INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Speech Rhythm in Parkinson's Disease: A Study on Italian

Massimo Pettorino[1], *Maria Grazia Busà*[2], *Elisa Pellegrino*[1]

[1] University of Naples L'Orientale, Italy
[2] Padua University, Italy

mpettorino@unior.it, mariagrazia.busa@unipd.it, epellegrino@unior.it

Abstract

Experimental studies on different languages have shown that neurogenetic disorders connected with Parkinson's disease (PD) determine a series of variations in the speech rhythm. This study aims at verifying whether the speech of PD patients presents rhythmic abnormalities compared to healthy speakers also in Italian. The read speech of 15 healthy speakers and of 11 patients with mild PD was segmented in consonantal and vocalic portions. After extracting the durations of all segments, the vowel percentage (%V) and the interval between two consecutive vowel onset points (VtoV) were calculated. The results show that %V has significantly different values in mildly affected patients as compared to controls. For Italian, %V spans between 44% and 50% for healthy subjects and between 51% and 58% for PD subjects. A positive correlation was found between %V and the number of years of PD since its insurgence. The correlation with the age at which the disease insurges is weak. With regard to VtoV, PD subjects do not speak at a significantly slower rate than healthy controls, though a trend in this direction was found. The data suggest that %V could be used as a more reliable parameter for the early diagnosis of PD than speech rate.

Index Terms: Parkinson's Disease, speech rhythm, early diagnosis

1. Introduction

Parkinsonian speech has been extensively described in medical literature. There is indeed a general consensus about the typical patterns occurring in the course of Parkinson's disease (PD). Patients affected by hypokinetic dysarthria present disorders of laryngeal, respiratory and articulatory functions that result in breathy or hoarse voice, reduced loudness, narrow pitch variability, imprecise articulation, abnormal speech rate, hesitant and disfluent speech [1-,4].

Phonetic studies on Parkinsonian speech have underlined its specific segmental and suprasegmental peculiarities. At the segmental level, researchers have focussed on the amplitude and duration of speech gestures. It has been shown that impairments in vowel articulation are present even in mildly affected patients and deteriorate significantly in the course of the illness [5], [6], [7], [8], [9]. Micro adjustments in the movements of the articulators determine changes in the frequency of the first two vowel formants in the production of vowels, with a modification of their acoustic characteristics. Skodda, Gronheit, and Schlegel [9] found a considerable reduction in the vowel space of PD patients and concluded that VAI (Voice Articulation Index), computed from the first and the second formant frequencies of the vowels /a/, /i/ and /u/, can be predictive of PD progression.

At the suprasegmental level, PD patients' reduced range of motions due to muscle rigidity seems to have an effect on pitch variability, speech fluency and speech rate. However, studies diverge as to which prosodic characteristics of speech can be related unequivocally to PD. For example, a comparison of German native speakers with PD and healthy controls showed the existence of a correlation between PD symptoms, the number of pauses in speech and lower pitch variability [10].

As regards to articulation rate, as underlined by Skodda and Schlegel [11], the data available in the literature do not seem to highlight a uniform pattern of alteration in Parkinsonian speech. If, on the one hand, Skodda and Schlegel [11] found no significant differences in overall articulatory rate between PD patients and controls, on the other hand, Blanchet and Snyder [12] claim that PD patients may speak at a rate that can be considered either too fast or too slow compared to that of unimpaired speakers.

