



Prosodic and Segmental Correlates of Spontaneous Dutch Speech in Patients with Parkinson's Disease: A Pilot Study

Vass Verkhodanova¹, Matt Coler¹

¹University of Groningen, Campus Fryslân, the Netherlands

v.verkhodanova@rug.nl, m.coler@rug.nl

Abstract

This study investigates the acoustic correlates of prosody and vowel articulation in Dutch individuals with Parkinson's Disease (PD). We compared prosodic and segmental acoustic measures in spontaneous monologues in PD patients to those in elderly healthy controls matched for age and gender. For the prosodic measurements of pitch variability, span and speech rate, we analysed fundamental frequency and intensity. For articulation measurements, the first two formants were calculated from Dutch corner vowels extracted from the speech signal. Results show a monopitch trend, reduced speech rate, centralization of the formant frequencies and reduced first formant variability in individuals with PD compared to control group.

Index Terms: Parkinson's Disease, hypokinetic dysarthria, dysprosody, vowel articulation, acoustic analysis, Dutch spontaneous speech

1. Introduction

Parkinson's Disease (PD) is a neurodegenerative disorder characterized by progressive loss of dopaminergic neurons [1, 2], affecting 1-2% of people older than 60 [3]. The progressive dopaminergic loss results in a range of motor and non-motor deficits. In addition to symptoms like muscular rigidity and tremor, up to 90% of PD patients develop a distinctive speech disorder referred as hypokinetic dysarthria [4]. The typical pattern of dysarthric manifestation involves monotone voice, hypophonia, reduced articulatory movements, "slurred" speech and bursts and rushes of speech [5, 6]. Prosodic and articulatory deficits due to hypokinetic dysarthria are commonly observed in various languages [3, 7, 8].

Monopitch is the most common and deviant prosodic correlate in hypokinetic dysarthria [5, 9, 10]. Monopitch is manifested as lack of normal pitch variability. A typical consequence of this deficit is the reduced ability to achieve certain intonation contours. Thus, individuals with PD may experience difficulties in expressing certain meanings. They may also be perceived by others as withdrawn and cold [11]. A common way to track prosodic deficits in dysarthric speech is through the analysis of disturbances in fundamental frequency [9, 12], intensity [13], stress [14], and speech rate and rhythm [15]. While speech rate and intensity has been shown to yield inconsistent results [12, 16], monopitch was reported to be consistent characteristic of hypokinetic dysarthria.

Among common articulation deficits in speech of individuals with PD is vowel "undershooting" [17]; that is, the reduced ability to achieve a certain vowel target due to slower and reduced movements of the articulatory organs. This "undershooting" leads to the centralization of the vowels, contributing to reduced speech intelligibility [18]. A common way to capture this phenomenon is with vowel space area (VSA) measurement calculated from the first two formants of the corner vowels, and

with ratios based on formant measurements. While VSA proved to be unreliable to separate pathological from non-pathological speech [19, 20], vowel articulation index (VAI) or F2 ratio of the vowels /i/ and /u/ have been demonstrated to be more sensitive to speech impairment and less to interspeaker variability [21, 22]. Speakers' relative stability of reaching a vowel target has been shown to account for speech intelligibility and to contribute to differentiating pathological from non-pathological speech [18, 22]

To the best of our knowledge, acoustic studies on dysarthric Dutch speech are scarce, as are acoustic studies on articulatory performance of PD during spontaneous speech. Thus, we aimed to explore acoustic correlates of prosodic and articulatory deficits in Dutch spontaneous speech. With this study we addressed the question of whether Dutch spontaneous speech reflects the common monopitch and vowel centralization trends of PD dysarthric speech.

2. Methods

Recordings of spontaneous speech used in the present study originate from [23]. The collection and analysis of the material was approved by the Medical Ethics Committee of the University Medical Center Groningen. All participants gave written informed consent.

