More than words: Recognizing speech of people with Parkinson's disease

Verkhodanova, Vass

DOI:
10.33612/diss.183425053

Publication date:
2021

CHAPTER 2
HYPOKINETIC DYSARTHRIA

2.1. INTRODUCTION

A large body of research has demonstrated that changes in speech and voice of people with Parkinson's disease (PD) are inevitable as the disease progresses (Forrest et al., 1989; Ho et al., 1999a; Miller et al., 2007; Tjaden, 2008). Such changes in speech and voice constitute a speech disorder known as *dysarthria*. "Dysarthria is a collective name for a group of speech disorders resulting from disturbances in muscular control over the speech mechanism due to damage of the central or peripheral nervous system. It designates problems in oral communication due to paralysis, weakness, or in coordination of the speech musculature"(Darley et al., 1969b, p. 246).

According to previous studies, 70% to 90% of people with Parkinson's disease (PwPD) demonstrate voice changes and 45% to 55% of them demonstrate speech changes (Tjaden, 2008). Changes in speech and voice of PwPD constitute *hypokinetic dysarthria* (HD), since it is associated with disorders of dopamine insufficiency in the basal ganglia control circuit. This lack of dopamine within the basal ganglia progressively limits the muscular control of the larynx, oral cavity and other physiological mechanisms involved in speech production (Darkins et al., 1988).

The typical manifestations of HD include monotonous voice, hypophonia, reduced articulatory movements, slurred speech, as well as bursts and rushes of speech (Duffy, 2013). The existing descriptions and analyses of speech affected by HD have been mostly based on speech characteristics described as *deviant dimensions* by Darley et al. (1969a,b). In those seminal studies, the authors reported on the results of auditory perceptual assessments of dysarthric speech performed by speech and language therapists (SLTs), who evaluated 212 people with different neurological diseases (Darley et al., 1969a). Their work laid the foundation for dysarthria classification into five classes with each of which them corresponding to a different neurological disorder with distinct speech characteristics. According to Darley et al. (1969b), the top ten deviant dimensions that

---

This chapter is partially adapted from:

Strinzel M., Verkhodanova V., Jalvingh F., Jonkers R., Coler M. (2017) Acoustic and Perceptual Correlates of Vowel Articulation in Parkinson's Disease With and Without Mild Cognitive Impairment: A Pilot Study. In *Speech and Computer. Lecture Notes in Computer Science*, volume LNAI 10458. pages 56-64. Springer International Publishing.

Verkhodanova V., Coler M. (2018) Prosodic and Segmental Correlates of Spontaneous Dutch Speech in Patients with Parkinson's Disease: a Pilot Study. In *Speech Prosody Conference*, pages 163-166.

constitute HD are: (1) lack of pitch variability or *monopitch*, (2) reduced stress, (3) lack of loudness variability or monoloudness, (4) imprecise consonants, (5) inappropriate silences, (6) short rushes of speech, (7) harsh and (8) breathy voice, (9) low pitch and (10) variable rate. The presence of such deviant speech dimensions may have serious effects on the quality of life of PwPD (Parveen and Goberman, 2017), leading to communication problems, which in turn may result in tension, depression, resignation and withdrawal from conversation, and eventually – feelings of social isolation (Miller et al., 2006; Jaywant and Pell, 2010). Voice and speech changes associated with HD are reported to affect the daily life of PwPD long before the impairment of intelligibility becomes apparent (Miller et al., 2006), and are described as one of the likely and typical early symptoms of PD (Rusz et al., 2013b; Defazio et al., 2016).

In general, there are two main types of dysarthric speech analysis. First one is subjective – speech of PwPD is assessed and evaluated by SLTs (often referred to as auditory perceptual assessment). The second one is objective – speech signal in its digital form is analyzed acoustically (Brabenec et al., 2017). According to many studies, (for instance, Bunton et al. (2007); Sussman and Tjaden (2012); Duffy (2013); Näsström and Schalling (2020)), the subjective auditory perceptual assessment of dysarthria continues to be the "gold standard" for clinical decisions. The means of assessment performed by listeners ranges from rating vowels (Sapir et al., 2007b) to evaluating spontaneous conversational speech (Bunton et al., 2007; Bunton and Keintz, 2008). One of the common measures of assessment and management of speakers with dysarthria is the use of speech intelligibility scores. These scores are commonly used as a measure of the severity of the speech disorder and as a source of information for treatment and monitoring (Yorkston and Beukelman, 1978; Bunton and Keintz, 2008). Another popular approach to auditory perceptual assessment of dysarthria is the use of component-specific perceptual judgements based on evaluating lists of deviant dimensions, as first described by (Darley et al., 1969b).

Objective acoustic analysis of speech of PwPD allows to quantify the speech features important for speech and language therapy. Due to its non-invasive nature and objective and replicable results, automatic acoustic analysis has been considered by many as a potential measure of PD (Little et al., 2008; de Lima et al., 2016; Zhan et al., 2018). Acoustic analysis of dysarthric speech may also be classified into two types based on speech parametrization methods: using conventional and non-conventional features (Brabenec et al., 2017). The difference between conventional and non-conventional speech features is that conventional features are clinically interpretable, while non-conventional features are more robust towards aperiodic, noisy and irregular voices (Brabenec et al., 2017). For a detailed description of the diverse parametrization techniques and a review of the studies using different feature sets for the objective acoustic analysis of speech of PwPD see Brabenec et al. (2017).

