Speech Prosody 2022
23-26 May 2022, Lisbon, Portugal

A longitudinal study on Italian speech rhythm in Parkinson's Disease

Massimo Pettorino[1], *Marta Maffia*[1], *Brigitte Bigi*[2]

[1]University of Naples "L'Orientale", Italy
[2]Laboratoire Parole et Langage, CNRS & Aix-Marseille Université, France
mpettorino@gmail.com, mmaffia@unior.it, brigitte.bigi@lpl-aix.fr

Abstract

Parkinson's Disease dysarthria affects the speech motor control, causing alterations at the suprasegmental level of speech. In previous researches, vowel percentage (%V) and the mean interval between two consecutive vowel onset points (VtoV) were effectively used in the synchronic description of the rhythmic variations of Italian PD speech, compared to healthy speech, even at a very early stage of the disease.

This study aims at verifying the early alteration of PD speech rhythm using a diachronic approach. To reach this goal, a corpus of read speech produced by a single PD subject (female, 66 years old) has been collected, consisting of 15 radiophonic speech samples (about 100 s each) on the same topic, recorded between 2001 and 2021.

The speech samples were manually segmented in consonantal and vocalic intervals by means of Praat, allowing the calculation of %V and VtoV.

The results show an alteration of %V values since 2018, two years before the diagnosis and the insurgence of motor symptoms.

Moreover, first results of the application of the automatic segmentation performed by SPPAS on a selection of PD speech samples will also be presented.

Index Terms: Parkinson's Disease, speech rhythm, longitudinal study, manual annotation, automatic annotation, radiophonic speech

1. Introduction

Parkinson's Disease (PD) is recognized as the second most common neurodegenerative disorder of aging after Alzheimer's disease and the most common movement disorder [1]. The gradual degeneration and loss of dopaminergic neurons of the *substantia nigra pars compacta* are associated to the emergence of the classic motor signs and symptoms of PD (i. e. bradykinesia, rigidity, tremor, and postural instability).

A large amount of experimental studies conducted on different languages demonstrated that the physiological and anatomical changes occurring in PD also affect the three major anatomic subsystems governing speech motor control: the respiratory, phonatory, and articulatory systems [2].

The most common speech abnormalities include hypophonia (or reduced loudness), changes in voice quality (breathy and/or harsh voice), and, at advanced stages of the disease, hesitant and disfluent communication [3], [4].

Moreover, PD dysarthria causes alterations at both the segmental and suprasegmental level of speech. From the segmental point of view, a reduction of vowel space area in PD speech has been observed [5], [6] and several studies focused on the amplitude and duration of speech gestures, underlying the occurrence of *target undershoot* [7] in the imprecise articulation of both vocalic and consonantal sounds by PD patients [8], [9].

At the suprasegmental level, PD speech is often characterized by narrow pitch variability and a reduction of the tonal range, in comparison to healthy speech [3], [10], [11].

1.1. Speech rhythm and Parkinson's Disease

A general dysrhythmia, tested in nonverbal rhythmic tasks, such as finger tapping and gait, has been reported in recent researches on PD [12], [13].

In the context of this global disorder, the rhythm of speech is also altered by PD, along with rate and speech/pause ratio. However, while there is an undoubted impact of PD on prosody, literature data are quite heterogeneous: as for the Articulation Rate, some studies observed an increase of AR in PD subjects, compared to healthy speakers (tachylalia) [14], [15]; in other researches, PD speech is described as slower than the healthy one [16], [17]; in some other cases no differences were found in AR between pathological and healthy speeches [18], [19].

Liss and colleagues [20] conducted a study on American English patients with PD and used several rhythmical metrics in order to identify significant variations in PD speech: standard deviation (SD) of vocalic intervals over a sentence (ΔV), the SD of consonantal intervals over a sentence (ΔC), the percentage of vocalic intervals (%V), the rate-normalized SD of vocalic and consonantal intervals (VarcoV and VarcoC), the pairwise variability indices (nPVI-v and rPVI-c), and the AR. In this study, the parameter %V was found to be one of the most effective in characterizing dysarthric speech.

