Acoustic characteristics of vowel sounds in patients with Parkinson disease

Young-Im Banga, Kyunghoon Minb, Young H. Solnh and Sung-Rae Choa, f, g, h, *
aDepartment and Research Institute of Rehabilitation Medicine, Yonsei University College of Medicine, Seoul, Korea
bSpeech-Language Clinic, Rehabilitation Hospital, Yonsei University Medical Center, Seoul, Korea
cDepartment of Rehabilitation Medicine, CHA Bundang Medical Center, CHA University, Korea
dDepartment of Neurology and Brain Research Institute, Yonsei University College of Medicine, Seoul, Korea
eBrain Korea 21 for Medical Science, Yonsei University, Seoul, Korea
fNano Science and Technology, Yonsei University, Seoul, Korea
gYonsei Biomedical Research Institute, Avison Biomedical Research Center, Seoul, Korea
hRehabilitation Institute of Neuromuscular Disease, Yonsei University College of Medicine, Seoul, Korea

Abstract: The purpose of this study was to define the acoustic voice and speech characteristics of patients with Parkinson disease (PD). Seven female patients with PD and seven female healthy controls participated in this study. Each subject was instructed to vocalize extended corner vowels (/a/, /e/, /i/, /u/) three times for at least 5 seconds at a comfortable voice loudness and tone. The voice was analyzed using the Praat program. As a result, female patients with PD showed a significant increase in jitter and noise-to-harmonics ratio (NHR). In addition, F1 and F2 among the PD patients demonstrated asymmetric centralization of unrounded vowels (/a/, /e/, /i/) in high/low/front/back positions of the tongue, consequently leading to a significant decrease in vowel space area, compared to healthy controls. This study showed the acoustic characteristics of vowel sounds not only by laryngeal variables such as abnormal jitter and NHR, but also by articulatory variables such as asymmetric centralization and reduced vowel space area in female patients with PD. Therefore, it is important to use these objective and sensitive variables to evaluate the status or severity of hypokinetic dysarthria in patients with PD.

Keywords: Parkinson disease, acoustic analysis, vowel space area, centralization

1. Introduction

Parkinson disease (PD) shows not only the cardinal symptoms of tremor, rigidity, and bradykinesia, but also the features of hypokinetic dysarthria [1]. Generally, deficits in speech and voice are the first common symptoms of neurodegenerative diseases such as PD [2, 3]. Hypokinetic dysarthria occurs in the majority of patients with PD, and the recognition thereof serves as confirmatory evidence of motor speech disorders, although only 3–4% of these patients receive speech and voice treatment [4]. Hypokinetic dysarthria is typically characterized by deficits in respiration, phonation, articulation, and prosody, caused by monopitch, monoloudness, underarticulation, and harsh voice [5, 6]. The combination of these deficits can lead to a loss of speech intelligibility [7–9]. Unfortunately, speech and voice characteristics in PD have not proven to be significantly or were only minimally improved by pharmacological or neurosurgical approaches [4, 10–13]. Such findings may represent that assessment of more sensitive acoustic variables is required to differentiate changes in patients with PD.

Auditory-perceptual evaluation has been commonly used to evaluate the quality of voice and the severity of voice and speech disorders. Furthermore, Darley...
et al. reported on auditory-perceptual evaluation for various types of dysarthria [6]. This method is convenient in terms of expense, time, and complexity of the procedure; however, it entails a limitation of inter-observer variability and relies on subjective judgment [13]. Fortunately, voice characteristics can be objectively analyzed with numeral values using an acoustic analyzer. Acoustic analysis programs are able to generate data related to speech disorders for the vocal cords easily and non-invasively in a relatively short time, as well as in a sensitive, objective, and quantitative manner [13]. Praat is one of such acoustic analysis software programs that not only can measure acoustic variables such as jitter, shimmer and noise-to-harmonic ratio (NHR), but also can express vocal functions as a separated spectrum of sound components [14].

The purpose of this study was to survey the sensitive acoustic characteristics of vowel sounds and to compare objective data obtained from patients with PD to those of normal healthy controls using the Praat program to overcome the shortcomings of auditory-perceptual evaluation in analyzing hypokinetic dysarthria associated with PD. We found that female patients with PD showed not only abnormal jitter and NHR, but also asymmetric centralization in high/low/front/back positions of the tongue, consequently leading to a significant decrease in vowel space area. The results demonstrated that reduced vowel space area can be a sensitive variable to assess the status or severity of hypokinetic dysarthria in patients with PD.