Studies that are concerned with acoustic analysis of dysarthric speech use a wide range of speech tasks, such as sustained vowel phonation (Tsanas et al., 2010b; Orozco-Arroyave et al., 2013), syllable repetition (Lowit, 2014; Skodda et al., 2011a), reading tasks (Skodda et al., 2009, 2011b; Rusz et al., 2013b), or spontaneous speech (Rusz et al., 2013b; Vásquez-Correa et al., 2018). Even though there is less research done on spontaneous speech of PwPD, this particular speech task shows significantly different phonetic features compared to non-spontaneous speech in PwPD (Kempler and Van Lancker, 2002) and, therefore, may be more sensitive for detecting certain speech changes. For example, the study by Rusz et al. (2013b) suggests that spontaneous speech is preferable to other speech tasks in detecting imprecise vowel articulation in Czech speakers with PD.

This chapter describes two pilot studies dealing with an acoustic analysis of articulatory performance of PwPD during spontaneous speech. The literature focusing on acoustic correlates of articulation of PwPD in spontaneous speech is relatively scarce, and the common aim of these two pilot studies is to replicate the findings of Rusz et al. (2013b). These two pilot studies are concerned with similar analysis of spontaneous speech of PwPD in two related Germanic languages, German (section 2.2) and Dutch (section 2.3). This allows us to draw conclusions about cross-linguistic sensitivity of the analysis proposed by Rusz et al. (2013b). Moreover, the pilot on German speech investigates effects of mild cognitive impairment on vowel articulation in PwPD and on intelligibility of speech produced by PwPD, while the Dutch pilot looks into several prosodic features alongside vowel articulation.

2.2. VOWEL ARTICULATION IN GERMAN PwPD WITH AND WITHOUT MILD COGNITIVE IMPAIRMENT

This pilot study investigates the added effects of cognitive impairment on vowel articulation precision in PwPD. It investigates both vowel acoustics in speech of PwPD and listeners' intelligibility ratings of speech of the same PwPD. A growing body of research shows that non-motor symptoms are common and clinically significant features in PD as well. These non-motor symptoms include first and foremost Mild Cognitive Impairment (MCI) and dementia. Cognitive impairments are prevalent in approximately 30% of the individuals with PD (Aarsland and Kurz, 2010; Riedel et al., 2008) and have been found to significantly contribute to disability and reduced quality of life in PwPD. The pattern of cognitive deteriorations in PD is heterogeneous, but typically comprises memory-based impairments, executive dysfunctions, visual-spatial impairments and attentional deficits (Meireles and Massano, 2012). Although there is evidence indicating a positive correlation between motor and cognitive symptoms (Papapetropoulos et al., 2004) and affected language function in PwPD with Alzheimer's disease (Goldman and Litvan, 2011), to the best of our knowledge at the time of writing this section, no study has investigated which effects (if any) cognitive impairments have on speech motor disorders in PD.

As mentioned earlier, up to 90% of PwPD manifest hypokinetic dysarthria. Apart from respiratory, phonatory and prosodic abnormalities, a common feature of dysarthria in PD is imprecise vowel articulation. PwPD are limited in the execution of articulatory movements. Accordingly, voluntary motions of lips, jaw and tongue tend to be smaller and slower than that of neurologically healthy controls (Forrest et al., 1989). A typical consequence is articulatory "undershooting" (Forrest et al., 1989), i.e., the reduced ability to achieve a certain vowel target. As a result, the vowels produced are more centralized and become less distinct from each other. This contributes to reduced speech intelligibility (Kim et al., 2011a). A common method to represent this phenomenon is with the vowel space area (VSA) based on first (F1) and second (F2) formants values of the corner vowels.

However, findings on the VSA have been inconsistent. While the VSA-based analyses managed to distinguish dysarthric from non-pathological speech in some studies (Kent and Kim, 2003), it yielded no significant differences in other studies (Weismer et al., 2001; Sapir et al., 2007b). Ratio based vowel measurements such as the F2 ratio of the vowels /i/ and /u/ or the vowel articulation index (VAI) (Sapir et al., 2010; Roy et al., 2009) have been found to be more sensitive towards speech impairments and more robust against inter speaker variability than the VSA (Sapir et al., 2010; Skodda et al., 2012). Apart from vowel space metrics, measurements of formant frequency overlap and a speaker's relative stability of reaching a vowel target is likely to account for speech intelligibility as well (Kim et al., 2011a).

Speech motor control requires more attention in PwPD than in neurologically healthy individuals and is more likely to deteriorate as the complexity of a verbal task increases (Ho et al., 2002; Walsh and Smith, 2011). Consequently, the characteristics of dysarthria differ depending on the type of verbal task that is performed (Rosen et al., 2005; Rusz et al., 2013b). In particular, spontaneous speech shows significantly different phonetic features compared to non-spontaneous speech in PwPD (Kempler and Van Lancker, 2002). Presumably due to the attention devoted to cognitive and linguistic processing, the control of articulatory movements decreases during spontaneous speech. A study by Rusz et al. (2013b) suggests that spontaneous speech is preferable to other speech tasks in detecting imprecise vowel articulation in Czech speakers with PD. Therefore, since acoustic studies on articulatory performance of PD during spontaneous speech are relatively scarce, we aim to replicate Rusz et al.'s findings for German speakers. In addition, we are interested in whether cognitive impairments would influence vowel articulation precision in PD. One of the features of cognitive impairments in PD is a reduced attention capacity. We, therefore, hypothesize that the control of articulatory precision during spontaneous speech is more compromised in PwPD with additional cognitive impairments than in individuals with PD only. We expect this pattern to be reflected by acoustic vowel measurements. Knowing that formant frequencies differ for female and male speakers (e.g., Whiteside (2001)) and since previous research on speech of PwPD has found gender-related patterns in several speech characteristics (Skodda et al., 2011c), we expected to see gender-related effects in the acoustic vowel measurements.