The description of the speech characteristics of PD patients has been accompanied by a search for the parameters that can serve as an index for the progression of the disease. Liss et al. [13] found that American English PD patients and controls could be classified with an 80% accuracy on the basis of a number of rhythmic metrics, such as the standard deviation of vocalic intervals over a sentence (ΔV), the standard deviation of consonantal intervals over a sentence (ΔC), the proportion of the vocalic intervals (%V), the rate-normalized standard deviation of vocalic and consonantal intervals (VarcoV, VarcoC), the Pairwise variability indices (nPVI-v, rPVI-c) and the articulation rate. %V, that is the percentage of vocalic portion in articulated speech, appeared to be one of the most important variables to maximize the distance between healthy controls and speakers affected by hypokinetic, hyperkinetic, flaccid-spastic or ataxic dysarthria. While %V was about 41% in healthy speakers of American English, in dysarthric speakers %V ranged from 45.87% to 51.82%, depending on the type of impairment. In a later study Liss, LeGendre and Lotto [14] found that the analysis of speech envelope modulation spectra (EMS) also allows to distinguish dysarthric patients from healthy controls. Six of the 48 EMS metrics analyzed were found to be robust indicators of speech signals coming from dysarthric vs. healthy speakers (with 95.3% accuracy on cross-validation). Orozco-Arroyave et al. [15] proved that the speech of German, Spanish and Czech PD patients can be classified automatically based on the systematic separation of voiced and unvoiced segments.

Studies on Italian Parkinsonian speech have focused both on the segmental and the suprasegmental aspects. As for the former, preliminary acoustic data on Italian PD patients show

http://dx.doi.org/10.21437/Interspeech.2016-74

a smaller difference in intensity between bilabial voiced plosives and the following vowel than healthy speakers [16]. It was suggested that PD patients cannot produce plosives with a complete closure of the articulators, possibly due to a reduction in the amplitude of speech gestures. Additionally, the results of an acoustic and kinematic study of Italian bilabial stops followed by the vowels [i] and [a] showed that the two mild-to-severe PD subjects of the study produce slightly velarized vowels and shorter consonants than the controls [7].

With regard to the suprasegmental level, in a study involving a sentence-repetition task, Bandini et al. [17] found no significant difference in noise level or F0 variability between Italian PD patients and controls. By contrast, the former have longer pauses between each sentence repetition as well as an overall lower percentage of articulated time during a whole repetition period than the controls.

The phonetic literature on Italian, however, does not seem to have paid particular attention to the rhythmic correlates of Parkinsonian speech. In particular, to the authors' knowledge, there are no studies that compare the productions of PD and healthy subjects in terms of a metric based on %V and articulation rate.

2. The study

2.1. Objective

The objective of this study is to verify whether there are rhythmic variations in the speech produced by Italian Parkinsonian and healthy subjects. Since PD is an age related neurodegenerative disease, in this study PD speech and healthy speech are compared according to a rhythmic metric that was proved to be connected with aging in Italian [18]. Such a metric is based on the computation of two acoustic parameters: the vowel percentage in the utterance (%V) and the mean interval between two consecutive vowel onset points (VtoV) [19]. This latter represents a perceptual counterpart of articulation rate. Indeed, rather than expressing the rate at which speech is produced (syllables/s), VtoV represents the rate at which speech is perceived. In this regard, the syllabic boundaries mark the most perceptually salient instants along the speech signal, that is the vowel onset points (VOPs). So the higher the VtoV value, the slower the articulation rate.

2.2. Subjects and Methods

A corpus of read speech produced by 11 PD subjects was collected. The subjects were males and females, aged between 49 and 75, diagnosed with the same degree of severity of the disease (mild), but differing with regard to the year of insurgence of the disease, ranging from 3 to 20 years.
A group of 15 neurologically healthy speakers of the same age range and regional area as the PD speakers were recorded reading the same text for control.

The subjects were instructed to read a text, a short description of the city of Milan (about 500 syllables), from a visual prompt on a computer screen. They were encouraged to produce speech in their normal, conversational voice. The speech samples were recorded in a silent room, with the software Praat, at a 44100 Hz sampling rate [20].