The same range of metrics was also used by Lowit et al. [21], who described the rhythmic performance of a group of British English speakers with PD, comparing reading, spontaneous speech and diadochokinetic tasks. The results of this study demonstrated that the %V measure is a consistent and sensitive metric to differentiate dysarthric and healthy speaker populations across different speech tasks.

In previous studies on Italian subjects with PD, the authors confirmed the usefulness of %V, associated to VtoV, the mean interval between two consecutive Vowel Onset Points (VOPs), for the detection of rhythmic variations in pathological speech, compared to healthy productions [22], [23]. The alteration of the %V parameter, with significant higher values in PD speech, was also observed at the very early stage of the

10.21437/SpeechProsody.2022-11

disease [24]. The metrics VtoV/%V was chosen by the authors because it has the advantage to describe the rhythmic characteristics of a speech sample from a perceptual perspective. The calculation of the mean value of VtoV takes into account the average distance between the signal discontinuities that guide listeners in the perception of rhythm (VOPs) in a specific speech sample. VtoV mean value can be therefore considered as the perceptual counterpart of AR, overcoming the intrinsic and frequent difficulties in the identification of the "traditional" syllables.

The parameter %V, on the other hand, is completely independent of AR. It gives information about the continuity perceived by the listener in a speech sample; the greater %V the greater the perception of continuity: "this is what, in musical terms, goes under the name of *legato*" [23, p. 3,173].

1.2. Synchronic and diachronic studies

Most of the studies on speech and acoustic correlates of PD, as well as of other pathologies, are generally conducted with a synchronic approach: the speech produced by a certain number of patients is usually compared to that produced by a healthy control group, adequately age and gender-matched. The significance of the results of this kind of studies is tested through statistical analyses.

Longitudinal studies are not frequent, instead, for a very understandable reason: it is certainly more difficult to follow the evolution of a person's speech and collect data over time. Yet the immense corpus of the web is now available, even if some critical issues persist. In a previous case study on a very famous PD speaker, for example, the authors tried to collect a corpus of speech samples produced by the Canadian actor Michael J. Fox before and after the diagnosis of the disease. Even if the results of this study appeared to be very promising, with data showing an abnormal increase in %V in the actor's speech since five years before the diagnosis, the corpus could obviously not be homogeneous. It consisted of interviews, readings, informal conversations, recited speech in different communication contexts etc. It is extremely rare to have the same type of speech, produced by the same PD speaker, on the same topic and in the same communicative situation, over a large span of years, before and after the diagnosis and the appearance of the first motor symptoms of the disease. Some peculiar circumstances gave us the possibility to come into contact with very precious PD speech diachronic data.

2. The study

In the present contribution, the application of a diachronic approach to the study of speech abnormalities caused by PD is proposed. The first objective of this research is to verify the early alteration of speech rhythm in a PD subject, through the VtoV/%V metrics described above.

Moreover, the second aim of the research is a preliminary test of the automatic segmentation procedure, performed by SPPAS and applied on a selection of PD speech samples.

2.1. Materials and methods

2.1.1. The corpus and the speaker

In order to reach the first goal of the study, a corpus of diachronic Italian read speech was collected. It includes 15 radiophonic speech samples recorded over 20 years, between 2001 and 2021.[1] The corpus consists in a series of broadcasts in which the same speaker reads texts on religious topics. The recordings are available on the web.[2]

The speaker is an Italian woman, 66 years old, who declared to speak Italian as her mother tongue and to be born in the Campania region (South of Italy). She has been diagnosed with PD without dementia in 2020, when the first motor symptoms (shaking hand) appeared. Until that date, she had no history of previous disorders. She gave written consent to the data collection procedure.

2.1.2. The acoustic analysis

From each of the 15 recordings of the corpus, samples of 100 s were selected and spectro-acoustically analysed by means of Praat [25].

