Journal of Parkinson's Disease 5 (2015) 291–299
DOI 10.3233/JPD-140451
IOS Press

Research Report

Speech Prosody Across Stimulus Types for Individuals with Parkinson's Disease

Joan K.-Y. Ma[a,1], Christine B. Schneider[b,c,*,1], Rüdiger Hoffmann[d] and Alexander Storch[b,c,*]
[a]*Clinical Audiology, Speech and Language Research Centre (CASL), Queen Margaret University, Edinburgh, UK*
[b]*Division of Neurodegenerative Diseases, Department of Neurology, Technische Universität Dresden, Dresden, Germany*
[c]*German Centre for Neurodegenerative Disease (DZNE), Dresden, Germany*
[d]*Institute of Acoustics and Speech Communication, Technische Universität Dresden, Dresden, Germany*

Abstract.
Background: Up to 89% of the individuals with Parkinson's disease (PD) experience speech problem over the course of the disease. Speech prosody and intelligibility are two of the most affected areas in hypokinetic dysarthria. However, assessment of these areas could potentially be problematic as speech prosody and intelligibility could be affected by the type of speech materials employed.
Objective: To comparatively explore the effects of different types of speech stimulus on speech prosody and intelligibility in PD speakers.
Methods: Speech prosody and intelligibility of two groups of individuals with varying degree of dysarthria resulting from PD was compared to that of a group of control speakers using sentence reading, passage reading and monologue. Acoustic analysis including measures on fundamental frequency (F0), intensity and speech rate was used to form a prosodic profile for each individual. Speech intelligibility was measured for the speakers with dysarthria using direct magnitude estimation.
Results: Difference in F0 variability between the speakers with dysarthria and control speakers was only observed in sentence reading task. Difference in the average intensity level was observed for speakers with mild dysarthria to that of the control speakers. Additionally, there were stimulus effect on both intelligibility and prosodic profile.
Conclusions: The prosodic profile of PD speakers was different from that of the control speakers in the more structured task, and lower intelligibility was found in less structured task. This highlighted the value of both structured and natural stimulus to evaluate speech production in PD speakers.

Keywords: Hypokinetic dysarthria, Parkinson's disease, prosody, intelligibility, stimulus effect

INTRODUCTION

The most common type of speech disorder associated with PD is hypokinetic dysarthria, and is most prominently characterized by disruption in prosody [1, 2]. Prosody refers to the aspects of the speech signal used in carrying both linguistic (e.g., intonation and stress) and paralinguistic (e.g., anger and sadness) meanings. As a result of rigidity of speech musculatures and reduced range of articulatory movements, speech associated with PD is generally

[1]These authors contributed equally to this work.

*Correspondence to: Christine Schneider, M.D. and Alexander Storch, M.D. Division of Neurodegenerative Diseases, Department of Neurology, Technische Universität Dresden, and German Centre for Neurodegenerative Diseases (DZNE) Dresden, Fetscherstrasse 74, 01307 Dresden, Germany. Tel.: +49 351 458 2532; Fax: +49 351 458 4352; E-mail: christine.schneider@uniklinikum-dresden.de. (Christine Schneider), E-mail: alexander.storch@neuro.med.tu-dresden.de. (Alexander Storch).

ISSN 1877-7171/15/$35.00

characterized by monopitch, inappropriate pitch level, reduced stress, monoloudness, inappropriate silences, short rushes of speech, variable speech rate, harsh and breathy voice qualities [2], all could have direct and/or indirect impact on speech prosody. Impairment in speech prosody not only affects the expression of linguistic intonations and emotions; but is also reported to negatively influence overall intelligibility [3, 4], and may thus result in reduced communication efficiency and listeners making negative personality judgement about the speakers [5]. However, only very few studies have explored the effects of impaired speech prosody on intelligibility, indicating that prosody might be the second most impacting factors on speech intelligibility, after articulation [3]. Assessment of speech prosody in PD speakers is thus crucial for diagnosis and speech/language therapy.