The speech signal was manually segmented and labeled in 3 different tiers: consonantal (C) and vocalic (V) segments, C and V portions, and VtoV intervals. The CV boundaries were identified through the visual inspection of speech spectrograms and waveforms. [w] and [j] were treated as consonants and the boundary was placed between the approximant and the vowel. Diphthongs were segmented as one or two vowels depending on their spectro-acoustic characteristics: if both vowels presented a specific steady-state formant pattern, the diphthong was divided into two VtoVs; otherwise, it was treated as a single interval. The voice onset time in stop consonants was included within the consonantal interval. In vowel+nasal consonant sequences, the nasalized portion of the vowel was assigned to the vowel. In the case of initial voiced stop consonants, the first boundary was considered to be the onset of the glottal pulses. Post-pausal voiceless plosives were assigned a duration equal to the mean value of single plosives in the same utterance. Respiratory and syntactic pauses were labeled as silent pauses (SP). Different kinds of disfluencies (laryngealizations, nasalizations, vocalizations, etc.) were treated as filled pauses (FP). In the VtoV tier, the pre-pausal vowels were not considered. The duration of all segments was extracted using a Praat script in order to calculate the %V and mean VtoV. Silent and filled pauses were not considered for the calculation of these two parameters. The significance of the data was verified with t-tests.

3. Results

The results of the analysis are presented in figure 1. %V is plotted on the x-axis, while mean VtoV is plotted on the y-axis. White circles represent the controls, grey circles the PD speakers.

The Figure shows a clear difference between the two groups in the distribution of the values along the x-axis. The difference is centered around 50%, with the PD patients having vowel percentages above it and the control group below it. The controls' %V ranges from 44 to 50 (mean values 46.6%, S.D. ± 1.88). The PD subjects' %V ranges from 51 to 58 (mean values 54.4; standard deviation. ± 2.52).

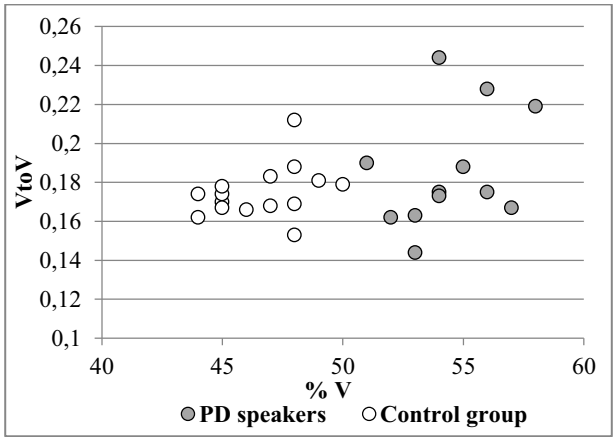

Figure 1: *Values of %V and VtoV of PD and controls.*

The statistical analysis shows that the difference between the two groups is significant as shown in Fig.2 (p= 9.28E-10)..

In order to verify whether there is a relationship between the %V and the speakers' age, we calculated a correlation index. Fig. 3 shows that the speakers' chronological age does not have a significant effect on the speakers' %V (Fig. 3). For both groups, there is only a weak positive correlation between

the two variables (control group: r = 0.10; experimental group r = 0.11).