2.1. Participants

A total of 30 Dutch native speakers participated in this study. The participants were split into two groups. The first group included 15 individuals clinically diagnosed with idiopathic PD: six males and nine females, mean age 65 (SD: ± 8) years. The second group was comprised of 15 healthy controls (hereafter HC): mean age 65 (SD: ± 8) years, matched for age and gender. Table 1 summarizes the demographic data of both groups.

Table 1: Summary of group demographics. Age and disease duration are given in years

		PD	HC
Male: Female		6:9	6:9
Age	M	65.1	65
	SD	7.8	8
Disease duration	M	7.3	-
	SD	3.6	-
Hoehn & Yahr scores	M	2	-
	SD	0.7	-

2.2. Speech task and recording procedure

Patients were recruited all over the Netherlands. Recordings were made at their homes. For the purpose of Harris et al. study [23] participants were asked to perform two speech tasks (monologue and recitation) and two music tasks (singing familiar melodies and improvised singing). For the current study only monologues were used. The duration of the monologues recordings ranged from 1.5 minutes to 10 minutes, the whole corpus is around 1 hour and 40 minutes. For more detailed information on data collection, participants profiles and speech tasks see [23].

2.3. Annotation

In this study prosodic analysis was performed automatically and did not require manual annotation. As for vowel articulation, each monologue was segmented and annotated for the occurrence of the three corner vowels /a, i, u/ and their respectively short or lax counterparts /a, ɪ, u/. The annotation was made manually based on visual observation of the waveform and the wideband spectrogram in Praat [24]. All annotation work was done by the same trained phonetician to keep segmentations and annotation consistent. Owing to privacy restrictions it was not possible to check agreement with another annotator. We used similar criteria for annotation as that in [22]. Suitable vowels were selected according the following criteria:

1. Only vowels occurring in intelligible, phonated words were annotated.
2. Only vowels with a stable part of at least 40 ms were selected. This stable part was the central part of each vowel, starting at least one period after vowel onset and ending one period before vowel offset.
3. Vowels preceded by a voiced sound were only selected if that sound matched the respective vowel's place of articulation, to ensure that formant transitions and co-articulation did not affect the vowel.
4. Vowels immediately following nasals, glides or other vowels were not selected.
5. Certain exceptions were made for the long vowels: in some cases they were annotated after the consonants not matching vowels' place of articulation. In such cases, the stable parts of these vowels were selected starting at least four periods after vowel onset.

2.4. Acoustic analysis

Acoustic measures were obtained with the Speech Signal Toolkit (SPTK) for Python [25] and with speech analysis software Praat [24]. SPTK toolkit was used to track fundamental frequency (F0) based on the robust algorithm for pitch tracking (RAPT) [26]. Praat scripts were used to estimate the speech rate [27] and to obtain frequencies of the first two formants for vowel articulation measures.

2.4.1. Prosodic analysis

In this study we investigated two prosodic characteristics: speech and articulation rates and pitch. Typically measuring speech and articulation rates requires annotation of phonemes or syllables, which is time-consuming and sometimes error-prone. Therefore, these measurements were done automatically by detecting syllable nuclei Praat script written by de Jong et al. [27]. In this algorithm, syllable nuclei correspond to peaks

in intensity preceded and followed by dips in intensity, with unvoiced peaks being discarded. This script has been shown to be informative for the study of French dysarthric speech [28]. In our study we have used -20 dB silence threshold, 4 dB dip and 70 ms as a minimal pause duration. Speech rate was computed as the number of syllables divided by total time. Articulation rate was computed as number of syllables divided by phonation time.