Through the visual inspection of speech spectrograms and waveforms, the speech samples were manually segmented in consonantal and vocalic intervals (C and V). In Figure 1 and Figure 2 examples of the manual annotation are provided, with red lines placed in correspondence with VOPs.

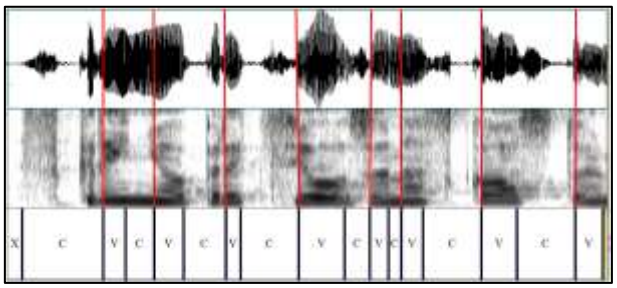

Figure 1: *Spectrogram and manual annotation of the Italian utterance "scrigno prezioso nascosto" (from 2001 sample).*

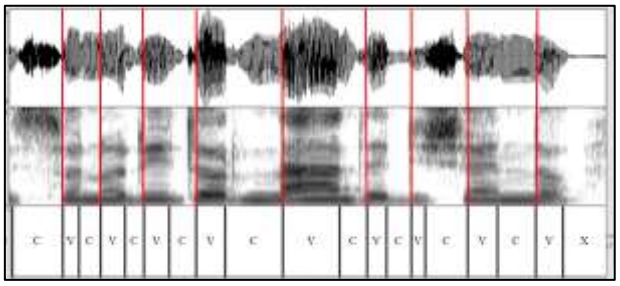

Figure 2: *Spectrogram and manual annotation of the Italian utterance "sono ritornate dicendo" (from 2007 sample).*

The segmentation procedure was the same already used in previous works [22], [24]: approximants were treated as vowels and diphthongs were always considered as a single

[1] As reported in Figure 3 and 4, two recordings were collected for 2019 and two for 2021.

[2] In order to preserve anonymity, links are not reported in the paper. They will be made available by the authors to qualified researchers, on specific and motivated instances.

vocalic interval; in V + nasal consonant sequences, the nasalized portion of the vowel was assigned to the V interval; in the case of initial voiced stop consonants, the first boundary was considered to be the onset of the glottal pulses; finally, post-pausal voiceless plosives were assigned a duration equal to the mean value of single plosives in the same utterance.

Once extracted the durations of all the segments, the values of VtoV and %V for each speech sample were obtained using a Praat script. Disfluencies and silent pauses were also labelled (as D and X, respectively) but they were not taken into consideration for the calculation of these two rhythmic parameters.

The manual labelling of speech samples was conducted by one of the authors, and doubts were discussed and resolved by consensus.

2.2. Results

The results of the analysis are shown in Figure 3. It can be observed in the graph how the VtoV values do not change significantly over the years, being located in an interval between 169 and 207 ms (mean value: 182 ms). The AR, therefore, does not seem to be an effective feature to significantly distinguish speech samples produced in the period preceding the diagnosis (2020) and those following it. A VtoV value equal to 182 ms, in fact, corresponds, from the point of view of the "traditional" syllable, to 5.5 syllables per second: on average, it is a speech produced at a rather rapid speed, however not excessive, which is a quite normal standard for a radio speaker.

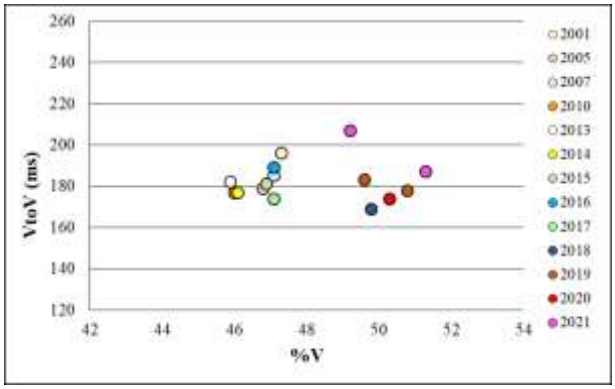

Figure 3: *Mean VtoV (in ms) and %V for all the speech samples in the corpus.*

On the contrary, the variation over the years of the %V values appears to be very interesting, as better highlighted in Figure 4.

