The Comparison of Cepstral Peak Prominence by Using Speech Tool and Praat for Voice Analysis of Patients With Vocal Cord Palsy

Jisu Kim, Min Young Cho, Dong Young Kim, Woongsang Sunwoo, Woori Park, and Joo Hyun Woo

Department of Otorhinolaryngology-Head and Neck Surgery, Gil Medical Center, Gachon University College of Medicine, Incheon, Korea

Background and Objectives This study aimed to evaluate the voice of patients with vocal cord palsy using the Praat compared to Speech tool.

Materials and Method The medical record of the patients with vocal cord palsy from 2013 to 2021 was analyzed retrospectively to validate Praat as a voice evaluation modality compared to the speech tool. Total 60 patients were enrolled in this study. Thirty control and 30 vocal cord palsy patients were selected to undergo recording of voice samples. The voice samples, /a/ and “Sancheck” were evaluated both groups and cepstral peak prominence was analyzed with both modalities, Praat and speech tool.

Results Statistically significant differences were observed between the control and vocal cord palsy groups in the speech tool and Praat. There was also a significant difference between pre- and post-treatment values in the vocal cord palsy group. A similar change in the voice value was observed using both methods. The Praat showed a lower value in 1st visit of patients with vocal cord palsy in the vowel test. The Praat vowel test may sensitively represent voice problems in patients with vocal cord palsy. This could contribute to decision-making regarding the treatment modality for vocal cord palsy.

Conclusion Praat is open-access software that is freely available. It can be easily and sufficiently used for voice evaluation in patients with vocal cord palsy.

Keywords Vocal cord palsy; Speech tool; Praat; Voice evaluation.

INTRODUCTION

Patients with vocal cord palsy experience multiple difficulties, one of which is voice problems: breathy, weak, and low-pitched dysphonia.

Over the years of research on voice and speech, cepstral peak prominence (CPP) has become an objective measure of breathiness and dysphonia. Guidance from the American Speech-Language-Hearing Association recommends CPP as a tool for measuring the overall level of noise in the vocal signal and as a general measure of dysphonia [1].

Voice problems can be measured and quantified using different software programs, such as Analysis of Dysphonia in Speech and Voice, speech tools, and Praat [2].

As described by Maryn et al. [3], CPP is an acoustic measure of voice quality that has been qualified as the most promising and perhaps robust acoustic measure of dysphonia.
severity. Speech tool’s CPP value has been used in many studies to evaluate patient’s voice [4-8].

Different CPP measurement software programs provide different CPP values. In addition, software programs differ in terms of cost. Praat is a program that is free and open to society, which features different types of speech analysis including cepstral, pitch, formant, jitter, shimmer and many more. Praat can be a beneficial tool for the local outpatient clinic to use as an indicator of whether the patient’s voice has improved after a certain treatment.

Therefore, the purpose of this study is to compared the CPP values of speech tools and Praat voice analysis in patients with vocal cord palsy to determine whether there is a correlation between the CPP values of the two software programs and whether Praat’s CPP value could be beneficial in measuring the degree of voice problems at local outpatient clinics.

**METHODS**

**Study design and population**

This study was a retrospective analysis based on data selected from a database of electronic medical records that recorded voices in the voice laboratory. The database for this study consisted of 60 patients who visited the otolaryngology head and neck surgery outpatient clinic between 2013 and 2021. All patients were first examined by an otolaryngologist using a 70° rigid endoscope (Stryker model FocESS Sinuscope, 70°) to diagnose abnormalities in the larynx. Thirty patients with vocal cord palsy were included in the study. None of the patients were diagnosed with bilateral vocal cord palsy. Patients were treated with any one of the following: injection laryngoplasty, arytenoid adduction or voice therapy. Patients who underwent treatment had to revisit the outpatient clinic for follow-up voice recordings. For comparison, a voice evaluation of 30 healthy speakers was performed. These 30 healthy speakers were volunteers who were recruited through advertisement at the outpatient clinic. They were examined thoroughly with 70° rigid endoscope and voice recordings were made as vocal cord palsy patients did. Thirty control (healthy speakers), 30 pre-treatment vocal cord palsy patients’, 30 post-treatment vocal cord palsy patients’ voice samples were analyzed using voice analyzing tools made by Hillenbrand and Gayvert [9] (Speech tool version 1.65) and Boersma and Weenink [10] (Praat version 6.2.01). Both tools analyzed the recorded voice samples of the patients and calculated their CPP.

