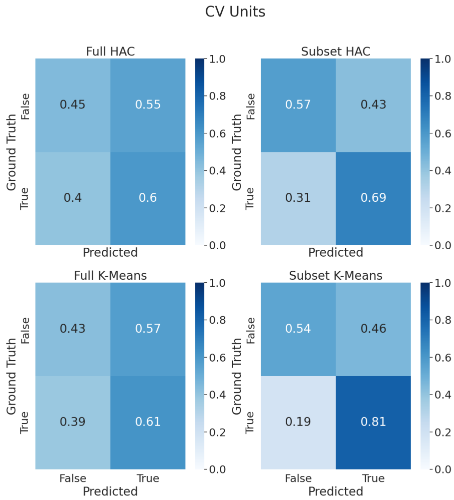

Figure 3: CV Clustering.

and HAC. However, the True negative detection rate is near 60% in the best case, while the true positive rate is at 80% in the best case.

Finally the CV setting (see Figure 3) shows performances which are comparable to a coin toss in most of cases. Only the K-means based on feature subset reaches a true positive detection rate of 81%, with a true negative detection rate of 54%.

In light of these results, we decided to also investigate the correlation between the considered features and the intelligibility score (from the above-described UPDRS) given by the specialists. As illustrated in figure 5, no strong correlation emerges between UPDRS scores and the analysed acoustic features with the exception of slopeV0-500 that negatively correlates with the specialists' ratings (see Appendix B for the correlation matrices concerning the features extracted from V and VC intervals, Figure 7).

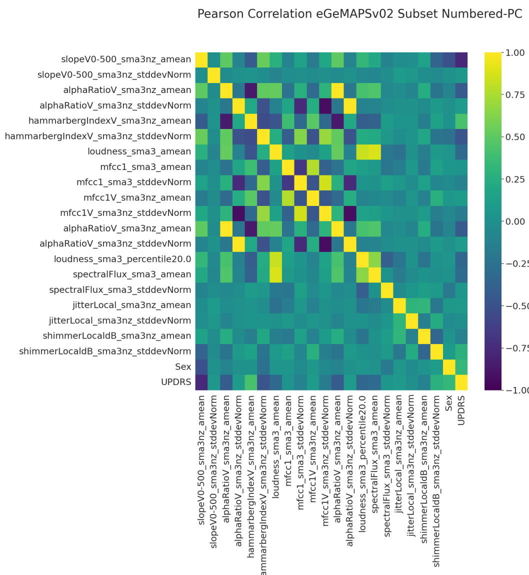

Figure 5: Feature correlation considering PC units.

5. Discussion and Conclusion

The present study provides relevant findings both for the development of PD detection systems and the analysis of Parkinsonian speech characteristics by integrating computational methods with domain-specific linguistic knowledge.

The correlation data between the UPDRS ratings concerning PD speakers' speech ability and the acoustic features automatically extracted from the speech signal corroborate the observation that the specialists' holistic assessment overlooks, or at least only partially and indirectly considers, acoustic features, which, nonetheless, prove to provide crucial information for the diagnosis. In fact, the speech signal is affected by the condition of the muscles involved in phonation. So, if the vocal apparatus is somewhat compromised as an effect of the muscular impairment due to the disease (dysarthria), the signal should show this. Hence, the relevance of including acoustic features in the assessment of the outbreak and severity of PD.

However, fully automated extraction and treatment of speech acoustic features is usually achieved with highly complex systems whose interpretation is quite difficult for both computational scientists, who might be not familiar with PD symptoms and the linguistic value of the features of the speech signal, and for domain experts, who might not be familiar with machine learning methods. Therefore, the design of models in a way that their predictions can be explainable and easily interpretable may actually be most sensible and economical. In fact, this study highlights that not all the possibly considerable acoustic features provide the same amount of information and are actually relevant for discrimination. Moreover, their contribution may vary as a result of the type and span of the linguistic unit used for the feature extraction.

More specifically, the classification results show that considering vowel intervals as units of reference for the features extraction is already quite effective. Most effective is, however, considering wider contexts as provided by the inter-pausal phonetic chain intervals, whereas enlarging the vocalic intervals only to the previous consonant (CV intervals as a basic unit) turns out to be noisy rather than informative.

Then, on average, the feature subset proved to be most informative, carrying out sufficient information to let the classifiers reach a reasonable detection rate in the considered medical scenario. In particular, the subset mainly includes features concerning spectral distribution, followed by those involving energy and amplitude and finally frequency features (MFCC above all).