Pitch tracking was performed with David Talkin's RAPT algorithm [26] implemented in the SPTK toolkit [25]. The RAPT algorithm identifies pitch candidates with the cross-correlation function and then attempts to select the "best fit" at each frame by dynamic programming [26, 29]. From the pitch trajectory we calculated pitch variance estimation as the average of the squared deviations from the mean of F0 (1) and pitch span (the estimation of speaker's range of frequencies) as difference between minimum and maximum of F0 values.

$$f0 \text{ variance} = \text{mean}|f0 - \text{mean}(f0)|^2 \quad (1)$$

2.4.2. Vowel articulation analysis

To determine vowel articulation differences, we calculated four measurements based on [22]: (1) F1 and F2 variability for each speaker, (2) the vowel space area (VSA), (3) the vowel articulation index (VAI), and (4) the F2 ratio of the vowels /i, ɪ/ and /u, u:/.

According to Kim et al. [18], the F1 and F2 contrasts reflect a speaker's relative stability in achieving vowel targets. These measurements were computed based on the description in [18, 22], but with introduced normalization to allow relative comparison of different vowels. For each speaker the mean normalized standard deviation of each vowel was calculated.

The following formula was used for VSA calculation [30]:

$$VSA = 0.5 \times |F1i \times (F2a - F2u) + F1a \times (F2u - F2i) + F1u \times (F2i - F2a)| \quad (2)$$

The VAI was based on the calculation of Roy et al. [31]:

$$VAI = \frac{F1a + F2i}{F1i + F1u + F2a + F2u} \quad (3)$$

For VSA, VAI and the F2 ratio measurements the formant frequencies were averaged over vowel and speaker.

2.5. Results and discussion

Table 2 summarizes the results of the prosodic measurements for each group. The pitch variance and pitch span were calculated for every 10 seconds within the recording. As expected, the PD group showed lower values of F0 variance (Fig. 1). Speech and articulation rates were calculated for the whole duration of each recording, and as expected prosodic measurements for F0 variation and span were lower for the PD group, except for the speech and articulation rates.

Table 3 summarizes the results of the vowel measurements for each group. We found the predicted pattern of vowel articulation precision: the values of VSA (see Fig. 2), VAI and F2 ratio were lower for the PD group in comparison with the HC group.

To determine differences across data we used Kruskal-Wallis rank sum tests for non-parametric data. The overall comparison of PD and HC subjects have shown significant differences for the measurements of F0 variance ($\chi^2 = 5.8$, $p <$

Table 2: Summary of prosodic measurements for each group, where F0 variance is estimation of pitch variability, F0 span is the estimation of speaker's range of frequencies

Group	Prosodic measurements			
	F0 variance	F0 span	Speech rate	Articulation rate
PD M	0.038	1.06	2.57	4.09
SD	0.027	0.19	0.29	0.54
HC M	0.04	1.08	2.81	4.23
SD	0.023	0.16	0.29	0.33

0.02), VAI ($\chi^2 = 5.1$, $p < 0.03$), F2 ratio ($\chi^2 = 4.2$, $p < 0.05$) and F1 variability ($\chi^2 = 7.3$, $p < 0.007$). Speech rate distribution showed to be significantly different for PD as well ($\chi^2 = 4.2$, $p < 0.04$). This finding, along with the lower values of speech and articulation rate for the PD group, is not in line with the previous studies [15, 32]. However, this inconsistency may be accounted for with the methodological differences and small sample size relative to [15], as well as possible differences in pause distribution that were not accounted for in this study. It was also shown that speech rate is heterogeneous within the population of PD speakers [15].

To assess if F0 variance is related to gender differences, we ran separate analysis for male and female participants. The most affected group was male individuals with PD. However, a comparison between group and gender pairs showed significant difference, except for healthy controls: the F0 variability did not differ significantly between male and female HC participants. This finding contradicts the previous study on gender-related patterns of dysprosody by Skodda et al. [3]. This inconsistency may be attributed to the smaller sample size or effect of the gender differences induced by the Hertz-based measures. Nonetheless, additional investigation is required since this might suggest the possibility of different gender-related dysprosody patterns.