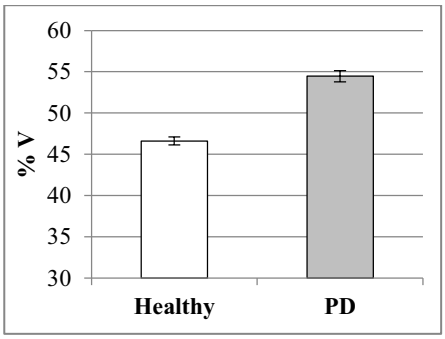

Figure 2: *Mean values of %V and standard error for controls and PD*

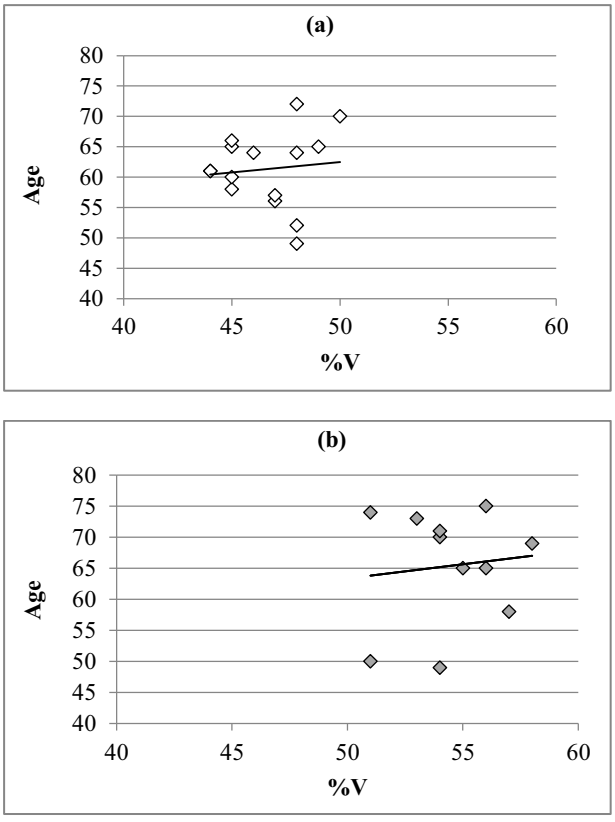

Figure 3: *%V and chronological age for controls* (**a**) *and PD* (**b**)

For the PD subjects, the increase in %V is, instead, positively correlated with the number of years since the insurgence of the disease (r=0.54). Looking at the correlation line in figure 4 it can be hypothesized that the duration of the disease in number of years has a greater effect than the age at which the disease insurges.

With regard to the variations in VtoV (Fig. 5), the present data do not show a significant difference between the PD and the controls (p = 0.09) in terms of articulation rate. This is possibly due to the fact that the PD subjects of the present experiment have all a mild form of the disease. However, the data in fig. 1 show a trend of greater variation in VtoV values for the PD subjects than for the controls. For the former, VtoV ranges from a minimum of 0.14 sec to a maximum of 0.24 sec (mean value = 0.18, standard deviation = 0.02), for the latter, VtoV ranges from 0.15 sec to 0.21 sec (mean value = 0.17, standard deviation = 0.13).

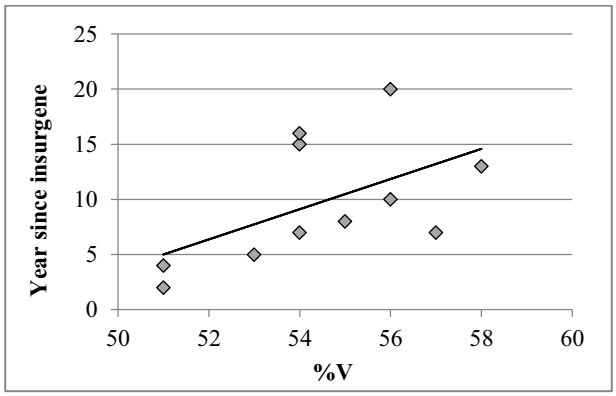

Figure 4: *Vowel percentage per year since PD insurgence.*

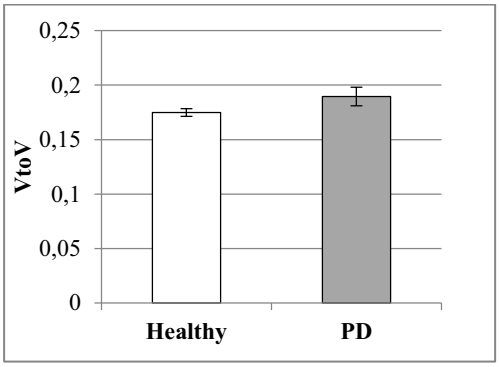

Figure 5: Mean values of *VtoV and standard error for controls and PD.*

4. Discussion and conclusions

The results of the investigation suggest that the speech of PD subjects is characterized by an abnormally high %V with respect to the speech of controls. For Italian, %V in an utterance spans between 44% and 50% for healthy subjects and between 51% and 58% for PD subjects. These data show that Liss et al.'s results [13], based on American English, also apply to Italian. Future work will look at PD patients with mother tongues different from Italian to examine whether abnormal variations in %V can be considered a general trait of PD.