It is worth noticing that the study has been conducted on continuous speech rather than on isolated phones, syllables or words, to get closer to the normal working dynamic of the vocal apparatus during utterance phonation and avoid artificial effects that may arise when producing single short items.

To conclude, supporting automated analytic methods and machine learning techniques with linguistic considerations allows for more accurate, economical and, most importantly, interpretable discrimination. Future work will be devoted to delving deeper into the linguistic analysis of the way the emergent features characterize PD speech and the investigation of the explainability of classification methods based on deep neural networks.

References

[1] L. M. De Lau, M. M. Breteler, Epidemiology of parkinson's disease, The Lancet Neurology 5 (2006) 525–535.

[2] W. Poewe, K. Seppi, C. M. Tanner, G. M. Halliday, P. Brundin, J. Volkmann, A.-E. Schrag, A. E. Lang, Parkinson disease, Nature reviews Disease primers 3 (2017) 1–21.

[3] F. L. Darley, A. E. Aronson, J. R. Brown, Clusters of deviant speech dimensions in the dysarthrias, Journal of speech and hearing research 12 (1969) 462–496.

[4] S. Pinto, A. Chan, I. Guimarães, R. Rothe-Neves, J. Sadat, A cross-linguistic perspective to the study of dysarthria in parkinson's disease, Journal of Phonetics 64 (2017) 156–167.

[5] C. G. Goetz, B. C. Tilley, S. R. Shaftman, G. T. Stebbins, S. Fahn, P. Martinez-Martin, W. Poewe, C. Sampaio, M. B. Stern, R. Dodel, et al., Movement disorder society-sponsored revision of the unified parkinson's disease rating scale (mds-updrs): scale presentation and clinimetric testing results, Movement disorders: official journal of the Movement Disorder Society 23 (2008) 2129–2170.

[6] J. R. Duffy, Defining, understanding, and categorizing motor speech disorders, Motor speech disorders—substrates, differential diagnosis, and management. 3rd edn. Saint Louis: Elsevier Mosby (2013) 3–13.

[7] D. G. Hanson, B. R. Gerratt, P. H. Ward, Cinegraphic observations of laryngeal function in parkinson's disease, The Laryngoscope 94 (1984) 348–353.

[8] M. Cernak, J. R. Orozco-Arroyave, F. Rudzicz, H. Christensen, J. C. Vásquez-Correa, E. Nöth, Characterisation of voice quality of parkinson's disease using differential phonological posterior features, Computer Speech & Language 46 (2017) 196–208.

[9] H. Ackermann, W. Ziegler, Articulatory deficits in parkinsonian dysarthria: an acoustic analysis., Journal of Neurology, Neurosurgery & Psychiatry 54 (1991) 1093–1098.