2. Subjects and methods

2.1. Subjects

In this study, seven female PD patients and seven female normal healthy controls were recruited. The normal controls consisted of female adults with no neurological disease or mental disease. PD was diagnosed by a specialized neurologist and the patients were treated with antiparkinsonian medications. All tasks were carried out during the “on” period of the medication effect. The baseline characteristics for the subjects concerning age, duration after diagnosis, mini-mental state examination (MMSE), functional independence measure (FIM), ambulatory function, as well as the Hoehn and Yahr scale are described in Table 1. The FIM comprises 18 items, a 7-level scale scored from 18 (total assistance) – 126 (complete independence), that assess the basic activities of daily living (ADL),

<table>
<thead>
<tr>
<th>Table 1 General characteristics of the subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
</tr>
<tr>
<td>No. of subjects</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
<tr>
<td>Duration (years)</td>
</tr>
<tr>
<td>MMSE (0–30)</td>
</tr>
<tr>
<td>FIM (18–126)</td>
</tr>
<tr>
<td>Ambulation (1–7)</td>
</tr>
<tr>
<td>Hoehn and Yahr scale (0–5)</td>
</tr>
</tbody>
</table>

Values are means ± standard deviation. MMSE: Mini-mental status examination, FIM: Functional independence measure.

2.2. Experimental procedure

Each subject was instructed to vocalize extended /a/, /e/, /i/, and /u/ vowels three times for at least 5 seconds at a comfortable voice loudness and tone. All vocalizations were recorded with a microphone (type: M-MS907, Sony Corporation, Tokyo, Japan) and a digital audio tape (Sony MZ-RH1–recorder: sampling rate 48,000 Hz: Sony Corporation, Tokyo, Japan) in a sound-proof room, and the distance of the microphone from the mouth was 5 cm. The vowel analysis of recorded aural signals was performed using the relatively stable vocalization period of the middle 1.5 seconds, excluding the initial and ending periods. Acoustic analysis was performed using Praat software version 4.1.2 (Boersume and Weenink, University of Amsterdam, The Netherlands) [15].

2.3. Acoustic variables

Jitter (%) was defined as the average variation in the frequency of adjacent cycles of the vocal note as a measure of pitch perturbation; shimmer (dB) as the average difference in the amplitude of consecutive periods of the vocal note as a measure of amplitude perturbation; and NHR as a ratio of the inharmonic (irregular and non-repetitive) sound wave components to the harmonic
(regular and repetitive) components. Vocal tract resonances comprised two formants. The first formant (F1) frequency was related to the constriction of the vocal tract and the capacity of the pharynx cavity. The second formant (F2) frequency was related to the length of the oral cavity. F2/F1 ratio was related to the listeners’ identification of vowels. In addition, the quadrilateral vowel space area was examined when the F1s and F2s for all vowels (a, e, i, u) were superimposed along the X- and Y-axes.

2.4. Statistical analysis

Statistical analysis was performed using the Statistical Package for Social Sciences (SPSS), Predictive Analytics Software version 18.0 for Windows. Each acoustic parameter of the vowel sounds was compared between patients with PD and normal controls using the Mann-Whitney U-Test. P-values <0.05 were considered statistically significant.

3. Results

3.1. Characteristics of subjects

The average age of the PD patients was 68.43 ± 5.74 years, and the average age of the normal controls was 61.29 ± 8.01 years. There was no statistically significant difference between the two groups. In the PD patients, the mean duration after diagnosis was 3.57 ± 1.90 years, and the mean score of MMSE was 23.29 ± 3.73, showing mild cognitive impairment. In addition, the mean Hoehn and Yahr scale score was 3.36 ± 1.11, reflecting mild to moderate disease with balance impairment (Table 1).

3.2. Jitter, shimmer, and noise-to-harmonics ratio (NHR)

When jitter was compared between the PD patients and the normal controls, the jitter values of the PD patients for all vowels (a, e, i, u) were significantly increased compared to those of the normal controls (p < 0.05) (Table 2). However, there was no statistically significant difference in the shimmer values between the two groups, although the shimmer values of the PD patients for all vowels (a, e, i, u) tended to be greater than those of the normal controls (0.26 ± 0.11, 0.34 ± 0.21). In regards to NHR, the NHR of the PD patients for only the u/ vowel (0.54 ± 0.69) was significantly greater than that of the normal controls (0.01 ± 0.15) (p < 0.05) (Table 2).