Examination of impairment in prosody has been regarded as one of the most difficult aspects for speech and language therapy, one of the main controversies being the type of speech material employed. Structured materials, such as sentence reading and standard passagehave been most commonly used [6]. However, the representativeness of prosody elicited in structured material compared to that of natural speech is questionable [7]. Previous studies have shown that there are differences in the prosodic patterns between structured speech tasks and spontaneous speech tasks: Speakers are more likely to pause at grammatical boundaries in a reading task than in a spontaneous speech task [8], while pauses in spontaneous speech are results of the higher cognitive demand in the planning process [9]. A faster speech rate and a higher average fundamental frequency (F0) were also reported for a reading task when compared to a spontaneous speech task [10].

Another factor which complicates the analysis of prosody in dysarthric speech is the lack of agreed method on how speech prosody should be analysed or quantified. Although perceptual evaluation has been criticized for its high level of subjectivity and interrater variability [11], it is still the "gold standard" in the clinical differential diagnosis of dysarthria including hypokinetic dysarthria [12]. Objective quantification of speech prosody by acoustic analysisis suggested as a more objective and reliable approach, employing different parameters such as average F0 and average intensity [7].

The overall aim of this study thus was to comparatively explore the effects of different types of speech materials on speech prosody and intelligibility in PD speakers. Acoustic analyses were used to examine prosodic variations across different types of speech materials, and direct magnitude estimation (DME) would be used to investigate the variation in speech intelligibility. The relationship between impairment in speech prosody and intelligibility was also be explored.

SUBJECTS AND METHODS

Speakers

Speaker recruitment was performed using a two-step approach: The first step included the recruitment of 38 native German speakers with PD according to the UK Brain Bank criteria [13] as a sample of opportunity by movement disorder physicians. Severity of PD symptoms was assessed using the Unified Parkinson's disease rating scale (UPDRS) [14]. Individuals who had an identifiable cause of parkinsonism or signs for atypical parkinsonian disorders, had history of psychosis or signs of dementia (Mini Mental State Examination [MMSE] $\leq$ 23 points) or other relevant conditions interfering with the study protocol were excluded. All 38 speakers included in this study had normal oral-peripheral structures, acceptable hearing level ($\leq$40 dBHL at 500, 1K and 2K Hz for the better ear), and normal language ability as screened by the Token Test in the Aachener Aphasia Test [15].

The second step included the dysarthria severity rating of all 38 recruited speakers by perceptual analysis of a thirty seconds passage-reading speech sample to exclude patients without dysarthria and to group patients according to dysarthria severity. These speech samples were presented to three expert listeners, who were qualified speech and language therapists with more than 5 years of clinical experience with dysarthric speech, to rate the degree of dysarthria as mild, moderate or severe based on the perception of the degree of deviation from the norm in respiration, phonation, articulation, resonance and prosody. Twelve speakers were rated to have no dysarthria and were thus excluded from the study. Twenty-six speakers were judged to have hypokinetic dysarthria (for details see Table 1). Since only two patients were judged as having severe dysarthria, they were combined into the moderate group (moderate-severe). 24% ($n = 9$) of patients received speech therapy (all according to Lee-Silverman-Voice-Treatment) prior to study inclusion.

Twelve native German speakers were recruited as controls, with ten males and two females (mean age 65.83 years, range 59–75 years). All control speakers had no previous history of speech or language disorders, normal oral-peripheral functions and acceptable

Table 1
Demographic and clinical data of speakers

Speaker	Age [years]	Gender	Disease duration [years]	Hoehn&Yahr stage	UPDRS part III	Dysarthria severity
S1	69	M	16	2	15	mild
S2	66	M	8	1	7	mild
S3	71	F	13	2	12	mild
S4	67	M	1	1	7	mild
S5	57	M	5	1	11	mild
S6	65	M	3	1	7	mild
S7	68	M	3	2	6	mild
S8	65	M	3	2	10	mild
S9	70	M	7	2.5	13	mild
S10	67	M	4	2	12	mild
S11	69	M	14	2.5	15	mild
S12	72	F	9	2	4	mild
S13	79	M	6	3	19	mild
S14	44	F	13	2	17	moderate
S15	72	M	5	2	22	moderate
S16	74	M	15	3	22	moderate
S17	66	M	23	3	13	moderate
S18	65	M	2	2	8	moderate
S19	78	M	7	4	12	moderate
S20	59	F	6	2.5	19	moderate
S21	46	M	4	2	16	moderate
S22	74	M	6	4	15	moderate
S23	75	M	13	4	41	moderate
S24	59	M	10	2	11	moderate
S25	47	M	1	1	14	severe
S26	75	M	4	3	37	severe

hearing level ($\leq$40 dBHL at 500, 1K, and 2K Hz for the better ear).