[10] D. Duez, Acoustic analysis of occlusive weakening in parkinsonian french speech, in: International Congress of Phonetic Sciences, Université de Saarebrücken, 2007, pp. 1–4.
[11] I. Eliasova, J. Mekyska, M. Kostalova, R. Marecek, Z. Smekal, I. Rektorová, Acoustic evaluation of short-term effects of repetitive transcranial magnetic stimulation on motor aspects of speech in parkinson's disease, Journal of Neural Transmission 120 (2013) 597–605.
[12] S. Sapir, L. O. Ramig, J. L. Spielman, C. Fox, Formant centralization ratio: A proposal for a new acoustic measure of dysarthric speech (2010).
[13] S. Skodda, W. Visser, U. Schlegel, Vowel articulation in parkinson's disease, Journal of voice 25 (2011) 467–472.
[14] J. Rusz, R. Cmejla, T. Tykalova, H. Ruzickova, J. Klempir, V. Majerova, J. Picmausova, J. Roth, E. Ruzicka, Imprecise vowel articulation as a potential early marker of parkinson's disease: Effect of speaking task, The Journal of the Acoustical Society of America 134 (2013) 2171–2181.
[15] A. M. Goberman, C. A. Coelho, M. P. Robb, Prosodic characteristics of parkinsonian speech: The effect of levodopa-based medication, Journal of medical speech-language pathology 13 (2005) 51–69.
[16] C. L. Ludlow, N. P. Connor, C. J. Bassich, Speech timing in parkinson's and huntington's disease, Brain and language 32 (1987) 195–214.
[17] H. Hirose, S. Kiritani, M. Sawashima, Velocity of articulatory movements in normal and dysarthric subjects, Folia Phoniatrica et Logopaedica 34 (1982) 210–215.
[18] H. Ackermann, J. Konczak, I. Hertrich, The temporal control of repetitive articulatory movements in parkinson's disease, Brain and language 56 (1997) 312–319.
[19] S. Skodda, U. Schlegel, Speech rate and rhythm in parkinson's disease, Movement disorders: official journal of the Movement Disorder Society 23 (2008) 985–992.
[20] J. M. Liss, L. White, S. L. Mattys, K. Lansford, A. J. Lotto, S. M. Spitzer, J. N. Caviness, Quantifying speech rhythm abnormalities in the dysarthrias (2009).
[21] M. Maffia, R. De Micco, M. Pettorino, M. Siciliano, A. Tessitore, A. De Meo, Speech rhythm variation in early-stage parkinson's disease: a study on different speaking tasks, Frontiers in Psychology 12 (2021) 668291.
[22] M. Maffia, M. Pettorino, Voce, età e parkinson: questioni di ritmo, in: CLUB Working Papers in Linguistics, volume 6, Alma Mater Studiorum Università di Bologna, 2022, pp. 66–78.
[23] J. Hlavnička, R. Čmejla, T. Tykalová, K. Šonka, E. Ržička, J. Rusz, Automated analysis of connected speech reveals early biomarkers of parkinson's disease in patients with rapid eye movement sleep behaviour disorder, Scientific reports 7 (2017) 12.
[24] J. I. Godino-Llorente, S. Shattuck-Hufnagel, J.-Y. Choi, Moro-Velázquez, J. A. Gómez-García, Towards the identification of idiopathic parkinson's disease from the speech. new articulatory kinetic biomarkers, PloS one 12 (2017) e0189583.
[25] M. Chronowski, M. Klaczynski, M. Dec-Cwiek, K. Porebska, Parkinson's disease diagnostics using ai and natural language knowledge transfer, arXiv preprint arXiv:2204.12559 (2022).
[26] L. Ali, C. Zhu, M. Zhou, Y. Liu, Early diagnosis of parkinson's disease from multiple voice recordings by simultaneous sample and feature selection, Expert Systems with Applications 137 (2019) 22–28.
[27] J. S. Almeida, P. P. Rebouças Filho, T. Carneiro, W. Wei, R. Damaševičius, R. Maskeliūnas, V. H. C. de Albuquerque, Detecting parkinson's disease with sustained phonation and speech signals using machine learning techniques, Pattern Recognition Letters 125 (2019) 55–62.
[28] R. Lamba, T. Gulati, H. F. Alharbi, A. Jain, A hybrid system for parkinson's disease diagnosis using machine learning techniques, International Journal of Speech Technology 3 (2021) 583–593.
[29] C. O. Sakar, G. Serbes, A. Gunduz, H. C. Tunc, H. Nizam, B. E. Sakar, M. Tutuncu, T. Aydin, M. E. Isenkul, H. Apaydin, A comparative analysis of speech signal processing algorithms for parkinson's disease classification and the use of the tunable q-factor wavelet transform, Applied Soft Computing 74 (2019) 255–263.
[30] F. Cordella, A. Paffi, A. Pallotti, Classification-based screening of parkinson's disease patients through voice signal, in: 2021 IEEE International Symposium on Medical Measurements and Applications (MeMeA), IEEE, 2021, pp. 1–6.
[31] G. Dimauro, F. Girardi, Italian parkinson's voice and speech, IEEE Dataport (2019).
[32] G. Dimauro, V. Di Nicola, V. Bevilacqua, D. Caivano, F. Girardi, Assessment of speech intelligibility in parkinson's disease using a speech-to-text system, IEEE Access 5 (2017) 22199–22208.
[33] F. Eyben, M. Wöllmer, B. Schuller, Opensmile: the munich versatile and fast open-source audio feature extractor, in: Proceedings of the 18th ACM international conference on Multimedia, 2010, pp. 1459–1462.
[34] F. Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. André, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan, et al., The geneva minimalistic

acoustic parameter set (gemaps) for voice research and affective computing, IEEE transactions on affective computing 7 (2015) 190–202.

[35] L. Moro-Velazquez, J. A. Gomez-Garcia, J. D. Arias-Londoño, N. Dehak, 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 66 (2021) 102418.

[36] J. Demšar, T. Curk, A. Erjavec, Črt Gorup, T. Hočevar, M. Milutinovič, M. Možina, M. Polajnar, M. Toplak, A. Starič, M. Štajdohar, L. Umek, L. Žagar, J. Žbontar, M. Žitnik, B. Zupan, Orange: Data mining toolbox in python, Journal of Machine Learning Research 14 (2013) 2349–2353. URL: http://jmlr.org/papers/v14/demsar13a.html.

