Acoustic Analysis of Intonation in Parkinson’s Disease

Joan K.Y. Ma\textsuperscript{1}, Rüdiger Hoffmann\textsuperscript{2}

\textsuperscript{1}Speech and Hearing Sciences, Queen Margaret University, United Kingdom
\textsuperscript{2}Institute of Acoustics and Speech Communication, Dresden University of Technology, Germany

\texttt{jma@qmuc.ac.uk, ruediger.hoffmann@ias.et.tu-dresden.de}

Abstract

The aim of this study was to explore the prosodic characteristics of speakers with Parkinson’s disease (PD) in the marking of intonations. Twenty-four German PD speakers with either mild or moderate degree of dysarthria were compared with twelve non-dysarthric control speakers on the production of imperatives, questions and statements. Acoustic analyses of fundamental frequency (average F0, F0 range and F0 envelop), intensity (average intensity, intensity range and intensity envelop) and speech rate (number of syllable per second) were conducted to investigate the effect of PD on intonation marking. The results showed that the dysarthric and non-dysarthric speakers differed significantly in all F0 measures, with higher average F0 and reduced F0 variability noted for the PD speakers. Although the PD speakers were more monotonous in prosody, they showed similar intonation contrasts between intonations as in non-dysarthric speakers.

Index Terms: intonation, dysarthria, Parkinson’s disease

1. Introduction

Parkinson’s disease (PD) is a progressive neurological disease which involves the loss of nerve cells in the substantia nigra and the reduction of dopamine content in the striatum, resulting in reduced muscular control over the oral-facial musculature, laryngeal muscles and other physiological mechanisms involved in speech production [1]. It has been estimated that at least 89% of PD patients experience speech problems over the course of the disease [2]. Hypokinetic dysarthria is the most common type of speech disorder associated with PD, and is most commonly characterized by disruption in speech prosody [3].

Intonation is a universal feature shared by different languages. It plays an important role in conveying information about various characteristics of human communication, such as the differences between imperatives, questions and statements, and information about the speaker’s state of mind and attitude [4]. The communication of intonation in speech is achieved through relative changes in fundamental frequency (F0), intensity and timing of speech. These acoustic features may be combined in various ways to express different intonations. For example, imperatives are generally expressed with lower F0 and increase in intensity, so as to convey the meaning of authoritativeness.

PD speech has been generally described as monotonic, inappropriate pitch level, reduced stress, monoloudness, inappropriate silences, short rushes of speech, variable speech rate, harsh and breathy voice qualities and reduced intelligibility [3]. The disturbances in speech prosody could have a significant impact on the marking of intonations in PD speakers. For example, a smaller than usual rise in the final-F0 of questions in PD speakers might lead to misperception of the intended intonation. Using perceptual experiments to investigate the intonations and emotions produced by PD speakers, Pell and colleagues [5] found that listeners had difficulties to identify questions from statements produced by PD speakers. This poses an interesting research question on how PD affects intonation marking at the acoustic level.

Although it is generally agreed that dysprosody is a common characteristic of PD speech, the intonation contrast in PD speakers was examined mainly in only one or two dimensions, such as average F0 and speech rate [6, 7]. The main findings of these studies suggested that the majority of the PD speakers were less effective in modulating the contrast parameters in achieving the intonation contrast, while speech rate was found to vary as a function of intonation. However, the results were inconclusive when only a small number of acoustic measures were included for two reasons. First, intonation marking is multidimensional, which allows the combination of various acoustic cues in the expression of intonations. In studying the marking of intonation contrast in dysarthric speakers with cerebral palsy, Patel [8] found empirical evidence of cue trading in some of the speakers. That is, some speakers might use loudness instead of F0 in marking certain intonation contrasts, owing to the reduced control over F0 variation. As a result, it is important that in studying the intonation production of PD speakers, different prosodic dimensions, such as F0, intensity and timing of speech, should be included. Second, Kent and Kim [9] argued that overall measures such as average F0 do not provide information about the F0 movement within sentences. Weismer and Liss [10] also criticized that it might not be entirely appropriate to use the same set of acoustic parameters in measuring speech of dysarthric speakers and non-dysarthric speakers, as what are important to speech production of nondysarthric speakers might not be relevant to speakers with dysarthria. As a result, in addition to overall measures such as average F0, parameters measuring range and variation within sentences need to be included. Leuschel and Docherty [11] proposed a total of 14 different dimensions to represent the most prominent prosodic aspects of dysarthric speech, and suggested that parameters representing different prosodic dimensions should be included to capture the speaker’s prosodic profile.