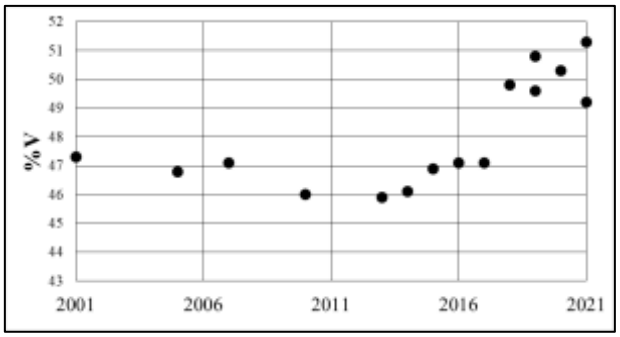

Figure 4: *%V values in all the speech samples from 2001 to 2021.*

As can be noticed, taking into account data from 2001 until 2017, the %V values are located in a range that goes from 45.9 to 47.3 (mean value: 46.7). Starting from 2018 until 2021, a consistent increase in %V is detected, with values that range between 49.2 and 51.3 (mean value: 50.2). These results confirm those already found in a previous synchronic study conducted on early-stage Italian PD speakers: "The threshold value for %V can be estimated at ~ 48%, with the patients with PD having vowel percentages always above it and the control group mostly below it." [24, p. 7]

Moreover, it is important to underline that the rhythmic variation, as in the study on Michael J. Fox's speech, seems to precede the emergence of first motor symptoms: in this case it was already detectable two years before the diagnosis.

2.3. The automatic annotation procedure

If the alteration of %V appears as a very early sign of PD insurgence, it should be emphasized that this change in speech cannot be easily perceived by a clinician or by a phonetician. Variations in %V can be instrumentally detected, instead, as already demonstrated.

These considerations suggest the importance of developing a non-invasive speech-test through which a person can record a specific utterance and receive an immediate feedback about the rhythmic quality of his/her speech. To accelerate the procedure, the speech-test device should be able to analyze the signal, automatically segment it into vocalic and consonantal intervals and then calculate the %V and VtoV values in real time. Such an automatic procedure would be of great help: it would be a support to the clinical diagnosis of PD, allowing the observation and monitoring of speech rhythm variation also in the absence of a phonetician.

As a first step towards this ambitious outcome of the research project, four speech samples were selected from the corpus, two recordings dated before the %V alteration (2010 and 2017) and two dated after it (2018 and 2021). In order to observe if and how the results obtained with an automatic procedure could diverge from those of the manual analysis, these speech samples were automatically annotated by means of the software SPPAS [26] - version 4.0, already used in researches on French dysarthric speech [27].

The automatic annotation was then manually checked by one of the authors of this contribution, who did not analyze the speech samples in the "manual" phase. The results of the two procedures in terms of %V values are reported in Table 1.

Table 1: *%V values obtained with manual and automatic annotation.*

Speech sample	Manual annotation	SPPAS
2010	46	47.7
2017	47.1	46.9
2018	49.8	51.5
2021	51.3	51.7

It is interesting to notice that, even if percentages are not perfectly identical - that would be unrealistic even in the case of two human annotators, the threshold value for the characterization of PD speech (~ 48%) is confirmed by the automatic data.

3. Discussion and conclusions

In order to observe early speech rhythm alterations due to the insurgence of PD, a longitudinal corpus of radiophonic speech samples produced by an Italian single speaker over 20 years was collected and manually analysed. Although the use of a case-study limits the generalization of the findings, the results of the analysis confirm that the %V value alteration could be an important cue for the detection of the disease insurgence, as reported in previous researches based on a synchronic approach. Much earlier than other evident speech disorders (disfluences or hesitations) and even before the first motor symptoms of PD, %V seems to be already altered, with values that exceed the threshold for Italian language of 48%.