[37] J. A. Hartigan, M. A. Wong, Algorithm as 136: A k-means clustering algorithm, Journal of the royal statistical society. series c (applied statistics) 28 (1979) 100–108.

[38] F. Murtagh, A survey of recent advances in hierarchical clustering algorithms, The computer journal 26 (1983) 354–359.

Appendix A: Further Features Analysis

List of the features included in the considered subset of the eGeMAPSv02 features.

Features concerning the spectral distribution:

- slopeV0-500_sma3nz_amean
- slopeV0-500_sma3nz_stddevNorm
- alphaRatioV_sma3nz_amean
- alphaRatioV_sma3nz_stddevNorm
- hammarbergIndexV_sma3nz_amean
- hammarbergIndexV_sma3nz_stddevNorm
- spectralFlux_sma3_amean
- spectralFlux_sma3_stddevNorm

Features concerning energy and amplitude:

- loudness_sma3_amean
- loudness_sma3_percentile20.0
- shimmerLocaldB_sma3nz_amean
- shimmerLocaldB_sma3nz_stddevNorm

Features concerning frequency:

- mfcc1_sma3_amean
- mfcc1_sma3_stddevNorm
- mfcc1V_sma3nz_amean
- mfcc1V_sma3nz_stddevNorm
- jitterLocal_sma3nz_amean
- jitterLocal_sma3nz_stddevNorm

Appendix B: Further Results

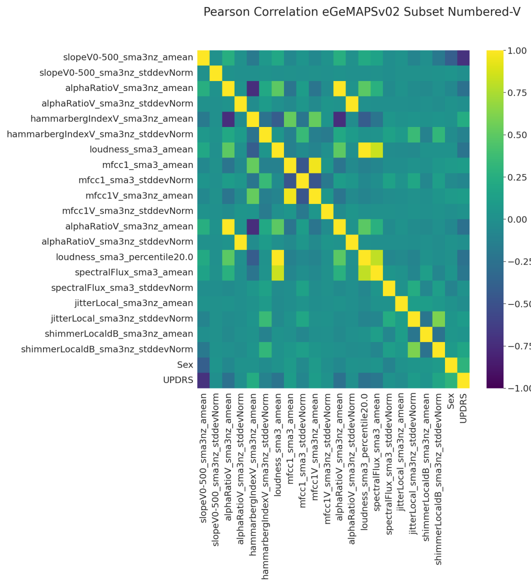

Figure 6: Feature correlation considering V units.

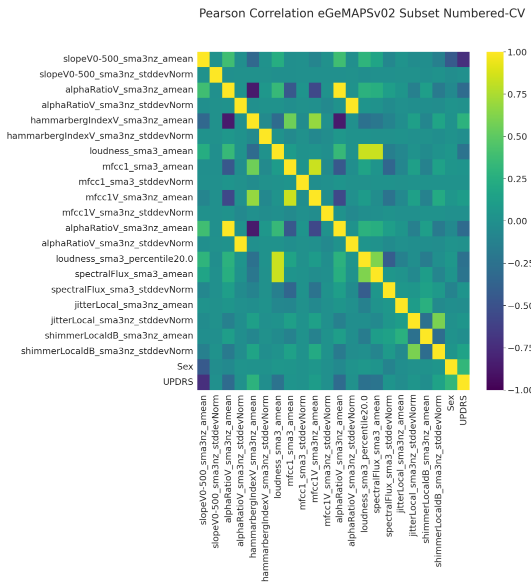

Figure 7: Feature correlation considering V and CV units.