Speech materials

Three types of speech materials were included: a set of 15 single sentences, a reading passage and a monologue, a German equivalent of the material described previously [12]. The speech recording was carried out in a quiet room, with an unidirectional digital microphone connected to a computer running Audacity software (version 1.2.6) and an E-MU USB sound card. A 10 cm mouth-to-microphone distance was maintained during the recording. All three types of speech materials were collected in the same session, which lasted for about 20 minute. The order of recording of the three speech tasks was counterbalanced across speakers.

An approximately 30-second speech sample was extracted for each stimulus type from each speaker using Praat software (version 5.1.07) for analysis. For single-sentence materials, sentences of different intonation and syllable length were combined with a one second pause separating the sentences. The total number of sentences in the samples differed across speakers owing to different speech rates. For the passage-reading materials, a 30 second sample was edited beginning from the second sentence of the passage. For the monologue task, a 30 second sample of sample was extracted starting from the second sentence of the monologue.

Intelligibility ratings

Five tasks were included in the perceptual ratings, including a global severity rating task, an intelligibility rating task for each of the three types of speech stimulus and the rating of voice quality on the passage reading task. The global severity rating was always presented as the first task, followed by the three intelligibility tasks. The sequence of the three perceptual tasks was counterbalanced across the three expert listeners (5-years experienced native German speech and language therapists). The listeners sat in front of a laptop and the speech samples were presented using an AKG K240 headphone in a quite room.

Global severity rating of the degree of dysarthria (mild, moderate and severe) was undertaken based on the perception of the degree of deviation from the norm in respiration, phonation, articulation, resonance and prosody in a thirty seconds passage-reading speech sample. For each intelligibility rating task, 42 speech samples of the same type of stimuli were presented, with one sample from each of the PD speakers (non-dysarthric, mild, moderate and severe) and four random samples were repeated for reliability measure. The presentation of the 42 stimulus was randomized automatically by the software. A free-modulus direct magnitude estimation (DME) was chosen for the intelligibility rating task, as it was found to be an appropriate scaling methods in perceptual evaluation of intelligibility in dysarthric speech [16]. In free-modulus DME, the listeners give a reference number to the first speech sample they heard, and rate the following samples with reference to the first sample in terms of intelligibility. For example, if 100 was given the first speaker, a rating of 50 for the second speaker would imply that the listeners judged the sample to be half as intelligible as the reference sample, while a rating of 200 would mean that the sample is double in intelligibility when compared to the reference. The reference sample was repeated every five items, and for each item, the listeners could opt to listen to it a maximum of two times. An introduction session was provided by the first author to all three listeners before the perceptual sessions, and written instructions were also provided at the beginning of each task. Additionally, a written definition of intelligibility was also provided, in which intelligibility was as described how easy it was for the listeners to understand the content of the speech. The intelligibility ratings were then converted using the transformation procedure outlined in Engen [17].

Acoustic analysis

In order to establish a comprehensive prosodic profile of intonation, the speech stimuli were analysed for its F0, intensity and rate. Both overall and variability measures were included for the F0 and intensity measures, as proposed by Leuschel and Docherty (1996). The F0 value of the stimuli was analysed using the autocorrelation algorithm of the Praat software (version 5.1.07) [18]. The voiced segment of each syllable was identified visually from a wideband spectrogram and an amplitude waveform display, and the F0 estimates were obtained at about 5 ms interval from the beginning to the end of the voiced segment. All F0 measurements were subsequently converted into semitone (ST) scale for further analysis, as speakers of both genders were involved. Three measures of F0 were then calculated for each stimulus. Average F0 was calculated across the entire speech sample. F0 range for each sample was tabulated by the difference between the maximum and the minimum values within the sample. Additionally, F0 envelop was included to examine the amount of F0 variation between the voiced segment of each syllable within the sentence. In calculating the F0 envelop, the mid-point between the maximum and the minimum value for each voiced segment was calculated, and the average difference between the neighbouring voiced segments was computed throughout the whole sentence. The average F0 envelop was then calculated across the sentences within each speech sample for each individual speaker.