The purposes of this pilot study were threefold: (1) we aimed to assess the utility of spontaneous speech as a task to detect imprecise vowel articulation often attested in hypokinetic dysarthria, (2) we evaluated the sensitivity of different vowel measurements in detecting imprecise vowel articulation and (3) we investigated whether cognitive impairments affect vowel articulation in PD.

2.2.1. METHODS

PARTICIPANTS

A total of 23 German native speakers participated in this study. They were split into three groups. The first group comprised of eight individuals who were clinically diagnosed with idiopathic PD (hereafter PD group), none of whom exhibited cognitive impairments as assessed with the Minimal Mental State Examination (MMSE). The second group included six individuals clinically diagnosed with idiopathic PD and MCI (hereafter MCI group).

Table 2.1 | Summary of speaker group demographics. Age and duration of disease are given in years

		PD	MCI	HC
Male: Female ratio		5:3	4:2	5:4
Age	(mean ± SD)	76 ± 6	81.8 ± 2.5	74.0 ± 5.9
Duration of disease	(mean ± SD)	12 ± 4.1	5.7 ± 2.4	-
MMSE	(mean ± SD)	29 ± 1	24.5 ± 1.9	29.8 ± 0.4

The third group was made up of nine elderly neurologically healthy controls (hereafter HC group) without a history of neurological disorders. Table 2.1 summarizes the demographic data of each group. All participants gave their written informed consent to the speech task and the recording procedure.

SPEECH TASK, RECORDING PROCEDURE AND ANNOTATION

Participant monologues were audio-recorded with a Zoom H2 recorder during a conversational interview with open-ended questions on a familiar topic such as hobbies, daily routines, family or prior jobs. The recordings were administered in an identical manner for each participant and were made available at the TalkBank web page (`https://sla.talkbank.org/TBB/dementia/German/Jalvingh`).

For each monologue, the occurrence of the three corner vowels /aː, i, u/ and their respectively short or lax counterparts /a, ɪ, ʊ/ was manually segmented and annotated based on visual observation of the waveform and the wideband spectrogram in Praat (Boersma and Weenink, 2017). All annotation work was done by the same trained German native speaker to keep segmentation and annotation consistent. Given the characteristics of continuous speech, we established criteria according to which suitable vowels were selected. First, only vowels occurring in intelligible, phonated words were annotated. Second, 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 at least one period before vowel offset. Third, 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. And fourth, vowels immediately following nasals, glides or other vowels were not selected.

ACOUSTIC ANALYSIS

Acoustic measures were obtained with Praat (Boersma and Weenink, 2017). Automatic scripts were run to determine the formant frequencies of F1 and F2 in Hertz (Hz) from the entire duration of the stable part of each selected vowel. With the obtained formant frequencies, we computed the following five vowel measurements: vowel formant contrasts for each speaker, F1 and F2 variability within each speaker, the vowel space area (VSA), the vowel articulation index (VAI) and the F2 ratio of the vowels /i, ɪ/ and /u, ʊ/.

To measure the vowel contrasts for each speaker individually, we ran ANOVAs and subsequent *post hoc* comparisons with the dependent variables F1 and F2 frequencies and vowel as independent variable. This measurement served as an index of whether the formant frequencies of different vowels are distinct or not. We expected F1 frequencies to differ between the vowels /a, aː/ and /i, ɪ, u, ʊ/ and F2 frequencies between /i, ɪ/ and /a, aː, u, ʊ/. Accordingly, we expressed this measurement as a ratio of expected contrasts to observed contrasts, with a ratio of 1.0 indicating full contrasts between vowels and a lower ratio indicating reduced vowel contrasts.

The F1 and F2 variabilities were computed for each speaker individually as the mean standard deviation of each vowel respectively. According to Kim et al. (2011a) this measurement reflects a speaker's relative stability of achieving vowel targets. For VSA, VAI and the F2 ratio the formant frequencies were averaged over vowel and speaker. VSA is expressed as the following formula (Liu et al., 2005):

$$\mathbf{VSA} = \frac{S(i,a,u) + S(a,u,i) + S(u,i,a)}{2}, \tag{2.1}$$

where

$$\mathrm{S}(v1, v2, v3) = \mathrm{F1}_{v1} \times (\mathrm{F2}_{v2} - \mathrm{F2}_{v3})$$

The VAI calculation was based on that of Roy et al. (2009):

$$\mathbf{VAI} = \frac{\mathrm{F1}_a + \mathrm{F2}_i}{\mathrm{F1}_i + \mathrm{F1}_u + \mathrm{F2}_a + \mathrm{F2}_u} \tag{2.2}$$

INTELLIGIBILITY RATING

As a rough measure of speech impairment severity, the intelligibility of each participant's speech was rated by 15 untrained listeners. The listeners were German native speakers, between 20 and 40 years of age who had no training in phonetics or background related to speech pathologies.