**Procedure**

After the evaluation of the larynx, the patients were recorded in a quiet and comfortable environment in the outpatient clinic using a microphone (SM48 Cardioid dynamic vocal microphone, Shure Asia, Island East, Hong Kong). The microphone was placed approximately 10 cm in front of each participant’s mouth. Two different sample types were used for each patient. The first sample made a sustained /a/ vowel sound at a comfortable pitch and loudness for approximately 3–4 seconds and the second sample read the Korean passage “Sancheck.” This sentence is considered well-balanced, phonetically and phonologically [8].

The voice samples, /a/ and “Sancheck” were evaluated both groups; once in control group and twice in vocal cord palsy group for comparing before and after treatment.

**Statistical analysis**

Statistical analyses were performed using SPSS statistics (version 22.0; IBM Corp., Armonk, NY, USA). p<0.05 was considered statistically significant.

The study was conducted in accordance with the guidelines of the Declaration of Helsinki and approved by the Institutional Review Board of Gachon University (Approval No. GDIRB 2022-114).

**RESULTS**

Sixty patients were enrolled in the study. The vocal cord palsy group was older than the control group (Table 1). There were various reasons for vocal cord palsy, the most common of which was postoperative complications of thyroid surgery being 17 patients (56.7%). Other causes of the vocal cord palsy included, post-concurrent chemoradiation therapy complication, cerebral infarct, viral infection and others were unknown. All the patients with vocal cord palsy were treated to alleviate their voice difficulties. Most of the patients underwent surgical procedures; 15 patients (50%) were treated with injection laryngoplasty, 13 patients (43.3%) were treated with injection laryngoplasty and arytenoid adduction, one patient was treated with laser cordec-

<table>
<thead>
<tr>
<th>Table 1. Subject characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total (n=60)</td>
</tr>
<tr>
<td>Sex</td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td>Age (years)</td>
</tr>
</tbody>
</table>

Data are presented as n (%) or mean±standard deviation.
Comparison of CPP From Different Tools
Jisu Kim, et al.

Table 2. Mean cepstral peak prominence values of control, pre-treatment, post-treatment group (p<0.001)

<table>
<thead>
<tr>
<th></th>
<th>CG (n=30)</th>
<th>PTG (n=30)</th>
<th>PoTG (n=30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>STV</td>
<td>21.879±2.169</td>
<td>11.219±2.612</td>
<td>15.321±3.761</td>
</tr>
<tr>
<td>STS</td>
<td>13.587±1.909</td>
<td>10.354±1.357</td>
<td>11.782±1.186</td>
</tr>
<tr>
<td>PS</td>
<td>11.247±1.301</td>
<td>6.646±1.495</td>
<td>8.959±1.591</td>
</tr>
</tbody>
</table>

CG, control group; STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; PTG, pre-treatment group; PoTG, post-treatment group.

Table 3. Pre- and post-treatment value of ST and Praat in vocal cord palsy patients

<table>
<thead>
<tr>
<th></th>
<th>Injection laryngoplasty (n=15)</th>
<th>Injection laryngoplasty and arytenoid adduction (n=13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-</td>
<td>Post-</td>
</tr>
<tr>
<td>STS</td>
<td>10.63±1.330</td>
<td>12.36±1.065</td>
</tr>
<tr>
<td>PS</td>
<td>6.87±1.423</td>
<td>9.48±1.240</td>
</tr>
</tbody>
</table>

ST, speech tool; STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence.

Fig. 1. Mean cepstral peak prominence values of control, PTG and PoTG. CG, control group; STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; PTG, pre-treatment group; PoTG, post-treatment group.

Fig. 2. Mean cepstral peak prominence values of control, PTG and PoTG. CG, control group; STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; PTG, pre-treatment group; PoTG, post-treatment group.

Vocal cord paralysis causes serious phonation, respiration, and psychological problems that degrade quality of life. Vocal cord palsy occurs due of dysfunction of the recurrent laryngeal or vagus nerve, which leads to dysphonia. Patients with vocal cord palsy usually present with a fairly sudden onset of breathiness and weak, low-pitched dysphonia [2].

In clinical settings, various assessments have been performed to determine the severity of voice disorders. Voice evaluation is a widely used method that can easily quantify the quality of a patient’s voice disorders [11,12].

Auditory perceptual evaluation using the Grades of the Dysphonia Rating Scale (GRBAS), grade, roughness, breathiness, asthenia, strain, score, and consensus auditory perceptual evaluation of voice (cape-V) have been reliable measurements of each voice in many countries. In addition, acoustic and aerodynamic evaluations are relevant measurements in unilateral vocal cord palsy, which shows worse jitter, shimmer, noise/harmonic ratio, and maximum phonation time compared with normal voice [13,14].