Intensity level was estimated using the averaging method of the Praat software calibrating speech intensity individually to the average level of spontaneous speach (Boersma & Weenink, 2009) at about 5 ms interval from the beginning to the end of each voiced segment, as explained above for the F0 estimates. Three measures of intensity were subsequently calculated, including average intensity, intensity range and intensity envelop, which were calculated in a comparable way to that of the F0 measures detailed above.

Speech rate (syllable per second) was calculated by dividing the total number of syllables in a sentence by the overall duration (in seconds) of the sentence. The average speech rate for the 30 second speech sample was then obtained by calculating the average speech rate across all the sentences.

Statistics

Validation of intelligibility was estimated by intra-rater reliability measured using Pearson's correlation and by intraclass correlation (ICC) for analysing inter-rater reliability. Differences in speech intelligibility and acoustic/prosodic parameters between speakers of different dysarthria severity and between different stimulus types were analysed using a two-way ANOVA with *post-hoc* Tukey Honest Significant Difference (HSD) test. Significant differences was assumed with $p < 0.05$ (two-tailed).

RESULTS

Reliability

Four speech samples (11%) were repeated for each listener in each perceptual task. Intra-rater reliability was measured using Pearson's correlation. The intra-rater reliabilities range from 0.92 to 0.99 among the three raters, with an average of 0.94. Intraclass correlation (ICC) was used to analyse the inter-rater reliability. The reliability coefficient for the three listeners was 0.84, which deemed to be at an acceptable level.

Table 2
Mean values of each acoustic parameter for each speaker groups across stimulus type

	Mild dysarthria			Moderate-severe dysarthria			Control		
	Sentence	Passage	Monologue	Sentence	Passage	Monologue	Sentence	Passage	Monologue
F0 average (semitone)	7.81	6.44	7.39	6.89	5.48	5.66	5.47	3.64	3.41
F0 range (semitone)	17.84	20.08	19.30	17.76	19.17	19.08	21.01	18.95	21.92
F0 envelope(semitone)	1.72	2.15	1.84	1.74	2.14	1.86	2.37	2.11	2.08
Mean Intensity (dB SPL)	72.34	70.65	70.09	69.89	69.58	66.54	69.44	67.13	64.55
Intensity range (dB SPL)	32.37	29.72	33.24	34.03	30.45	30.06	35.87	30.77	32.38
Intensity envelope (dB SPL)	2.74	3.03	3.31	3.55	3.23	3.42	3.26	3.13	3.27
Speech rate (syllable per second)	4.93	3.48	3.06	4.73	3.31	2.5	4.84	3.34	3.31

Table 3
Summary of the statistical analysis for the acoustic parameters

	Group	Stimulus	Group x Stimulus
F0 average	$F = 1.83, p = 0.18$	$F = 36.12, p <$ **0.001**	$F = 3.30, p <$ **0.05**
F0 range	$F = 0.92, p = 0.41$	$F = 0.99, p = 0.38$	$F = 1.29, p = 0.28$
F0 envelope	$F = 1.38, p = 0.86$	$F = 1.08, p = 0.35$	$F = 3.09, p <$ **0.05**
Average intensity	$F = 6.15, p <$ **0.05**	$F = 27.42, p <$ **0.01**	$F = 2.25, p = 0.07$
Intensity range	$F = 0.56, p = 0.58$	$F = 8.10, p <$ **0.05**	$F = 1.60, p = 0.18$
Intensity envelope	$F = 1.72, p = 0.19$	$F = 3.69, p <$ **0.05**	$F = 2.10, p = 0.10$
Rate	$F = 1.22, p = 0.31$	$F = 122.48, p <$ **0.001**	$F = 1.93, p = 0.12$

Results from two-way ANOVA for each acoustic parameter (average F0, F0 range, F0 envelope, average intensity, intensity range, intensity envelope and difference in speech rate) with speaker group (mild, moderate-severe and control) as the between-group factor and stimulus type (sentence, passage and monologue) as the within-group factor.