For the intelligibility ratings two words and two short phrases were randomly selected from each monologue resulting in 46 words and 46 phrases in total. Both words and phrases included at least one of the selected vowels. The listeners were instructed to rate the intelligibility of each word and phrase on a scale from 1 (very poor intelligibility) to 6 (very high intelligibility). No time restrictions were imposed on the rating tasks and listeners were allowed to listen to the words and phrases as many times as needed.

2.2.2. RESULTS

Table 2.2 summarizes the results of the vowel measurements and averaged intelligibility rating scores for each group divided by gender. For the male groups we found the predicted pattern of vowel articulation precision: the MCI group yielded lower values for VSA, VAI, F2 ratio and formant frequency contrasts than the PD and the HC groups. As expected, higher values were found in the formant frequency variabilities for the MCI group compared to the PD and HC groups. The vowel measurement results of the female groups are more ambiguous: the female MCI group scored lowest only in the VAI. In all other measures, except for the F2 contrast, the PD group performed the poorest among the female participants.

Kruskal-Wallis rank sum tests for non-parametric data were conducted to determine group differences across the data. The overall comparison of PwPD (including both PD group and MCI group) and control speakers yielded significant differences for the

Table 2.2 | Summary of vowel measurements for each group divided by gender, where **F2 i/u ratio** is the ratio of /i/ and /u/ vowels, F1/F2 contrast are F1 and F2 contrasts (as ratios), **F1/F2 variability** are F1 and F2 variabilities (mean(sd)), **IS** is the intelligibility score

			Vowel measurements							
Gender	**Group**		**VSA**	**VAI**	**F2 i/u ratio**	**F1 contrast**	**F2 contrast**	**F1 variability**	**F2 variability**	**IS**
M	**HC**	M	105 105	0.86	1.76	1	0.93	50	162	5.2
		SD	63 348	0.09	0.42	0	0.15	4.5	41	
	PD	M	60 707	0.74	1.43	0.83	0.7	48	139	4.1
		SD	76 967	0.09	0.46	0.24	0.33	12	29	
	MCI	M	42 742	0.72	1.35	0.38	0.56	118	297	3.2
		SD	19 217	0.04	0.12	0.49	0.36	116	95	
F	**HC**	M	246 025	1.00	2.42	1	1	42	130	5.3
		SD	30 710	0.04	0.23	0	0	5.5	31	
	PD	M	194 988	0.96	2.13	0.79	1	68	227	4.3
		SD	97 948	0.11	0.57	0.36	0	36	44	
	MCI	M	221 496	0.94	2.39	1	1	48	167	4.1
		SD	89 311	0.02	0.33	0	0	7.4	31	

measurements VAI (H(2) = 4.6, $p < .05$) and F1 contrast (H(2), $p < .05$). When MCI was included as a factor, subsequent *post hoc* tests showed a significant difference between the F2 variability values of the MCI group and the control group. To assess how this finding is related to gender differences, and since Wilcoxon rank sum tests revealed differences for the measures of VSA (W = 118, $p < .05$), VAI (W = 121, $p < .05$), F2 ratio (W = 115, $p < .05$) and F2 contrast (W = 94.5, $p < .05$) between men and women in our data, we ran separate analyses for the male and female participants.

Male participants' vowel measurements (except for the F1 variability measures) confirm the expected pattern of decrease of vowel space, formant contrasts and stability of achieving vowel targets in the MCI group compared to the PD and HC groups (see Figure 2.1). This finding is further reflected by a lower intelligibility score for the MCI group. Kruskal-Wallis tests were run to assess the significance of the observed trend. Significant differences were found for the vowel measurements of VAI (H(2) = 6.2, $p < .05$), F1 contrast (H(2) = 6, $p < .05$) and F2 variability (H(2) = 8, $p < .05$). Subsequent *post hoc* analyses revealed that differences between the MCI group and the HC group accounted for the significance. The PD group did not differ from the MCI or the HC group.