The aim of this study was to investigate how PD affects the expression of intonations (imperatives, questions and statements) on different prosodic dimensions. It was hypothesized that PD speakers were less effective in modulating the contrast in different acoustic parameters, especially in F0 variability and overall intensity level, as monopitch and reduced loudness are the most prominent features of PD speech.

2. Method

2.1. Speakers

Twenty-four native German speakers were recruited from the Clinic and Polyclinic of Neurology, Carl Gustav Carus Dresden University Hospital. They were all diagnosed by neurologists as having idiopathic Parkinson’s disease. All speakers were judged to demonstrate hypokinetic dysarthria, based on a passage reading speech sample, by three speech-
language pathologists, who are native German speakers and have more than 5 years of clinical experience working with dysarthric speech. A global severity rating of the degree of dysarthria (mild, moderate and severe) was also undertaken based on the judges’ perception of the same passage reading speech sample. The mild dysarthric group consisted of 13 speakers and the moderate group included 11 speakers. All speakers had normal oral-peripheral structures, acceptable hearing ability (≤ 40 dBHL at 500, 1K, 2K and 4K Hz for the better ear), and normal language ability as screened by the Token Test in Aachener Aphasia Test [12].

Twelve native German speakers were recruited as control, with two female and ten male. The mean age of the control group was comparable with the dysarthric groups. All control speakers had no history of speech or language disorders, normal oral-peripheral functions and acceptable hearing (≤ 40 dBHL at 500, 1K, 2K and 4K Hz for the better ear). Table 1 summarizes the demographic data of the speakers.

Table 1. Demographic data of the speakers.

<table>
<thead>
<tr>
<th></th>
<th>No. of speakers</th>
<th>Gender</th>
<th>Age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mild</td>
<td>13</td>
<td>11 M, 2 F</td>
<td>57-79</td>
</tr>
<tr>
<td>Moderate</td>
<td>11</td>
<td>9 M, 2 F</td>
<td>44-78</td>
</tr>
<tr>
<td>Control</td>
<td>12</td>
<td>10 M, 2 F</td>
<td>59-75</td>
</tr>
</tbody>
</table>

2.2. Speech Materials

A set of single sentences consisted of 15 German sentences of three intonation types (statement, question and imperative) were recorded for each speaker. The five sentences of each intonation type varied in the number of syllables (7, 9, 11, 13 and 15 syllables). Data collection was carried out in a quiet room with a Shure (SM48) microphone connected to a computer running the Wavesurfer software (version 1.8.5) and a Creative E-MU 0404 USB sound card. A 10-cm mouth-to microphone distance was maintained. During the recording, one sentence was presented on the computer screen with the appropriate punctuation for statement, question or imperative. Speakers were instructed to read each sentence carefully and produce the appropriate intonation accordingly. The stimuli were presented using Microsoft PowerPoint®, and the order of the sentences was randomized automatically by the program.

2.3. Data analysis

In order to establish a comprehensive prosodic profile of intonation, the speech stimuli were analyzed 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 [11]. The F0 value of the stimuli was analyzed using the autocorrelation algorithm of the Praat software (version 5.1.07) [13]. The voiced segment of each syllable was identified visually from a waveband 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 sentence. F0 range for each sentence was tabulated by the difference between the maximum and the minimum values within the sentence. 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.

Intensity level was estimated using the averaging method of the Praat software [13] 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. As the speech rate of PD speakers has been described as both faster [14] and slower [15] than non-dysarthric speakers, instead of comparing the speech rate across the speaker groups directly, the difference in speech rate between the three intonations was calculated to investigate if the relative relationship between intonations was maintained by the speakers. That is, the absolute difference in speech rate between questions and statements, statements and imperatives, and imperatives and questions across stimuli with the same number of syllables for each speaker was calculated for further analysis.

3. Results

Tables 2 and 3 summarize the average values of the F0, intensity and rate measures across intonations for the three groups of speakers. A two-way ANOVA was used to analyze the data for each acoustic parameter (average F0, F0 range, F0 envelop, average intensity, intensity range, intensity envelop and difference in speech rate), with intonation (imperative, question and statement) and speaker group (mild, moderate and control) being the independent variables.