3.3. F1, F2, and F2/F1 ratio

When F1 was compared between the PD patients and the normal controls, the F1 of the PD patients for the /i/ vowel (605.97 ± 135.85) was significantly less than that of the normal controls (859.46 ± 47.91), whereas the F1 of the PD patients for the /u/ vowel (483.80 ± 118.41) was significantly greater than that of the normal controls (350.12 ± 40.00) (p < 0.05) (Table 3). For F2, the F2 of the PD patients for the /i/ vowel (2042.90 ± 189.40) was significantly less than that of the normal controls (2343.21 ± 304.71) (p < 0.05) (Table 3). On the other hand, the F2/F1 of the PD patients for the /i/ vowel (2.15 ± 0.43) was significantly greater than that of the normal controls (1.63 ± 0.07), while the F2/F1 of the PD patients for the /u/ vowel (4.40 ± 0.96) was significantly less than that of the normal controls (6.81 ± 1.50) (p < 0.05) (Table 3).

3.4. F1 and F2 ratios of each vowel

Examining F1 ratios for each vowel, the F1/i/F1a (0.83 ± 0.28), F1/e/F1e (0.84 ± 0.16), F1/u/F1u (0.73 ± 0.22), and F1/e/F1a (0.99 ± 0.26) ratios of the PD patients were significantly greater than that of the normal controls (0.41 ± 0.06; 0.57 ± 0.04; 0.52 ± 0.07; 0.70 ± 0.07) (p < 0.05), whereas the F1/i/F1 and F1/u/F1 ratios of the PD patients were not different from those of the controls (Table 4). This result

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Jitter, shimmer, and noise-to-harmonics ratio in PD patients and normal controls</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Normal</td>
</tr>
<tr>
<td>Jitter</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>0.30 ± 0.55</td>
</tr>
<tr>
<td>e</td>
<td>0.22 ± 0.76</td>
</tr>
<tr>
<td>i</td>
<td>0.30 ± 0.32</td>
</tr>
<tr>
<td>u</td>
<td>0.41 ± 0.30</td>
</tr>
<tr>
<td>Shimmer</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>0.26 ± 0.11</td>
</tr>
<tr>
<td>e</td>
<td>0.21 ± 0.06</td>
</tr>
<tr>
<td>i</td>
<td>0.31 ± 0.27</td>
</tr>
<tr>
<td>u</td>
<td>0.34 ± 0.21</td>
</tr>
<tr>
<td>NHR</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>0.01 ± 0.00</td>
</tr>
<tr>
<td>e</td>
<td>0.00 ± 0.00</td>
</tr>
<tr>
<td>i</td>
<td>0.01 ± 0.01</td>
</tr>
<tr>
<td>u</td>
<td>0.01 ± 0.15</td>
</tr>
</tbody>
</table>

Values are means ± standard deviation, *p<0.05 as compared with normal controls.
demonstrated that the ratios of the F1s for the unrounded vowels of /i/, /e/, and /a/ were significantly different between the PD patients and the normal controls, suggesting asymmetric centralization of the vowels. However, when the F2 ratios for each vowel were examined between the two groups, none of F2 ratios for the PD patients was significantly different from those of the normal controls (p > 0.05) (Table 4).

### 3.5. Vowel space area

Quadrilateral vowel space area was compared between PD patients and normal controls by superimposing the F1s and F2s for all vowels (/a/, /e/, /i/, /u/) on X- and Y-axes. As a result, a pattern of asymmetric centralization of the vowels was characteristically demonstrated in PD patients with hypokinetic dysarthria (Fig. 1). Consequently, the vowel space area of the PD patients (119954.30 ± 107323.01) was significantly less than that of the normal controls (312856.59 ± 143937.33) (p < 0.05) (Fig. 2).