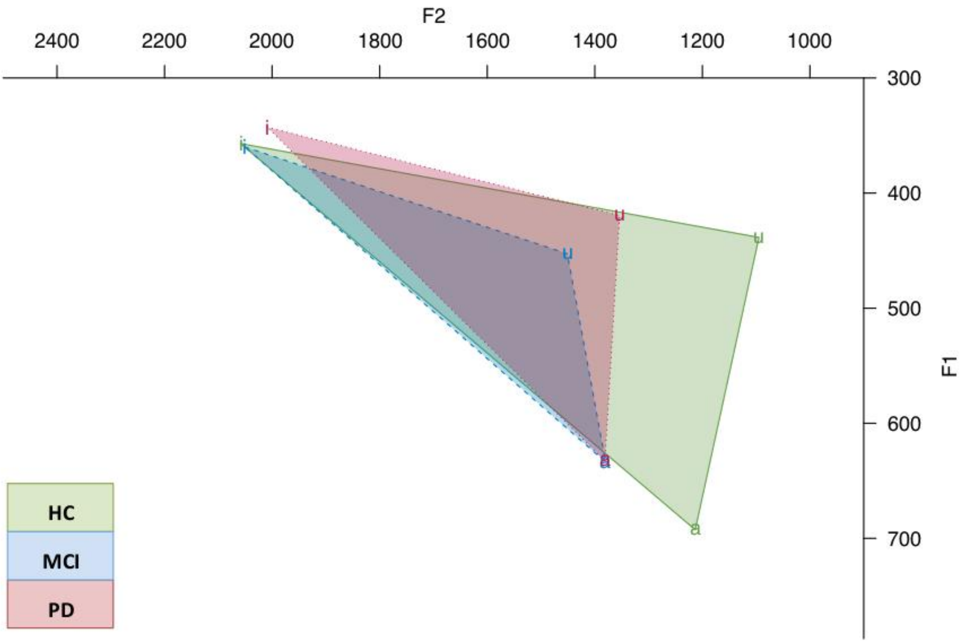

Figure 2.1 | Male vowel space areas. VSA plots with dotted lines reflect the PD group, VSA plots with scattered lines reflect the MCI group and solid lined VSA plots correspond to the HC group

The pattern of vowel measurement results was less consistent for the female participants (see Figure 2.2). Although the intelligibility scores for the three female groups show the expected trend, with the MCI group being the least intelligible one, the analysis of the PD group yielded the poorest results in almost all vowel measurements (except for the

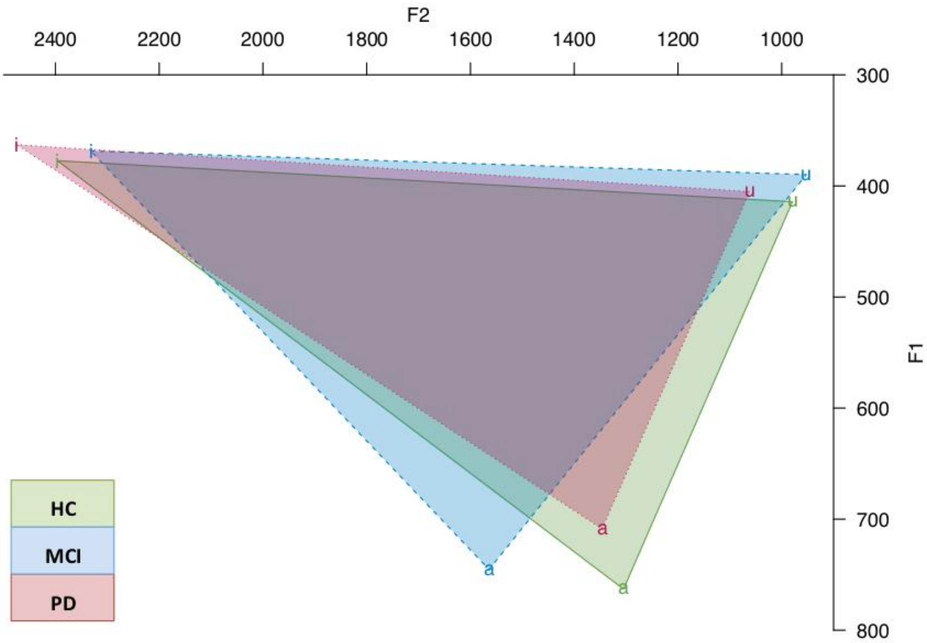

Figure 2.2 | Female vowel space areas. VSA plots with dotted lines reflect the PD group, VSA plots with scattered lines reflect the MCI group and solid lined VSA plots correspond to the HC group.

VAI and the F2 contrast). Accordingly, the significant difference found for F2 variability (H(2) = 6.2, $p < .05$) was between the PD group and the HC group. Vowel measures of the MCI group did not differ significantly from the HC and PD group.

Intelligibility ratings between groups differed as expected: the intelligibility of the MCI group was rated the lowest and the intelligibility of control participants was rated the highest. Kruskal-Wallis tests and subsequent *post hoc* tests showed significant differences ($p < .05$) between the HC and MCI group and between the HC and PD groups. No correlation was found between intelligibility and vowel measures.

2.2.3. DISCUSSION

Regarding the first aim of the current study, our findings are in line with a previous study by Rusz et al. (2013b), indicating that acoustic analysis of spontaneous speech is sensitive enough to separate dysarthric from non-pathological speech at the group level. Even with the small sample size in this study, we were able to acoustically detect imprecise vowel articulation.

Concerning the second aim, vowel measurements that proved to be most sensitive in this study were the vowel articulation index (VAI) and the F1 contrast. While the first measurement is related to vowel space, the F1 contrast is an index of how distinct a speaker's formant frequencies between different vowels are. Moreover, the separate analyses for men and women yielded significant effects of the F2 variability measurement, which reflects speaker's steadiness in achieving vowel targets (Kim et al., 2011a).

Following our third aim, we found that the effects of cognitive impairments on speech of PwPD remain inconclusive. As speech motor control requires more attention capacity in PwPD than in neurologically healthy individuals, we expected individuals with PD and additional cognitive impairments to exhibit less precise vowel articulation than individuals with PD only, because of their reduced attention capacity. While vowel measurements and intelligibility rating showed the expected trend for male participants, the pattern of vowel measurements and intelligibility rating was less clear for the female speakers. This lack of clarity could be attributable to in-group variation relative to the scarcity of data, especially among the female speakers. Although this study focused on vowel articulation precision, we stress that metrics of vowel articulation should not be treated as a single parameter to differentiate dysarthric from healthy speech and to investigate the effects of cognitive decline. Thus, future research should include a larger sample size, more balanced sets of groups and further acoustic measurements to better understand the effects of cognitive impairment on speech motor control in PD. Moreover, future studies investigating the effects of cognitive impairments on speech of PwPD should also explore the effects of speakers' gender.