A plausible explanation of %V variation in PD subjects can be found in the motor impairments caused by the disease: akinesia, i. e. the difficulty at initiating movements; bradykinesia, the slowing of the velocity in the execution of movements once initiated; muscular rigidity.

Such motor impairments have different effects on the articulation of vocalic and consonantal sounds, because of the different degree of neuromuscular effort required for their production. From the articulatory perspective, in fact, vowels are produced with a static configuration of articulators, without any obstruction of the vocal tract. On the contrary, consonants are dynamic, requiring fast and synchronized movements of the phonatory organs.

As a consequence, in the dysarthric speech, vocalic gestures are found to be sustained once they have been started, with a delay of the consonantal dynamic phase. This prolonging of vowels would be the reason for the greater %V in PD speech, in comparison to healthy speakers.

To reach the second aim of the research, a preliminary step in the direction of an automatic speech-test for the detection of %V variation was made, with extremely interesting results. Even if the segmentation performed by SPPAS still presents some criticalities, with no doubt it already represents a very great advantage for the operator who, manually, can quickly and easily correct any anomaly in the segmentation instead of producing a completely new one. In this phase, the identification of the particular and more critic contextual situations, in which SPPAS does not work as expected, is necessary and propaedeutic to the construction of a new procedure of data elicitation and recording. Furthermore, the quality of the audio signal is certainly a factor to keep in mind when preparing the kit for the speech-test. We are currently working to address all these issues, hoping that such an automatic speech-test for %V calculation could be, in future, a supportive tool for clinicians in the diagnosis of PD.

4. Acknowledgements

The authors would like to thank the speaker involved in this research, for being very helpful in the data collection procedure.

5. References

[1] L. M. de Lau and M. M. Breteler, “Epidemiology of Parkinson’s disease,” *Lancet Neurol.*, vol. 5, pp. 525-535, 2006.

[2] F. L. Darley, A. E. Aronson and J. R. Brown, “Differential diagnosis patterns of dysarthria,” *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246–269, 1969.

[3] F. L. Darley, A. E. Aronson and J. R. Brown, “Cluster of deviant speech dimension in the dysarthrias,” *Journal of Speech and Hearing Research*, vol. 12, no. 3, pp. 462-469, 1969.

[4] L. O. Ramig, C. Fox and S. Sapir, “Speech treatment for Parkinson disease,” *Expert Review of Neurotherapeutics* vol. 8, pp. 299-311, 2008.

[5] S. Skodda, W. Visser and U. Schlegel, “Vowel articulation in Parkinson’s disease,” *Journal of voice : official journal of the Voice Foundation*, vol. 25, no. 4, pp. 467-472, 2011.

[6] S. Sapir, L. O. Ramig, J. L. Spielman and C. Fox, “Formant Centralization Ratio: A proposal for a new acoustic measure of dysarthric speech,” *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 1, pp. 114-125, 2010.

[7] B. Lindblom, “Spectrographic study of vowel reduction,” *J. Acoust. Soc. America*, vol. 35, pp. 1773–1781, 1963.

[8] B. Gili Fivela, M. M. Iraci, V. Sallustio, M. Grimaldi, C. Zmarich and D. Patrocinio, “Italian Vowel and Consonant (co)articulation in Parkinson’s Disease: extreme or reduced articulatory variability?”, *10th International Seminar of Speech Production (ISSP), 5-8 May 2014*, Colon, Germany, pp. 146-149, 2014.

[9] J. Rusz, R. Cmejla, T. Tykalova, H. Ruzickova, J. Klempir and V. Majerova 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.

[10] A. M. Goberman and C. A. Coelho, “Prosodic characteristics of Parkinsonian speech: The effect of levodopa-based medication,” *Journal of Medical Speech-Language Pathology*, vol. 13, pp 51-68, 2005.

[11] M. K. MacPherson, J. E. Huber and D. P. Snow, “The intonation-syntax interface in the speech of individuals with Parkinson's disease,” *Journal of Speech, Language, and Hearing Research*, vol. 54, no. 1, pp. 19-32, 2011.