Maryn et al. [3] described that CPP is one of the methods to measure voice quality, which has been qualified as the most promising and perhaps robust acoustic measure of breathy voice. The CPP value is a spectral representation of the voice spectrum, which is calculated by using the Fourier transformation of the voice spectrum. As CPP values are calculated by a double Fourier transformation and not by pitch tracking algorithms, as they do with noise-to-harmonic ratios, jitter, and shimmer values, they remain a robust measure for severely dysphonic voices [15-17].

As shown in this study, a normal voice without any diseases would have a well-defined harmonic structure with a strong cepstral peak compared to a breathy and hoarse voice, which is a poorly defined harmonic structure that leads to a weak cepstral peak with medial arytenoidectomy and one patient was treated with voice therapy.

Speech tool vowels (STV) and speech tool sentences (STS) were compared with Praat vowels (PV) and sentences (PS) in the control, pre-treatment, and post-treatment groups (Table 2, Supplementary Figs. 1, 2 and 3 [in the online-only Data Supplement]). The STV and STS values were higher in the control group than in the pre-treatment vocal cord palsy group. After treatment, the patients showed better STV and STS scores. Similar results were observed in the PV and PS groups, with the highest scores observed in the control group and better scores after treatment in the vocal cord palsy group.

Vowel tests of both software programs showed more differences between the three groups than the sentence test (Fig. 1).

Most patients with vocal cord palsy undergo surgical procedures to improve their voice. Injection laryngoplasty was performed in almost all the patients. Patients who had unsatisfactory outcome after injection laryngoplasty underwent additional surgery such as arytenoid adduction. Patients whom underwent arytenoid adduction was chosen not based on Praat CPP scores but by their clinical evaluation and patients’ needs. The result showed lower ST and Praat values in the pre-treatment group than those in the injection laryngoplasty-only group. Better voice status was observed after treatment with ST and Praat. The PV presented the lowest value in the pre-treatment of patients with injection laryngoplasty and arytenoid adduction. Both software programs were statistically significant in representing voice improvement (Table 3).
tral peak. Similar to our study, Balasubramanium et al. designed a study on CPP in patients with unilateral vocal fold palsy and their results revealed lower CPP values in the clinical group than in the control group [2,18].

The results of this study showed that both indices have a high level of validity in discriminating between different GRBAS [19,20].

The CPP values in Praat were consistently significantly lower than those of the speech tool across all the data pools included in the analysis, which shows that there is a correlation between the CPP values achieved by the Praat and speech tools. The PV test reflects a more serious voice problem, which could contribute to determining the type of treatment.

Although the Multidimensional Voice Program is still the tool of choice worldwide by most researchers in Vocology, with much research done with Praat, its use will soon become more valuable [21].

With more studies on Praat being performed in the field of vocology and phonetics, this software program will be a valuable tool for clinicians to use as an indicator of voice quality.

The use of Praat has several advantages to using Praat. First, the program is free for anyone in need. Second, because it is software that can be downloaded from the internet, there is no need for additional space or equipment other than computers. Third, the pitch analysis algorithm is the most accurate in the world, as mentioned by Boersma and Weenink [10] which is the only linguistically oriented learning algorithm that can handle dynamic length changes (ejectives), non-glottal myo-elastics (trills), and sucking effects (clicks and implosives); the gradual learning algorithm is the only linguistically oriented learning algorithm that can handle free variation. Forth, there is a possibility to use to decide the treatment modality to vocal cord palsy patients.

This study has several limitations. The first limitation is the small sample size. In only 30 patients with vocal cord palsy, CPP values did not represent all patients with vocal cord palsy. Second, the results of the study show a correlation between the two software programs, yet there is no certain numerical index that indicates whether a patient has a normal voice or a voice disorder. Therefore, further studies are needed to support and provide indices to validate people’s voice problems.

In conclusion, praat is an open-access software freely available and sufficiently used for voice evaluation in patients with vocal cord palsy. The data of this study supports that the values obtained by Praat and ST are positively correlated which can lead to further studies reflecting the severity of vocal cord palsy and could contribute to decide the type of treatment.

**Supplementary Materials**

The online-only Data Supplement is available with this article at https://doi.org/10.22469/jkslp.2023.34.2.45.

**Acknowledgments**

None

**Conflicts of Interest**

The authors have no financial conflicts of interest.

**Authors’ Contribution**


**REFERENCES**

Supplementary Fig. 1. Mean cepstral peak prominence values of control, pre-treatment, post-treatment group. CG, control group; STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; CPP, cepstral peak prominence.
Supplementary Fig. 2. Pre-post injection laryngoplasty treatment CPP value comparison. STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; CPP, cepstral peak prominence.
Supplementary Fig. 3. Pre-post injection laryngoplasty & arytenoid adduction treatment CPP value comparison. STV, speech tool vowel; PV, Praat vowel; STS, speech tool sentence; PS, Praat sentence; CPP, cepstral peak prominence.