2.3. PROSODY AND VOWEL ARTICULATION IN SPONTANEOUS SPEECH OF DUTCH PWPD

This pilot study investigates the acoustic correlates of prosody and vowel articulation in Dutch PwPD. Following the pilot study on German PwPD described in section 2.2, we explored acoustic correlates of articulatory deficits in the spontaneous speech of Dutch PwPD with additional focus on acoustic correlates of prosodic deficits related to HD. A common way to track prosodic deficits in dysarthric speech is through the analysis of disturbances in fundamental frequency (f_0) (Skodda et al., 2009; Holmes et al., 2000), intensity (Ramig et al., 2001), stress (Cheang and Pell, 2007), and speech rate and rhythm (Skodda and Schlegel, 2008). While speech rate and intensity have been shown to yield inconsistent results (Skodda et al., 2009; Rosen et al., 2005), monopitch has been described as the most common deviant prosodic dimension in hypokinetic dysarthria (Darley et al., 1969b; Holmes et al., 2000; Anand and Stepp, 2015). Monopitch is manifested as a lack of f_0 variability, and the typical consequence of this deficit is the reduced ability to achieve certain intonation contours, such as questions and statements. Therefore, PwPD may experience difficulties in expressing linguistic or emotional prosody. Such monotonous speech of PwPD has been described as withdrawn or cold by listeners (Jaywant and Pell, 2010).

To the best of our knowledge, at the time of writing this section, acoustic studies on spontaneous speech of Dutch PwPD were scarce. Thus, with this study we address the question of whether spontaneous speech reflects the common monopitch and vowel centralization trends of dysarthric speech of Dutch PwPD. Moreover, based on the previous findings that discussed gender effects in prosodic measurements (Skodda et al., 2011c), we also explore the gender effects on f_0 variability.

Table 2.3 | Summary of speaker group demographics. Age and disease duration are given in years.

		PD	HC
Male: Female ratio		6:9	6:9
Age	(mean ± SD)	65.1 ± 7.8	65 ± 8
Disease duration	(mean ± SD)	7.3 ± 3.6	-
Hoehn & Yahr scores	(mean ± SD)	2.0 ± 0.7	-

2.3.1. METHODS

Recordings of spontaneous speech used in the present pilot originate from the study by Harris et al. (2016). 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.

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. The second group was comprised of 15 neurologically healthy controls matched with the PD group for age and gender. Table 2.3 summarizes the demographic data of both groups.

SPEECH TASK, RECORDING PROCEDURE AND ANNOTATION

Patients were recruited all over the Netherlands, and the recordings were made at their homes. For the purpose of the study of Harris et al. (2016) 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. For more detailed information on data collection, participants profiles and speech tasks, see Harris et al. (2016).

In this study the 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 /ɑ, ɪ, uː/. The procedure for annotation and vowel selection criteria were the same as described in the pilot study on German PwPD (section 2.2).

ACOUSTIC ANALYSIS

Acoustic measures were obtained with the Speech Signal Toolkit (SPTK) for Python (Imai et al., 2017) and with speech analysis software Praat (Boersma and Weenink, 2017). SPTK toolkit allowed us to track f_0 based on the robust algorithm for pitch tracking – RAPT (Talkin, 1995). Praat scripts were used to estimate the speech rate (De Jong and Wempe, 2009) and to obtain formant frequency values that were subsequently used in calculating the vowel articulation measures. In this study we investigated two prosodic characteristics – rate (speech and articulation rates) and pitch. Usually 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 with a Praat script written by De Jong and Wempe (2009). In this algorithm, syllable nuclei correspond to peaks in intensity preceded and followed by dips in intensity, with unvoiced peaks being discarded. This script was shown to be informative for the study on French dysarthric speech (Looze et al., 2012). In our study we have used a -20 dB silence threshold, 4 dB dip and 70 ms as a minimal pause duration. Speech rate measurement was defined in terms of number of syllables divided by total time of the recording and articulation rate was defined as number of syllables divided by speaking time in the recording.

Pitch tracking was performed with the David Talkin's RAPT algorithm (Talkin, 1995) implemented in the SPTK toolkit (Imai et al., 2017). The RAPT algorithm uses cross-correlation to identify pitch candidates and dynamic programming to select the "best fit" pitch value for each frame (Talkin, 1995; Morrison et al., 2007). From the pitch trajectory we calculated pitch variance estimation as the average of the squared deviations from the mean of f_0 (see equation 2.3) and pitch span (the estimation of the speaker's range of frequencies) as difference between minimum and maximum of f_0 values.

$$\mathrm{Var}(f_0) = \mathrm{mean}|f_0 - \mathrm{mean}(f_0)|^2 \tag{2.3}$$

To determine vowel articulation differences, we have calculated four measurements in the same way as described in the German pilot study (see section 2.2): F1 and F2 variability for each speaker, the vowel space area (VSA), the vowel articulation index (VAI), and the F2 ratio of the vowels /i, ɪ/ and /u, uː/.