[12] V. Cochen De Cock, D. Dotov, P. Ihalainen et al., “Rhythmic abilities and musical training in Parkinson’s disease: do they help?” *Parkinson's Disease*, vol. 4, no.8, 2018.

[13] F. Puyjarinet, V. Bégel and C. Gény, “Heightened orofacial, manual, and gait variability in Parkinson’s disease results from a general rhythmic impairment,” *Parkinson's Disease*, vol. 5, no. 19, 2019.

[14] H. Ackermann, J. Konczak and I. Hertrich, “The temporal control of repetitive articulatory movements in Parkinson’s disease,” *Brain and Language*, vol. 56, no. 2, pp. 312-319, 1997.

[15] H. Hirose, S. Kiritani and M. Sawashima, “Velocity of articulatory move-ments in normal and dysarthric subjects,” *Folia phoniatrica*, vol. 34, no. 4, pp. 210-215, 1982.

[16] J. A. Logemann, H. B. Fisher, B. Boshes and E. Blonsky, “Frequency and co-occurrence of vocal tract dysfunction in the speech of a large sample of Parkinson patients,” *Journal of Speech and Hearing Disorders*, vol. 43, pp. 47-57, 1978.

[17] C. L. Ludlow, N. P. Connor and C. J. Bassich, “Speech timing in Parkinson’s and Huntington’s disease,” *Brain and Language*, vol. 32, no. 2, pp. 195-214, 1987.

[18] D. Duez, “Syllable structure, syllable duration, and final lengthening in Parkinsonian French speech,” *Journal of Multilingual Communication Disorders*, vol. 4, pp. 45-57, 2006.
[19] S. Skodda and U. Schlegel, “Speech rate and rhythm in Parkinson's disease,” *Movement Disorders*, vol. 23, no. 7, pp. 985–992, 2008.
[20] J. M. Liss, L. White, S. L. Mattys, K. Lansford, A. J. Lotto, S. M. Spitzer and J. N. Caviness, “Quantifying speech rhythm abnormalities in the dysarthrias,” *Journal of Speech, Language, and Hearing Research*, vol. 52, pp. 1334–1352, 2009.
[21] A. Lowit, A. Marchetti, S. Corson, and A. Kuschmann, “Rhythmic performance in hypokinetic dysarthria: Relationshipbetween reading, spontaneous speech and diadochokinetic tasks,” *Journal of Communication Disorders*, vol. 72, pp. 26-39, 2018.
[22] M. Pettorino, M. G. Busà and E. Pellegrino, “Speech rhythm in Parkinson’s disease: a study on Italian,” *Proceedings of the 17th Annual Conference of the International Speech Communication Association*, San Francisco, USA, pp. 1958–1961, 2016.
[23] M. Pettorino, W. Gu, P. Półrola and P. Fan, “Rhythmic characteristics of parkinsonian speech: a study on mandarin and polish,” *Proceedings of the 18th Annual Conference of the International Speech Communication Association*, Stockholm, Sweden, pp. 3172–3176, 2017.
[24] M. Maffia, R. De Micco, M. Pettorino, M. Siciliano, A. Tessitore and A. De Meo, “Speech rhythm variation in early-stage Parkinson’s disease: a study on different speaking tasks,” *Frontiers in Psychology*, vol. 12, 2021, https://www.frontiersin.org/articles/10.3389/fpsyg.2021.668291/full.
[25] P. Boersma and D. Weenink, *Praat: doing phonetics by computer* [Computer program]. Version 6.1.39, 2021, http://www.praat.org/.
[26] B. Bigi, “SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech,” *The Phonetician*, vol. 111–112, pp. 54–69, 2015.
[27] B. Bigi, K. Klessa, L. Georgeton and C. Meunier, “A syllable-based analysis of speech temporal organization: a comparison between speaking styles in dysarthric and healthy populations,” *Sixteenth Annual Conference of the International Speech Communication Association*, Dresden, Germany, pp. 2977-2981, 2015.