Intelligibility rating

Intelligibility rating was compared between the two groups of speakers with dysarthria (between-subject factor) and across the three presentation conditions (within-subject factor) using a 2-way ANOVA. Results showed significant difference between the two groups [$F(1,24) = 19.94$, $p < 0.05$], with the mild group (mean $\pm$ SD $= 106.7 \pm 15.9$) showed significantly higher speech intelligibility rating than the moderate-severe group (82.2 ± 18.2). The statistical analysis also showed a significant difference between the three types of speech stimulus [$F(2,48) = 24.35$, $p < 0.05$], with significantly lower intelligibility rating for monologue (104.6 ± 19.1) than for sentence reading (90.5 ± 17.7) and passage reading (88.36 ± 22.7; *post-hoc* Tukey HSD test, $p < 0.05$ for both). No statistically significant difference was observed between sentence reading and passage reading tasks ($p > 0.05$). The interaction effect between speaker group and stimulus type was not significant [$F(2,48) = 0.75$, $p > 0.05$], indicating that the intelligibility rating varied across the three types of speech stimuli in similar ways across the two speaker groups.

Acoustic analysis

Table 2 shows the average values of the F0, intensity and rate measures across intonations for the speaker groups. Two-way ANOVA was used to analyse the data for each acoustic parameter with speaker group (mild, moderate-severe and control) as the between-group factor and stimulus type (sentence, passage and monologue) as the within-group factor (Table 3; Tukey HSD test: $p < 0.001$ for both).

There were main effects for speaker group across the three F0 measures (average, range and envelope) suggesting that speakers with dysarthria had similar overall F0 level and variability compared to that of controls. A main effect for stimulus type was found for F0 average, but not for range and envelope. *Post-hoc* analysis showed that F0 average in the sentence reading task was significantly higher than that in both passage reading and monologue tasks, which could be explained by final F0 increase in the intonation contour of the question stimuli in sentence reading tasks. Consistently, no differences were found in F0 average between passage reading and monologue tasks, as explained by the fact that both stimulus types were made up primarily by statements. Significant interaction effects between speaker group and stimulus type were observed for F0 average and envelope, but not for F0 range. F0 average in the sentence reading task was higher than in the passage reading task for all three speaker groups, but higher than in the monologue task for only the moderate-severe group and the control speakers. For F0 envelope, control speakers showed higher F0 envelope than both the mild and the moderate-severe groups in the sentence reading task. The control speakers produced higher F0 envelope in the sentence reading task than in monologue. While both mild and moderate-severe groups had higher F0 envelope values in the passage reading task than sentence reading task, control speakers showed no contrast in F0 envelope between these two tasks.

For intensity measures, an effect for speaker group was found for average, but not for range and intensity envelope. *Post-hoc* analysis showed that the average intensity of the mild dysarthric speakers was significantly higher than that of the control speakers. No difference was found between the mild dysarthric group and the moderate-severe dysarthric group, or between the moderate-severe group and the control group. The main effect of stimulus type was significant for all intensity measures. *Post-hoc* analysis showed that sentence reading had the highest average intensity, followed by that of passage reading and monologue. The higher intensity for sentence reading could be explained by the higher intensity of both questions and imperatives. Passage reading had a significantly smaller F0 range than sentence reading and a smaller F0 envelope than the monologue, suggesting a smaller degree of intensity variability in the passage reading task than the two other tasks. No difference was observed between the sentence reading and the monologue task in the intensity variability and no interaction effect was noted for the three intensity measures.

The main effect of stimulus type was significant in speech rate measurements. *Post-hoc* analysis showed the highest speech rate for the sentence reading task, followed by that of passage reading and monologue. There was no main effect for speaker group or interaction effect between speaker group and stimulus type.