3.1. Fundamental frequency

Statistical analyses revealed that the three groups of speakers differed in terms of average F0 [$F (2, 531) = 46.50, p < 0.05$], F0 range [$F (2, 531) = 7.29, p < 0.05$] and F0 envelop [$F (2, 531) = 21.94, p < 0.05$]. Post-hoc analysis showed that the average F0 values of speakers with mild dysarthria (mean = 7.74, SD = 3.56) and moderate dysarthria (mean = 7.12, SD = 4.87) were significantly higher than that of control speakers (mean = 6.48, SD = 3.72) (Tukey HSD test, $p < 0.05$ for both). The F0 range of speakers with mild dysarthria (mean = 11.32, SD = 2.80) and moderate dysarthria (mean = 11.10, SD = 3.56) was significantly narrower than that of the control speakers (mean = 13.66, SD = 3.56) (Tukey HSD test, $p < 0.05$ for both). Significantly smaller F0 envelop values were also observed for both speakers with mild dysarthria (mean = 1.67, SD = 0.82) and moderate dysarthria (mean = 1.72, SD = 0.80) than that of the control speakers (mean = 2.34, SD = 0.88) (Tukey HSD test, $p < 0.05$ for both), suggesting that speakers with dysarthria had less F0 variation within the sentence when compared with control speakers. No statistically significant difference was found between the speakers with mild and moderate dysarthria in average F0, F0 range and F0 envelop (Tukey HSD test, $p > 0.05$ for all).

Significant main effect of intonation was found for average F0 [$F (2, 531) = 4.41, p < 0.05$] but not for F0 range [$F (2, 531) = 2.21, p > 0.05$] and F0 envelop [$F (2, 531) = 1.38, p > 0.05$]. The average F0 of questions (mean = 7.31, SD = 4.13) was significantly higher than that of statements (mean = 6.07, SD = 4.22) (Tukey HSD test, $p < 0.05$), but there was no significant difference between questions and imperatives (mean = 6.95, SD = 4.07) and statements and imperatives (Tukey HSD test, $p > 0.05$ for both).
Table 2. The average acoustic values for F0 (in ST) and intensity measures across imperatives (I), questions (Q) and statements (S) for the three groups of speakers.

<table>
<thead>
<tr>
<th></th>
<th>Mild dysarthria</th>
<th>Moderate dysarthria</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>Q</td>
<td>S</td>
</tr>
<tr>
<td>Mean F0</td>
<td>7.94</td>
<td>8.14</td>
<td>7.14</td>
</tr>
<tr>
<td>F0 range</td>
<td>12.02</td>
<td>11.01</td>
<td>10.94</td>
</tr>
<tr>
<td>F0 envelop</td>
<td>1.75</td>
<td>1.65</td>
<td>1.63</td>
</tr>
<tr>
<td>Mean Intensity</td>
<td>72.62</td>
<td>72.99</td>
<td>71.37</td>
</tr>
<tr>
<td>Intensity range</td>
<td>23.36</td>
<td>22.01</td>
<td>22.99</td>
</tr>
<tr>
<td>Intensity envelop</td>
<td>2.80</td>
<td>2.66</td>
<td>2.83</td>
</tr>
</tbody>
</table>

Table 3. The average differences in rate between imperatives (I), questions (Q) and statements (S) across the speaker groups

<table>
<thead>
<tr>
<th></th>
<th>Mild</th>
<th>Moderate</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q Vs S</td>
<td>0.83</td>
<td>0.82</td>
<td>0.84</td>
</tr>
<tr>
<td>S Vs I</td>
<td>0.64</td>
<td>0.82</td>
<td>0.82</td>
</tr>
<tr>
<td>I Vs Q</td>
<td>0.69</td>
<td>0.61</td>
<td>0.84</td>
</tr>
</tbody>
</table>

3.3. Speech rate

The difference in speech rate between imperatives, questions and statements showed no significant contrast between the three speaker groups \(F(2, 531) = 0.95, p > 0.05\) and between intonations \(F(2, 531) = 2.36, p > 0.05\). No significant interaction effect was noted for speaker group and intonation \(F(4, 531) = 1.07, p > 0.05\), indicating that the differences in speech rate between intonations were comparable across speaker groups.