F0 variance, speech rate, VAI, F2 ratio and F1 variability proved to be sensitive to differentiate pathological from non-pathological speech on a group level. The first two measurements, F0 variance and speech rate, account for clear dysprosody patterns, suggesting that monopitch and abnormal speech rate are common feature for Dutch dysarthric speech as well. VAI and F2 ratio are related to the vowel space, confirming the hypothesis of vowel centralization. The significant difference of F1 variability reflects a speaker's steadiness in achieving vowel targets [18].

Overall, these results are in line with previous studies [7, 22]. Dutch spontaneous speech reflected the expected reduced trend in F0 variability for the PD group, confirming the monopitch tendency common for the hypokinetic dysarthria. An acoustic analysis of Dutch vowel articulation in spontaneous speech was sensitive enough to differentiate pathological and non-pathological speech, as it was previously shown for German spontaneous speech [22].

The lack of consistency with previous studies was expected in certain measures and could be attributed to in-group variation due to scarcity and imbalance of data, as well as the difference in methodology. Thus, future research should include larger sample size, more balanced groups and corpus, and further acoustic and perceptual measurements to better understand Dutch spontaneous dysarthric speech.

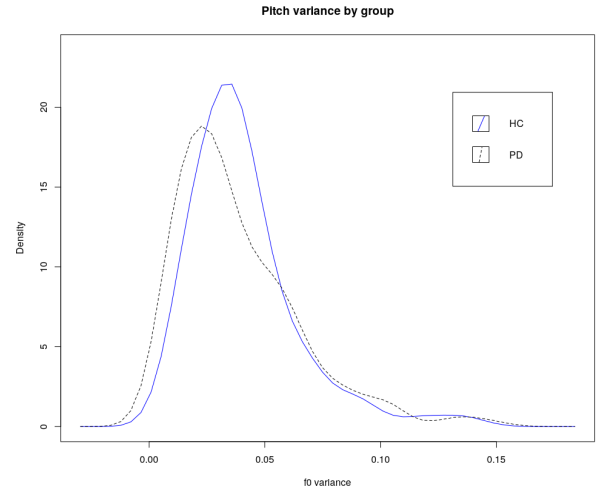


Figure 1: F0 variance for PD and HC groups

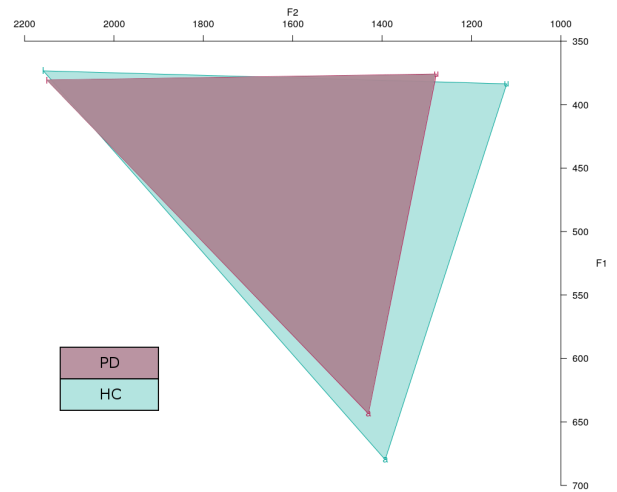


Figure 2: VSA for PD and HC groups

3. Conclusions

With this pilot study we demonstrated the adequacy of acoustic measurements of prosody and vowel articulation to differentiate Dutch dysarthric from non-pathological speech. The common monopitch trend was confirmed. Additionally the acoustic correlates of imprecise vowel articulation were shown to be significantly different for PD and HC groups. This study contributes to the growing body of research on both acoustic correlates of vowel articulation in spontaneous dysarthric speech, as well as on acoustic analysis of speech of Dutch individuals with PD.

4. Acknowledgements

We are very grateful to Dr. Robert Harris for giving access the speech material used for this study.