DISCUSSION

Our examination of the prosodic profile showed differences between PD and control speakers only in average intensity, but not in the other six acoustic parameters. The observation of higher average intensity level in speakers with mild dysarthria than that of the control speakers contradicted the general report of reduced loudness in PD speech [1, 2]. Therefore, the mean intensity of each speaker across the three stimulus types was examined qualitatively to compare the differences among individual speakers as suggested by Leuschel and Docherty [19]. The average intensity of the speakers with mild dysarthria was 71 dB SPL, and eleven of the thirteen speakers with mild dysarthria showed mean intensity values above 69 dB SPL, which is higher than the mean intensity values for both the moderate-severe dysarthric group (69 dB SPL) and controls (67 dB SPL). Four speakers showed consistently high mean intensity values across the three sampling tasks ($\geq$ 69 dB SPL for all three tasks). These qualitative observations suggested that the higher intensity level in mild dysarthric speakers was not related to individual variability. This might be explained by the fact that the two groups of speakers with dysarthria compensated for their speech volume to different degrees, resulting in speakers with mild dysarthria having a higher average intensity than the other two speaker groups.

Acoustic analysis showed no differences in F0 variability measures (both F0 range and envelope) between PD and the controls. This result was intriguing as PD speech is often described as monotonous perceptually [2]. However, the perceptual characteristic of monotonous speech in PD speakers is not always mirrored in studies using acoustic analysis. Harel and colleagues reported that differences in F0 measures in read speech and free speech between PD speakers and control speakers were observed in the "off" state but not in the "on" state [20]. F0 variation increased in PD speakers after the administration of levodopa [21]. In addition to the on/off state of the speaker, Metter and Hanson reported a measurable reduction in F0 variability in PD speakers with severe dysarthria, while no consistent difference in F0 variability was found between PD speakers with mild dysarthria and control speakers [22]. The effects of on/off state and severity of dysarthria could have contributed to the lack of F0 variability differences between PD and controls in this study, as all speakers were in "on" state at the point of speech recording, and most of our participants were rated as mild to moderate in severity of dysarthria. In addition, 24% of PD speakers received Lee-Silverman-Voice-Treatment, a speech treatment that aims at increasing vocal loudness, which might also have influenced F0 variability in these subjects. Other possible explanations for lack of F0 variability between PD speakers and controls could be related to the limited sample size or the used speech materials.

Using DME, speakers with moderate-severe dysarthria had significantly lower speech intelligibility than speakers with mild dysarthria. The difference in speech intelligibility between the two groups was not unexpected, as speech intelligibility tends to deteriorate with the progression of dysarthria. The main prosodic difference between the three speaker groups in the present study was in the average intensity, but not F0 variability. This suggests that the lower intelligibility of the moderate dysarthric speakers compared with the mild dysarthric speakers is likely to be caused by other common deviant speech characteristics in individuals with hypokinetic dysarthria, such as imprecise consonants, harsh voice and hypernasality. While it is beyond the scope of this study, in future it would be interesting to explore the relationship between the relative impairments in articulation, voice quality and resonance together with prosody, and their impact on overall speech intelligibility. The differences between the three types of speech stimulus were compared across the seven prosodic parameters. Differences between the three tasks were found in F0 average, intensity average, range and envelope and speech rate. Speech stimuli produced in the sentence reading task were found to have higher F0 average, higher mean intensity and larger intensity variation than in the passage reading and monologue tasks. This could be explained by the inclusion of various intonations in the sentence reading task, such as questions and imperatives. Generally speaking, questions have a higher mean F0 value than statements, and imperatives are produced with a higher intensity. The difference in speech rate was also consistent with previous reports in the literature, where speech rate was reported to be faster in reading than in spontaneous speech [10].

The three speaker groups varied differentially across the three types of speech stimulus: When the F0 envelope values of the three groups of speakers were compared across the three types of speech materials, the interaction effect showed that the two speaker groups with dysarthria had lower F0 envelope than the control group in the sentence reading task. Controls showed a higher degree of F0 modulations in sentence reading task than in monologue task. This could be explained by the inclusion of question intonations in the sentence reading task, which were typically marked by a gradual F0 decline followed by a final F0 increase; while monologues were dominated by statements marked by gradual F0 decline. However, both groups of speakers with dysarthria showed lower F0 envelope values in sentence reading than passage reading, and no significant contrast between sentence reading and monologue. This supported previous study of intonation patterns of PD speakers, which showed that while PD speakers are able to use similar intonation patterns as controls, they showed reduction in pitch excursion [23].