According to Kim et al. (2011a), the F1 and F2 contrasts reflect the speaker's relative stability in achieving vowel targets. These measurements were computed based on the descriptions from Kim et al. (2011a) and Strinzel et al. (2017). However, we introduced normalization to allow relative comparison of different vowels. For each speaker we calculated mean normalized standard deviation of each vowel.

2.3.2. RESULTS AND DISCUSSION

Table 2.4 summarizes the results of the prosodic measurements for each group. The f_0 variability and span were calculated for every 10 seconds within the recording. As expected, the PD group showed lower values of f_0 variability (see Figure 2.3). Speech and articulation rates were calculated for the whole duration of each recording, and as expected prosodic measurements for f_0 variability and span were lower for the PD group, except for the speech and articulation rates.

Table 2.5 summarizes the results of the vowel measurements for each group. We found the predicted pattern of vowel articulation precision in the data: the values of VSA (see Figure 2.4), 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 groups showed significant differences for the measurements of f_0 variability ($\chi^2 = 5.8,\ p < 0.02$), 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 distribu-

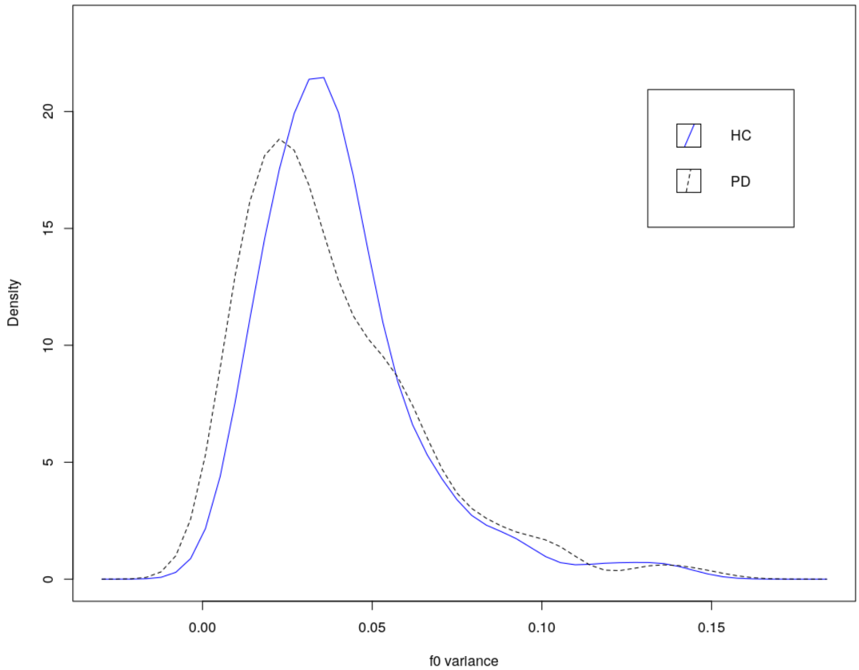

Figure 2.3 | f_0 variance for the PD and HC groups.

Table 2.4 | Summary of prosodic measurements for each group, where f_0 variance is estimation of pitch variability, f_0 span is the estimation of speaker's range of frequencies.

Group	**Prosodic measurements**			
	f_0 variance	**f_0 span**	**Speech rate**	**Articulation rate**
PD (mean ± SD)	0.038 ± 0.027	1.06 ± 0.19	2.57 ± 0.29	4.09 ± 0.54
HC (mean ± SD)	0.040 ± 0.023	1.08 ± 0.16	2.81 ± 0.29	4.23 ± 0.33

tion 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 (Skodda and Schlegel, 2008; Canter, 1963). However, this inconsistency may be accounted for with the methodological differences and small sample size relative to Skodda and Schlegel (2008), as well as possible differences in pause distribution that was not accounted for in this study. On the other hand, Skodda and Schlegel (2008) demonstrated that speech rate is heterogeneous within the population of PD speakers.

To assess if f_0 variability is related to gender differences, we ran separate analyses