4. Discussion

The results of the present study showed significant differences between the speakers with mild and moderate dysarthria and the control speakers for all F0 measures, between speakers with mild dysarthria and the control speakers for all intensity measures, but no differences were found between the three speaker groups for the differences in speech rate. Speakers with both mild and moderate dysarthria had significantly higher average F0 than the control speakers. This agreed with the previous reports on the F0 level of PD speakers [14, 16]. Additionally, speakers with dysarthria were also observed to have a narrower F0 range and a smaller F0 envelop than the control speakers. That is, the PD speakers demonstrated less F0 variation within a sentence than non-PD speakers. This was not surprising as monopitch is one of the most prominent characteristics of the speech problem associated with PD [3-16]. It was hypothesized to be related to the reduced laryngeal control in PD speakers [6] and to the reduced amplitude in motor planning as a result of basal ganglia dysfunction [17].

For the intensity measures, speakers with mild dysarthria had a higher average intensity than the control speakers, and speakers with moderate dysarthria showed a comparable average intensity level when compared with the control speakers. This was unexpected as reduced loudness in speech production, or “hypophonia”, in PD speech has been commonly reported in the literature [3, 16, 17]. Nevertheless, Ludlow and Bassich [18] claimed that only 42% of the PD speakers in their study were perceived as hypophonic. Adams [19] argued that it might not be conclusive to study the intensity level of PD speakers through formal speech task (e.g., reading task), as the speakers tend to compensate for their hypophonia. Instead, Leuschel and Docherty [11] suggested the use of intensity variability measures rather than overall measures, as variability measures tended to be more revealing in dysarthric speech. The analysis of intensity range and intensity envelop in this study showed that speakers with mild dysarthria had smaller intensity variability than the control speakers, but it revealed no statistically significant difference between speakers with moderate dysarthria and the control speakers. Although the speech characteristics of individuals with dysarthria were argued to be highly heterogeneous [11], it is intriguing that speakers with...
moderate dysarthria showed speech performance more similar to the control speakers than speakers with mild dysarthria. As PD speakers’ ability to compensate for speech loudness could have played a role in the data of this study, further study to explore the speech loudness of PD speakers in different speech task might reveal interesting findings how intensity is affected in PD speakers of different severities.

This study also showed no significant difference between the PD speakers and the control speakers in how they vary their speech rate across the three intonations. Although the speech rate of PD speakers have been found to be abnormally fast [3], abnormally slow [16] and same as non-dysarthric speakers [2], Logemann and colleagues [2] reported that only 20% of the PD speakers showed abnormal speech rate. A review of the speech rate of the PD speakers in the present study showed that most speakers had relatively normal speech rate in comparison to the control speakers. This explained the current results, that the PD speakers were able to use duration to mark the intonation contrast between imperatives, questions and statements in the same manner as the control speakers.

The comparison of the acoustic parameters between the three intonations across the three speaker groups showed that PD speakers were able to generate intonation contrast between imperatives, questions and statements in the same manner as control speakers, even though the PD speakers showed significantly reduced F0 variability than control speakers. This finding agreed with Le Dorze and colleagues [7] that the PD speakers retained the cognitive competence of intonation contrast, and that their prosodic deficit in speech production is related to the deficit in motoric skills as a result of the neurological condition.

5. Conclusions

The current study explored the prosodic characteristics of PD speakers in the marking of intonations. The results showed that PD speakers differed from the control speakers with higher average F0 and reduced F0 variability within the sentence. Reduced intensity variability was also noted in speakers with mild dysarthria. However, the PD speakers also showed that they were able to mark the differences between intonations in similar ways as the control speakers, suggesting that dysprosody is mainly associated with motoric deficit.

6. Acknowledgements

The authors would like to thank all the subjects for their participation. We would also like to thank Prof. Dr. Alexander Storch and Dr. Christine Schneider for their help in subject recruitment, Ms Dorothea Senf, Ms Friederike Röhle and Ms Susanne Wächter for their participation in the perceptual evaluation, Dr. Matthias Eichner and Mr. Guntram Strecha for experimental setup. This project was funded by Alexander von Humboldt Foundation.

7. References