The effects of speech materials were also observed in speech intelligibility. Speech materials produced as part of a monologue was less intelligible than those in sentence and passage reading tasks. The difference in intelligibility between speech stimuli could be attributed to the demands of the speech task on the speaker, as well as the effect of the speech materials on perception. The relatively higher cognitive demand of the monologue task may have reduced speech intelligibility. Similar findings have been reported in the study of dysfluency, where normal dysfluency was found to be more prevalent in free speech than in reading as a result of the increased cognitive demands [24].

The results of the current study revealed no significant difference in the effects of speech stimulus type on intelligibility between speakers with different severity. This contradicted findings reported in the literature that speech materials have differential effects on intelligibility of dysarthric speakers of different severity [25–27]. Although intelligibility ratings were carried out with different approaches in the above-mentioned studies (orthographic transcription) compared to our study, Yorkston and Beukelman pointed to a close approximation of the two measures [26]. This indicates that different means of quantifying intelligibility probably do not account for the lack of interaction effect between severity and stimulus type. However, in the current study, only two speakers were judged to be severe in their overall intelligibility ratings. As the most significant differences were found between speakers with mild impairment and profound impairment, the slightly skewed spread of severity rating in this current study might explain the differences. Additionally, intelligibility contrast was found to be prominent between word and sentence intelligibility tasks, while all the stimulus types in the current study involved a sentence context.

The findings of our study highlight the need to include different types of speech materials in assessing speech impairment in PD. While the overall measures did not reveal major differences between PD and controls, differences in speech parameters could be observed by comparing their performance across various types of speech stimulus. In clinical practice, intelligibility of speakers with dysarthria is often assessed at sentence and single word level (e.g., Assessment of Intelligibility of Dysarthric Speech [26], Diagnostic Intelligibility Testing [28]). While these assessments provide clinicians with detailed information about the specific breakdown in individuals with dysarthria, they do not take into account speech problems such as dysprosody and its effect on intelligibility in connected speech. It is important to emphasize the need to include more naturalistic materials in clinical assessments. For the purposes of clinical assessment, the results of this study suggest that there is a need to balance between materials that provide detailed information and those that accurately reflect the speech problem in individuals with dysarthria, and that different types of speech stimuli should be included in clinical assessment of dysarthria.

ACKNOWLEDGMENTS

We thank all the speakers for their participation in the study. We gratefully acknowledge Dorothea Senf, Susanne Wächter and Friederike Röhle for serving as expert listeners. We are also grateful to Guntram Strecha for setting up the perceptual experiment and Steffen Kurbis for technical support. This project was in part funded by Alexander von Humboldt Foundation.

CONFLICT OF INTEREST

The authors have no conflict of interest to report. The authors alone are responsible for the content and writing of the paper.