### 4. Discussion

In this study, the Praat software program was used to quantitatively and objectively assess acoustic characteristics including jitter, shimmer, noise, and vibration measurements in PD patients [12–15]. As these values can be measured in a relatively short time, they are quite useful for voice and speech analysis of PD patients whose vocalization duration time is very short. This study recruited only female subjects to exclude gender-specific differences in the acoustic study. Moreover, the four vowels of /a/, /e/, /i/, and /u/ were used in this study.
tongue as well as labial and tongue-tip dysfunction. Therefore, the vowel characteristics reflected with the movement of the utterance organ in various positions can be better inferred from this study than previous research that studied only one vowel [14, 15]. Abnormal jitter, shimmer, and NHR are not always found in hypokinetic dysarthria. Kent et al. reported that the female PD patients differed from controls in shimmer, but they found no differences between male PD patients and controls [16, 17]. In this study, jitter and NHR were significantly greater in female PD patients than in normal controls. Shimmer also tended to be greater in PD patients than in normal controls. The present findings of abnormal values for frequency and amplitude perturbation as well as NHR in PD patients are consistent with those of a previous report [18]. Therefore, jitter, shimmer, and NHR proved to be objective variables for which to assess the status or severity of speech disorders and to measure the quality of a harsh or hoarse voice in PD patients whose laryngeal control is insufficient and unstable.

Articulatory undershoot is defined as a reduced range of articulatory movements and is an important symptom in most types of dysarthria. This undershoot usually results in vowel formant centralization represented by a reduced vowel space area, which has been reported in several studies on dysarthria [17–20]. Expansion of vowel space area following natural recovery or voice treatment has also been documented [21]. Capacity changes are known to occur depending on the position of the tongue, where F1 and F2 change systematically according to the high/low/front/back positions of the tongue for each vowel. This study drew quadrilateral space areas with the vertexes of the corner vowels /a/, /e/, /i/, and /u/ using F2 and F1, respectively, on X- and Y-axes. By doing so, hypokinetic dysarthria can be interpreted by calculating the size of the vowel space area. Through comparison of the vowel space areas, this study was able to assess the impact of changes in tongue movement on vowel production. The unrounded vowels /a/, /e/, and /i/ moved in the central direction, but rounded vowel /u/ showed a slightly upward movement. F2 frequencies for the PD patients were lower than that of the normal controls, demonstrating that the tongue was moving backwards toward the mouth floor for all four vowels. Moreover, a previous study described that PD patients showed back tongue as well as labial and tongue-tip dysfunction whereas most previous researchers analyzed /a/ vocalization alone. The /a/ is a low back unrounded vowel; the /e/ is a low front unrounded vowel; the /i/ is a high front unrounded vowel; and the /u/ is a high back rounded vowel. Therefore, the vowel characteristics reflected with the movement of the utterance organ in various positions can be better inferred from this study than previous research that studied only one vowel [14, 15].

In addition, F2/F1 ratios for the PD patients significantly deviated from those in normal controls in this study. The F1 ratios for each vowel, F1i/F1a, F1i/F1e, F1u/F1a, and F1u/F1i, in the PD patients were significantly greater than that of the normal controls, whereas the F2 ratio for each vowel was not. Although this study demonstrated that the ratios of F1 for the unrounded vowels of /i/, /e/, and /a/ were significantly different between the PD patients and normal controls, a previous study showed that the F2 ratio F2i/F2u was the only acoustic index of vowel articulation that significantly differed between hypokinetic dysarthria and normal speech [13]. This suggests that PD patients with hypokinetic dysarthria have articulatory disorders derived from deficits or weaknesses of high/front and low/back tongue movements and asymmetric centralization of vowel formants, consequently affecting a listener’s perception of a PD patient’s vowels [8, 17]. Taken together, F1 and F2 of PD patients exhibit an asymmetric central position and a reduced vowel space area that lead to a loss of speech intelligibility. This study also assessed acoustic dysfunction as laryngeal and articulatory variables. Consequently, we found a co-occurrence of laryngeal dysfunction and articulatory dysfunction in female PD patients.

5. Conclusion

This study described the acoustic characteristics of vowel sounds not only by laryngeal variables such as abnormal jitter and NHR, but also by articulatory variables such as asymmetric centralization and reduced vowel space area in female patients with PD, who were shown to exhibit deficits of laryngeal function and tongue movement. Acoustic analysis using Praat software was used to conduct objective and sensitive assessments for which to evaluate the status or severity of hypokinetic dysarthria in patients with PD. Our results provide new insights into the necessity of such assessment for describing and monitoring changes in acoustic characteristics following therapeutic intervention such as the Lee Silverman Voice Treatment (LSVT) [21].
Acknowledgment

This study was supported by grants from the National Research Foundation (NRF-2010-0020408; 2010-0024334) funded by the Ministry of Education, Science and Technology, Republic of Korea.

Conflicts of interest

There are no conflicts of interest regarding this research.

References