With regard to VtoV, PD subjects do not appear to speak at a significantly slower rate than the healthy subjects, though a trend in this direction was found. Additionally, the greater variability within the experimental group confirms what reported in Blanchet and Snyder [12]. They argue that PD patients may speak at a rate that can be considered either to fast or to slow compared to that of unimpaired speakers. Differences in VtoV may be more relevant when comparing healthy subjects with PD patients with a greater degree of severity of the disease than the present study, which examines the speech of subjects with a mild form of PD. Investigations

of PD subjects with more severe forms of PD may in fact throw light on the acoustic characteristics of what are perceived as differences in speech rate between PD and healthy subjects.

The abnormal increase of %V in PD patients seems to be related to the duration of the disease (i.e., in number of years since its insurgence) and not to the patients' chronological age. Considering that an increase in %V and VtoV distinguishes the speech of older from that of younger speakers [18], it is possible to hypothesize that PD causes an acceleration of aging. Thus Parkinsonian speech is much more affected by the progression of the disease than by chronological age. This also calls for future investigations with different degrees of severity of the disease, that is, intermediate and advanced.

We speculate that the reported difference in %V has a cause in the motor impairments characterizing PD, like the difficulty at initiating movements (acinesis), the slowing down of the velocity in execution of movements after they are initiated (bradicinesis), and muscular rigidity [21-23]. The motor impairments of PD patients may have different effects on the production of vowels and consonants, as these sounds require different types of movements of the articulatory organs for their realization. While in fact vowels are produced with a more static configuration, i.e., with no obstruction of the vocal tract, consonants require the articulation of fast movements of the phonatory organs. Thus, because speech production requires a precise temporal coordination of articulatory movements, the motor impairments due to Parkinson would make it difficult for the phonatory organs to rapidly pass from a static phase (for vowels) to a dynamic phase (for consonants). Vowel gestures, instead, can be sustained once they have been initiated. The prolonging of the static phase over the dynamic phase accounts for the greater percentage of vocalic portion in the PD subjects (51-58%) with respect to that of the controls (44%-50%).

Finally, %V shows significantly different values in mildly affected patients as compared to controls. Thus, it could be used for the early diagnosis of PD. %V would so present an advantage as a PD detection parameter over speech rate that, at least at a mild level, does not change drastically.

5. Acknowledgements

This work is supported partly by the National Social Science Fund of China 13&ZD189. The data were collected through a collaboration with a regional non-profit Parkinson's Disease Association (Associazione Parkinsoniani Sud Pontino – Formia – Southern Italy). The authors thank the neurologist and the patients who provided the data and the members of the association for their collaboration.

6. References

[1] F. L. Darley, A. E. Aronson, and J. R. Brown, "Clusters of deviant speech dimension in the dysarthrias," *Journal of Speech and Hearing Research*, vol. 12, pp. 462-469, 1969a.

[2] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no 2, pp. 246-269, 1969b.

[3] F. L. Darley, A. E. Aronson, and J. R. Brown, *Motor Speech Disorders*. Philadelphia: W. B. Saunders, Co, 1975.

[4] L. O. Ramig, C. Fox, and S. Sapir, "Speech treatment for Parkinson's Disease," *Expert Rev. Neurotherapeutics,* vol. 8, no 2, pp. 299-311, 2008.

[5] K. G. Forrest, G. Weismer, and S. Turner, "Kinematic, acoustic, and perceptual analyses of connected speech produced by Parkinsonian and normal geriatric adults," *Journal of the Acoustical Society of America,* vol. 85, pp. 2608-2622, 1989.

[6] S. Skodda, W. Visser, and U. Schlegel, "Vowel articulation in Parkinson's Disease," *Journal of Voice*, vol. 25, no 4, pp. 467-72, 2011.