5. References

- [1] L. Kalia and A. Lang, "Parkinson's Disease." *The Lancet*, vol. 386, no. 9996, pp. 896–912, 2015.

Table 3: Summary of vowel measurements for each group, where F2-ratio is ratio of /i/ and /u/ second formants, F1-var and F2-var are normalized F1 and F2 variabilities (mean(sd/mean))

Group	Vowel measurements				
	VSA	VAI	F2-ratio	F1-var	F2-var
PD M	115500	0.79	1.6	0.12	0.13
SD	59552	0.06	0.29	0.0005	0.0002
HC M	155100	0.87	1.9	0.12	0.13
SD	66700	0.08	0.36	0.0004	0.0001

- [2] J. M. Fearnley and A. J. Lees, "Ageing and Parkinson's Disease: substantia nigra regional selectivity," *Brain*, vol. 114, no. 5, pp. 2283–2301, 1991.
- [3] S. Skodda, W. Visser, and U. Schlegel, "Gender-Related Patterns of Dysprosody in Parkinson Disease and Correlation Between Speech Variables and Motor Symptoms," *Journal of Voice*, vol. 25, no. 1, pp. 76–82, 2011.
- [4] A. K. Ho, R. Iansek, C. Marigliani, J. L. Bradshaw, and S. Gates, "Speech impairment in a large sample of patients with Parkinsons Disease," *Behavioural neurology*, vol. 11, no. 3, pp. 131–137, 1999.
- [5] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech, Language, and Hearing Research*, vol. 12, no. 2, pp. 246–269, 1969.
- [6] J. Duffy, *Motor Speech Disorders. Substrates, Differential Diagnosis, and Management*, 3rd ed. Elsevier Mosby, 2013.
- [7] J. Ruzs, R. Cmejla, H. Ruzickova, and E. Ruzicka, "Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's Disease," *The journal of the Acoustical Society of America*, vol. 129, no. 1, pp. 350–367, 2011.
- [8] J. K.-Y. Ma, T. L. Whitehill, and S. Y.-S. So, "Intonation contrast in Cantonese speakers with hypokinetic dysarthria associated with Parkinsons Disease," *Journal of Speech, Language, and Hearing Research*, vol. 53, no. 4, pp. 836–849, 2010.
- [9] R. J. Holmes, J. M. Oates, D. J. Phyland, and A. J. Hughes, "Voice characteristics in the progression of Parkinson's Disease," *International Journal of Language & Communication Disorders*, vol. 35, no. 3, pp. 407–418, 2000.
- [10] S. Anand and C. E. Stepp, "Listener perception of monophth, naturalness, and intelligibility for speakers with Parkinson's Disease," *Journal of Speech, Language, and Hearing Research*, vol. 58, no. 4, pp. 1134–1144, 2015.
- [11] A. Jaywant and M. D. Pell, "Listener impressions of speakers with Parkinsons Disease," *Journal of the International Neuropsychological Society*, vol. 16, no. 1, pp. 49–57, 2010.
- [12] S. Skodda, H. Rinsche, and U. Schlegel, "Progression of dysprosody in Parkinson's Disease over time: a longitudinal study," *Movement Disorders*, vol. 24, no. 5, pp. 716–722, 2009.
- [13] L. O. Ramig, S. Sapir, C. Fox, and S. Countryman, "Changes in vocal loudness following intensive voice treatment (lsvt®) in individuals with Parkinson's Disease: A comparison with untreated patients and normal age-matched controls," *Movement Disorders*, vol. 16, no. 1, pp. 79–83, 2001.
- [14] H. S. Cheang and M. D. Pell, "An acoustic investigation of parkinsonian speech in linguistic and emotional contexts," *Journal of Neurolinguistics*, vol. 20, no. 3, pp. 221–241, 2007.
- [15] S. Skodda and U. Schlegel, "Speech rate and rhythm in Parkinson's Disease," *Movement Disorders*, vol. 23, no. 7, pp. 985–992, 2008.