REFERENCES

[1] Logemann JA, Fisher HB, Boshes B, & Blonsky ER (1978) Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *J Speech Hear Disord*, **43**, 47-57.
[2] Darley FL, Aronson AE, & Brown JR (1975) *Motor Speech Disorders*, W. B, Saunders, Philadelphia,
[3] De Bodt MS, Hernandez-Diaz HM, & Van De Heyning PH (2002) Intelligibility as a linear combination of dimensions in dysarthric speech. *J Commun Disord*, **35**, 283-292.
[4] Klopfenstein M (2009) Interaction between prosody and intelligibility. *Int J Speech Lang Pathol*, **11**, 326-331.
[5] Pitcairn TK, Clemie S, Gray JM, & Pentland B (1990) Impressions of parkinsonian patients from their recorded voices. *Br J Disord Commun*, **25**, 85-92.
[6] Penner H, Miller N, Hertrich I, Ackermann H, & Schumm F (2001) Dysprosody in Parkinson's disease: An investigation of intonation patterns. *Clin Linguist Phon*, **15**, 551-566.
[7] Lowit-Leuschel A, & Docherty GJ (2001) Prosodic variation across sampling tasks in normal and dysarthric speakers. *Logoped Phoniatr Vocol*, **26**, 151-164.
[8] Goldman-Eisler F. (1968) *Psycholinguitics: Experiments in spontanous speech*, Academic Press, London.
[9] Barr DJ. (2001) Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In *Oralité et gestualité: Interactions et comportements multimodaux dans la communication*, Cavé C, Guaitella I, Santi S, eds. L'Harmattan, Paris, pp. 597-600.
[10] Hirschberg J. (1995) Prosodic and other acoustic cues to speaking style in spontaneous and read speech. In *Proceedings of the International Congress of Phonetic Sciences*, Stockholm, pp. 36-43.
[11] Kent RD (1996) Hearing and believing: Some limits to auditory-perceptual assessments of speach and voice disorders. *Am J Speech Lang Pathol*, **5**, 7-23.
[12] Ma JK, Whitehill T, & Cheung KS (2010) Dysprosody and stimulus effects in Cantonese speakers with Parkinson's disease. *Int J Lang Commun Disord*, **45**, 645-655.
[13] Hughes AJ, Daniel SE, Kilford L, & Lees AJ (1992) Accuracy of clinical diagnosis of idiopathic Parkinson's disease: A clinico-pathological study of 100 cases. *J Neurol Neurosurg Psychiatry*, **55**, 181-184.
[14] Fahn S, Elton RL, & Committee UD (1987) The Unified Parkinson's Disease Rating Scale. In *Recent developments in Parkinson's disease*., Fahn S, Marsden CD, Calne DB, eds. Macmillan Healthcare Information, Florham Park, pp. 153-163 293-304.
[15] Huber W, Poeck K, Weniger D, Willmes K. (1983) *Aachener Aphasia Test*, Hogrefe, Göttingen.
[16] Weismer G, & Laures J (2002) Direct magnitude estimate of speech intelligibility in dysarthria: Effects of a chosen standard. *J Speech Lang Hear Res*, **45**, 421-433.
[17] Engen T. (1971) Psychophysics. In *Woodworth and Schlosberg's Experimental Psychology*. Methuen, London, pp. 47-86.
[18] Boersma P, Weenink D, Praat: Doing phonetics by computer, http://www.fon.hum.uva.nl/praat/.
[19] Leuschel A, Docherty GJ. (1996) Prosodic assessment of dysarthria: Effects of sampling task and issues of variability. In *Disorders of motor speech: Assessment, treatment, and clinical characterization Paul* H. Brookes Publishing Co, Baltimore, pp. 155-178.
[20] Harel B, Cannizzaro M, & Snyder PJ (2004) Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: A longitudinal case study. *Brain Cogn*, **56**, 24-29.
[21] Skodda S, Gronheit W, & Schlegel U (2011) Intonation and speech rate in Parkinson's disease: General and dynamic aspects and responsiveness to levodopa admission. *J Voice*, **25**, e199-e205.
[22] Metter EJ, & Hanson WR (1986) Clinical and acoustical variability in hypokinetic dysarthria. *J Commun Disord*, **19**, 347-366.
[23] Mennen I, Schaeffler F, Watt N, & Miller N (2008) An autosegmental-metricalinvestigation of intonation in people with Parkinson's Disease. *Asia Pac J Speech Lang Hear Res*, **11**, 205-219.
[24] Mulligan HF, Anderson TJ, Jones RD, Williams MI, & Donaldson IV (2001) Dysfluency and involuntary movements: A new look at developmental stuttering. *Int J Neurosci*, **109**, 23-46.
[25] Hustad KC (2007) Effects of speech stimuli and dysarthria severity on intelligibility scores and listener confidence

ratings for speakers with cerebral palsy. *Folia Phoniatr Logop*, **59**, 306-317.

[26] Yorkston KM, & Beukelman DR (1978) A comparison of techniques for measuring intelligibility of dysarthric speech. *J Commun Disord*, **11**, 499-512.

[27] Carter CR, Yorkston KM, Strand EA, Hamme VL. (1996) Effects of semantic and syntactic context on actual and estimated sentence intelligibility of dysarthric speakers. In *Disorders of motor speech: Assessment, treatment and clinical characterization*, Paul Brookes Publishing Co., Baltimore, pp. 67-87.

[28] Kent RD, Weismer G, Kent JF, & Rosenbek JC (1989) Toward phonetic intelligibility testing in dysarthria. *J Speech Hear Disord*, **54**, 482-499.