[7] B. Gili Fivela, 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 ISSP, May 5-8, Cologne, Germany, Proceedings,* 2014, pp. 146-149.

[8] S. Sapir, L. Ramig, J. Spielman, and C. Fox, "Formant Centralization Ratios (FCR): A proposal for a new acoustic measure of dysarthric speech," *Journal of Speech and Hearing Research*, vol. 53, pp. 114-125, 2010.

[9] S. Skodda, W. Gronheit, and U. Schlegel, "Impairment of vowel articulation as a possible marker of disease progression in Parkinson's Disease," *PLoS ONE*, vol. 7, no 2, pp. 1-8, 2012.

[10] S. Skodda, W. Gronheit, and U. Schlegel, "Intonation and speech rate in Parkinson's Disease: General and dynamic aspects and responsiveness to levodopa admission," *Journal of Voice*, vol. *25*, no 4, pp. 199-205, 2011.

[11] S. Skodda, and U. Schlegel, "Speech Rate and Rhythm in Parkinson's Disease", *Movement Disorders*, vol. 23, no 7, pp. 985-992, 2008.

[12] P. G. Blanchet, and G. J. Snyder, "Speech rate deficit in individuals with Parkinson's disease: A review of the literature," *Journal of Medical Speech-Language Pathology*, vol. 17, pp. 1-7, 2009.

[13] 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.

[14] J. M. Liss, S. LeGendre, and A. J. Lotto, "Discriminating dysarthria type from envelope modulation spectra," *Journal of Speech, Language, and Hearing Research*, vol. 53, pp. 1246-1255, 2010.

[15] J. R. Orozco-Arroyave, F. Hönig, J. Arias-Londoño, J. F. V. Bonilla, S. Skodda, J. Rusz, and E. Nöth, "Automatic detection of Parkinson's Disease from words uttered in three different languages," in *INTERSPEECH 2014* - 15th Annual Conference of the International Speech Communication Association, September 14-18, Singapore, *Proceedings*, 2014, pp. 1573-1577.

[16] B. Gili Fivela, V. Sallustio, S. Pede, M. Grimaldi, and D. Patrocinio, "Intelligibility assessment of dysarthric Italian speech: Correlations between acoustic measures and auditory perceptual ratings," *29th World Congress of the International Association of Logopedics an Phoniatrics*, *August 25-29, Turin, Italy,* 2013.

[17] A. Bandini, F. Giovannelli, S. Orlandi, S. D. Barbagallo, M. Cincotta, P. Vanni, R. Chiaramonti, A. Borgheresi G. Zaccara, and C. Manfredi, "Automatic identification of dysprosody in idiopathic Parkinson's disease," *Biomedical Signal Processing and Control*, vol. 17, pp. 47-54, 2015.

[18] M. Pettorino, and E. Pellegrino, "Age and rhytmic variations. A study on Italian," *INTERSPEECH 2014 - 15th Annual Conference of the International Speech Communication Association, September 14-18, Singapore, Proceedings,* 2014, pp. 1234-1237.

[19] M. Pettorino, M. Maffia, E. Pellegrino, M. Vitale, and A. De Meo,"VtoV: A perceptual cue for rhythm identification", in P. Mertens and I.C. Simon (eds.) *Proceedings of the Prosody-Discourse Interface Conference*, 2013, pp. 101-106.

[20] P. Boersma, "Praat, a system for doing phonetics by computer," *Glot International*, vol. 5, pp. 341-345, 2001.

[21] A. Berardelli, and A. Currà, "*Fisiopatologia della bradicinesia nella malattia di Parkinson*," LIMPE, 2000.

[22] I. Donaldson, C. D. Marsden, S. Schneider, and K. Bhatia, "*Marsden's Book of Movement Disorders*". Oxford: Oxford University Press, 2012.

[23] H. Sagar, "*The new Parkinson's Disease Handbook: The essential guide for sufferers and Carers*", London: Random House, 2002.