- [16] K. M. Rosen, R. D. Kent, and J. R. Duffy, "Task-based profile of vocal intensity decline in Parkinsons Disease," *Folia Phoniatrica et Logopaedica*, vol. 57, no. 1, pp. 28–37, 2005.
- [17] K. Forrest, G. Weismer, and G. S. Turner, "Kinematic, acoustic, and perceptual analyses of connected speech produced by parkinsonian and normal geriatric adults," *The Journal of the Acoustical Society of America*, vol. 85, no. 6, pp. 2608–2622, 1989.
- [18] H. Kim, M. Hasegawa-Johnson, and A. Perlman, "Vowel contrast and speech intelligibility in dysarthria," *Folia Phoniatrica et Logopaedica*, vol. 63, no. 4, pp. 187–194, 2011.
- [19] G. Weismer, J.-Y. Jeng, J. S. Laures, R. D. Kent, and J. F. Kent, "Acoustic and intelligibility characteristics of sentence production in neurogenic speech disorders," *Folia Phoniatrica et Logopaedica*, vol. 53, no. 1, pp. 1–18, 2000.
- [20] S. Sapir, J. L. Spielman, L. O. Ramig, S. BH, and C. Fox, "Effects of intensive voice treatment (the lee silverman voice treatment [lsvt]) on vowel articulation in dysarthric individuals with idiopathic Parkinson's Disease: acoustic and perceptual findings," *Journal of Speech, Language, and Hearing Research*, vol. 50, no. 4, pp. 899–912, 2007.
- [21] 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.
- [22] M. Strinzel, V. Verkhodanova, F. Jalvingh, R. Jonkers, and M. Coler, "Acoustic and perceptual correlates of vowel articulation in Parkinsons Disease with and without mild cognitive impairment: A pilot study," in *International Conference on Speech and Computer*. Springer, 2017, pp. 56–64.
- [23] R. Harris, K. L. Leenders, and B. M. de Jong, "Speech dysprosody but no music dysprosody in Parkinsons Disease," *Brain and language*, vol. 163, pp. 1–9, 2016.
- [24] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program], version 6.0.14," 2017.
- [25] S. team, "SPTK: The speech signal processing toolkit, version 3.11," <http://sp-tk.sourceforge.net/>, 2017.
- [26] D. Talkin, "A robust algorithm for pitch tracking (RAPT)," *Speech coding and synthesis*, vol. 495, p. 518, 1995.
- [27] N. H. De Jong and T. Wempe, "Praat script to detect syllable nuclei and measure speech rate automatically," *Behavior research methods*, vol. 41, no. 2, pp. 385–390, 2009.
- [28] C. D. Looze, A. Ghio, S. Scherer, G. Pouchoulin, and F. Viallet, "Automatic analysis of the prosodic variations in parkinsonian read and semi-spontaneous speech," in *Speech Prosody 2012*, 2012, pp. 71–74.
- [29] D. Morrison, R. Wang, and L. C. De Silva, "Ensemble methods for spoken emotion recognition in call-centres," *Speech communication*, vol. 49, no. 2, pp. 98–112, 2007.
- [30] H.-M. Liu, F.-M. Tsao, and P. K. Kuhl, "The effect of reduced vowel working space on speech intelligibility in Mandarin-speaking young adults with cerebral palsy," *The Journal of the Acoustical Society of America*, vol. 117, no. 6, pp. 3879–3889, 2005.
- [31] N. Roy, S. L. Nissen, C. Dromey, and S. Sapir, "Articulatory changes in muscle tension dysphonia: Evidence of vowel space expansion following manual circumlaryngeal therapy," *Journal of Communication Disorders*, vol. 42, no. 2, pp. 124–135, 2009.
- [32] G. J. Canter, "Speech characteristics of patients with Parkinson's Disease: Intensity, pitch, and duration," *Journal of Speech & Hearing Disorders*, 1963.