Speaker	Slope	alphaRatio	H-Index	Shimmer	Loudness	MFCC
01PDm	0,0003 ± 0,0159	-21,2933 ± 2,9959	32,5541 ± 3,1311	1,2979 ± 0,2976	0,3965 ± 0,1292	36,3328 ± 5,8431
02PDm	0,0074 ± 0,0111	-21,3531 ± 2,7926	32,3354 ± 3,5172	1,0792 ± 0,2899	0,4698 ± 0,1215	34,4790 ± 5,6093
03PDm	0,0012 ± - 0,0124	-21,6803 ± 3,7769	29,2366 ± 4,5719	1,0642 ± 0,1860	0,5931 ± 0,1269	38,3075 ± 5,1538
04PDf	0,0083 ± 0,0107	-25,8529 ± 3,5913	37,0023 ± 3,0799	0,7232 ± 0,1815	0,3373 ± 0,0763	33,3890 ± 5,8998
05PDf	0,0283 ± 0,0102	-16,0985 ± 2,3628	26,4895 ± 1,8925	0,9988 ± 0,1635	0,6615 ± 0,1167	29,3979 ± 5,7566
06PDf	0,0080 ± 0,0077	-19,9178 ± 3,7987	32,1550 ± 3,5414	1,2773 ± 0,2916	0,2556 ± 0,0906	30,4934 ± 6,5089
07PDf	0,0322 ± 0,0088	-18,0467 ± 1,7098	28,5596 ± 2,2559	0,9679 ± 0,2658	0,2350 ± 0,0626	30,9480 ± 4,0652
08PDm	0,0020 ± 0,0113	-22,3596 ± 3,0759	31,7079 ± 2,7711	1,4473 ± 0,2784	0,2608 ± 0,0790	34,6625 ± 5,7257
09PDm	0,0002 ± 0,0074	-20,2667 ± 6,1511	29,4351 ± 8,4180	1,0641 ± 0,4419	0,3720 ± 0,1468	33,0758 ± 8,7907
10PDm	0,0006 ± 0,0117	-22,4290 ± 3,8050	31,4671 ± 3,5929	1,1519 ± 0,2819	0,1562 ± 0,0651	26,6755 ± 6,4457
11PDm	0,0074 ± 0,0134	-22,3154 ± 6,3170	31,3600 ± 8,7320	1,0987 ± 0,4362	0,1010 ± 0,0460	24,7227 ± 7,9015
12PDm	0,0000 ± 0,0115	-29,8513 ± 3,0748	40,0691 ± 3,5827	1,1172 ± 0,3162	0,1761 ± 0,0754	30,2178 ± 6,2936
13PDm	0,0051 ± 0,0098	-18,6693 ± 6,1645	25,7068 ± 8,2491	0,9146 ± 0,4247	0,2866 ± 0,1100	33,8904 ± 8,3735
14PDm	0,0185 ± 0,0111	-23,1260 ± 2,3925	32,0836 ± 2,9519	1,2185 ± 0,2678	0,4287 ± 0,1531	32,0228 ± 4,6341
15PDm	0,0090 ± 0,0108	-21,9545 ± 3,0236	30,2927 ± 3,3761	1,1309 ± 0,3094	0,2977 ± 0,0723	32,9187 ± 6,1951
16HCf	0,0633 ± 0,0096	-18,6750 ± 3,0968	27,9920 ± 3,7985	1,2668 ± 0,2783	0,2887 ± 0,1005	29,5724 ± 7,6025
17HCf	0,0926 ± 0,0076	-18,0667 ± 3,9360	25,8280 ± 3,6499	1,1462 ± 0,2630	0,2700 ± 0,0491	30,1793 ± 7,9545
18HCm	0,0711 ± 0,0060	-10,3776 ± 2,8562	21,0355 ± 3,7580	1,2385 ± 0,2646	0,7188 ± 0,1946	22,0280 ± 5,7435
19HCf	0,0783 ± 0,0091	-16,3902 ± 3,5428	27,4936 ± 5,7825	1,2023 ± 0,2160	0,4559 ± 0,1197	30,9898 ± 7,4353
20HCf	0,0846 ± 0,0059	-21,5727 ± 3,8459	34,0136 ± 4,6024	1,6192 ± 0,2362	0,1724 ± 0,0470	33,4514 ± 6,8876
21HCf	0,0716 ± 0,0097	-17,2437 ± 3,2172	26,4872 ± 4,1833	1,0707 ± 0,3338	0,4527 ± 0,1017	29,4803 ± 6,0164
22HCm	0,0662 ± 0,0110	-14,8324 ± 4,1724	25,3376 ± 4,3715	1,5356 ± 0,3035	0,4509 ± 0,1726	32,5912 ± 8,1394
23HCf	0,0948 ± 0,0073	-20,3646 ± 4,4190	29,3404 ± 4,9695	1,1668 ± 0,2565	0,4223 ± 0,0784	33,8704 ± 6,1713
24HCm	0,0715 ± 0,0090	-11,9822 ± 3,3146	22,3366 ± 4,3487	1,3189 ± 0,2434	0,4978 ± 0,1417	33,4114 ± 7,8817
25HCf	0,0876 ± 0,0066	-16,7485 ± 4,4084	27,8112 ± 5,7054	1,2287 ± 0,2130	0,4543 ± 0,0990	27,9522 ± 8,3363

Table 3
Mean value and Standard Deviation of the most relevant features in PC units per speaker.

Appendix C: Individual Variability