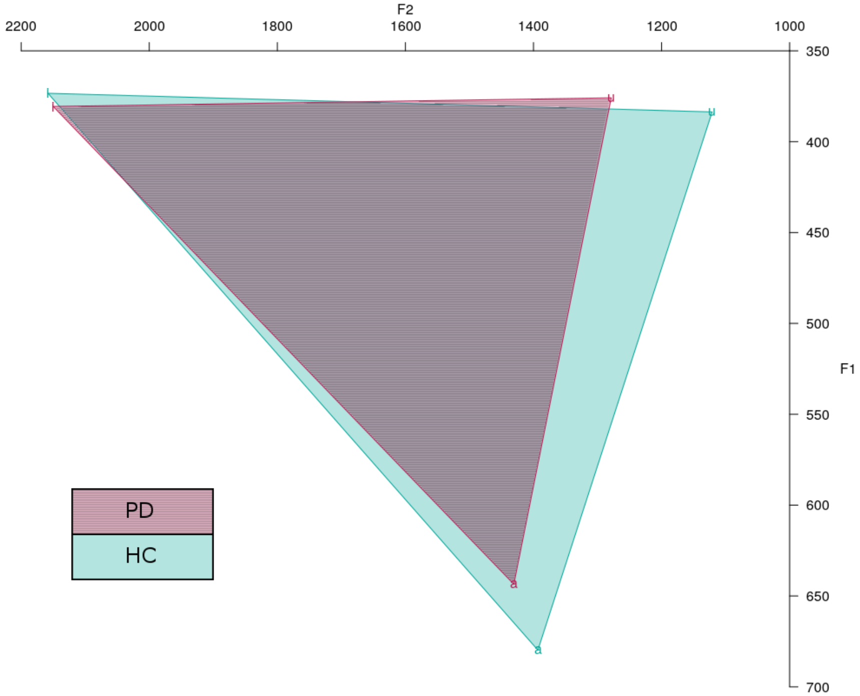

Figure 2.4 | VSA for the PD and HC groups.

for male and female participants. The most affected group was male PwPD. The comparison between group and gender pairs showed significant differences except for the HC speakers: the f_0 variability didn't differ significantly between male and female HC participants. This finding is in contrast with the previous study on gender-related patterns of dysprosody by Skodda et al. (2011c), who attributed significant differences in f_0 variability of their male and female HC speakers to the gender differences of intonation range. This inconsistency may be attributed to the smaller sample size but not to the potential effect of the gender differences induced by the Hertz-based measures, since Skodda et al. (2011c) also relied on Hertz-based measures. Additional research is required to explore the possibility of a different gender-related dysprosody patterns.

f_0 variability, speech rate, VAI, F2 ratio and F1 variability proved to be sensitive to differentiate dysarthric from non-pathological speech on a group level. The first two measurements, f_0 variability and speech rate, account for clear dysprosody patterns, suggesting that monopitch and abnormal speech rate are also a common feature for Dutch dysarthric speech. The vowel articulation index (VAI) and F2 ratio are related to the vowel space, confirming the hypothesis of vowel centralization, while the significant differences in F1 variability reflect differences in speakers' "steadiness" in achieving vowel targets (Kim et al., 2011a).

Overall, these results are in line with Rusz et al. (2013b) and Strinzel et al. (2017).

Table 2.5 | Summary of vowel measurements for each group, where **F2 ratio** is the ratio of /i/'s and /u/'s second formants, **F1/F2 variability** are normalized F1 and F2 variabilities (mean(sd/mean)).

Group		Vowel measurements				
		VSA	VAI	F2 i/u ratio	F1 variability	F2 variability
PD	M	115 500	0.79	1.6	0.12	0.13
	SD	59 552	0.06	0.29	0.0005	0.0002
HC	M	155 100	0.87	1.9	0.12	0.13
	SD	66 700	0.08	0.36	0.0004	0.0001

Dutch spontaneous speech reflected the expected reduced trend in f_0 variability for the PD group, therefore confirming the monopitch tendency common for HD. An acoustic analysis of Dutch vowel articulation in spontaneous speech proved to be sensitive enough to differentiate dysarthric and non-pathological speech, as it was previously shown for German spontaneous speech (Strinzel et al., 2017).

The lack of consistency with previous studies was expected in with regards to the effects of gender on acoustic measures and could be attributed to in-group variation due to scarcity and imbalance of the data, as well as methodological differences. Thus, future research should include a larger sample size, more balanced groups and further explore acoustic measurements to better understand the nature of deficits in Dutch spontaneous dysarthric speech.

2.4. CONCLUSIONS

This chapter provides an general overview of the manifestations of HD and research methodologies employed in the studies focusing on HD. The two pilot studies contribute to the growing body of research on both acoustic correlates of vowel articulation in spontaneous dysarthric speech, as well as on acoustic analysis of spontaneous speech of Dutch and German individuals with HD. The reported findings highlight the complexity of the interplay between diverse factors that may influence HD manifestations.

More specifically, these two pilot studies on German and Dutch dysarthric spontaneous speech are in line with the findings of Rusz et al. (2013b) who found that vowel articulation measurements were able to capture even minor abnormalities in speech of speakers in early stages of PD. Both pilot studies have demonstrated the sensitivity of acoustic measurements of vowel articulation to differentiate dysarthric and non-pathological speech. As expected, the study on spontaneous speech of Dutch PwPD confirmed a common monopitch and centralization trends unequivocally demonstrated by a wealth of research (Darley et al., 1969b; Goberman, 2005; Sapir et al., 2011; Skodda et al., 2011d; Anand and Stepp, 2015; Brabenec et al., 2017). The pilot study on German PwPD also demonstrated the adequacy of acoustic analysis as a methodological approach to detect cognitive decline in PD. Thus, the main contribution of the pilot study on German PwPD is a potential negative effect of cognitive impairment on the speech impairment in PwPD. This finding confirmed our hypothesis that MCI can lead to aggravation of HD symptoms, which is in line with reports of affected language function in PwPD with more severe cognitive issues such as Alzheimer's disease (Goldman and Litvan, 2011). The findings in the pilot study of German PwPD demonstrated that the effect of MCI on speech was both acoustically measurable and recognizable by listeners.

Yet further research is warranted in order to explore the nature of the different effects on the speech production of PwPD. Future research into acoustic characteristics of speech of PwPD should further explore gender-related patterns of articulation and prosody deficits as well as the effects of non-motor symptoms of PD such as MCI on HD manifestations in speech.