<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>The evaluation of dysphonia for Cantonese school-age children: a multiparametric approach</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Lam, Ngai-yan; 林艾茵</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>Lam, N. [林艾茵]. (2013). The evaluation of dysphonia for Cantonese school-age children: a multiparametric approach. (Thesis). University of Hong Kong, Pokfulam, Hong Kong SAR.</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2013</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/238551">http://hdl.handle.net/10722/238551</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>The author retains all proprietary rights, (such as patent rights) and the right to use in future works.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.</td>
</tr>
</tbody>
</table>
The Evaluation of Dysphonia for Cantonese School-age Children:

A Multiparametric Approach

Lam Ngai Yan, Yann

A dissertation submitted in partial fulfillment of the requirements for the Bachelor of Science
(Speech and Hearing Sciences), The University of Hong Kong, June 30, 2013.
The Evaluation of Dysphonia for Cantonese School-age Children: A Multiparametric Approach

Abstract

The present study investigated the feasibility of using multiparametric approach to predict the presence of dysphonia in Cantonese school-age children. A total of 30 children (16 dysphonic and 14 control) aged 6;02 to 12;07 participated in the study. The voice of each child was recorded and evaluated acoustically, aerodynamically, and using voice range profiles. A minimal set of instrumental voice parameters was identified based on the predictive power of each parameter in discriminating dysphonic and normal voices. Results showed that the two groups performed significantly differently in six parameters, including relative average perturbation, shimmer percent, maximum phonation time, mean airflow rate, intensity range and profile area in voice range profile. The combined use of shimmer percent and maximum phonation time correctly predicted 93.33% of the voices. These results agreed with many previous studies, indicating the value of clinical use and further investigation of multiparametric approach in evaluating voice quality in children.

Keywords: pediatric voice, perceptual evaluation, acoustic parameters, aerodynamic parameters, multi-dimensional evaluation
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

Introduction

A voice disorder is present when the quality of sound produced is disturbed due to vocal fold damage or some compensatory involvement of other vocal tract structures (Sapienza & Hoffman-Ruddy, 2009). The presence of voice problems in children is common, with a reported prevalence ranged from 2% to 38% (Lee, Stemple, Glaze, & Kelchner, 2004). In Hong Kong, 7.4% (33 out of 445) of children were reported to have a voice problem either currently or previously (Ma & Mo, in print). Reportedly, phonotrauma was one of the most common causes of dysphonia and vocal fold nodule was one of the most commonly seen pathologies in children (Andrews, 2002). Children with dysphonia may experience not only limitations on performing vocal activities but also negative perceptions from listeners towards their non-speech characteristics, namely, less intelligent, less confident, and less attractive (Ma & Yu, in print). As children grow older, they become intellectually mature enough to be aware of the negative impacts brought by dysphonia and hence their voice problems no longer only influence them physically but also socially and emotionally. According to Connor, Cohen, Theis, Thibeault, Heatley, and Bless (2008), dysphonic children aged 5 to 7 years have already started to be aware that they have a different voice when compared to their peers and started to feel embarrassment; dysphonic adolescents aged 13 to 18 years feel frustrated and angry because of their deviant voice quality. Therefore, close attention should be warranted to children with voice problems in order to prevent the
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

worsening of their dysphonia, to minimize the adverse impacts brought by dysphonia to their quality of life, and to prevent these impacts from further influencing the children as they grow.

In order to identify children with potential voice problems for proper management, a comprehensive evaluation of voice quality is crucial. To be comprehensive, the structures, functions, and activities and participations of voice production under The World Health Organization’s International Classification of Functioning, Disability, and Health framework (World Health Organization, 2001) should be evaluated. Laryngoscopic examinations like endoscopy examine for the structural abnormalities for voice production (Ma, Yiu, & Verdolini-Abbott, 2007); perceptual voice rating and instrumental voice analyses evaluate the functions of voice; and self-reported questionnaires like Pediatric Voice Handicap Index (Zur, Cotton, Kelchner, Baker, Weinrich, & Lee, 2007) evaluate the impacts of voice problems on activities and participations.

Among these aspects, the evaluation of the functions of voice draws major attention due to the controversy of the use of perceptual rating versus instrumental voice analysis. Perceptual evaluation has always been considered as the “gold standard” because voice quality is essentially perceptual in nature and the need of clinical attention is always based on the perceptual acceptability of the quality of sound (Kreiman, Gerratt, Kempster, Erma, & Berke, 1993). However, perceptual rating is subjective. In other words, it depends greatly
on sources of variability including raters’ internal standards, the types of rating scale (Kreiman et al., 1993), the provision of anchors using standard voice references (Chan & Yiu, 2002), and the types and qualities of the voice samples (Gerratt, Kreiman, Antonanzas-Barroso, & Berke, 1993). These variables increase the variability of the inter- / intra-rater and test-retest agreement and reliability and hence make the rating less consistent and less meaningful. Therefore, the complementary use of objective instrumental evaluations is crucial for more reliable assessment for the function of voice.

Major instrumental measures evaluating voice qualities are voice range performances, aerodynamic, and acoustic analyses (Sapienza & Hoffman-Ruddy, 2009). Van den Berg (1958)’s myoelastic-aerodynamic theory paved the foundation of how the mechanism and efficiency of voice production can be quantified by instrumental analyses. During phonation, or voicing, the pair of elastic vocal folds is set into vibrations when air stream is pushed from the lung to the oral cavity through the glottal region (Seikel, King, & Drumright, 2010). In this process, the efficiency of one’s vocal folds in converting the aerodynamic power into acoustic energy (Hirano, 1989) leads to various quantifiable characteristics. The characteristics measured in voice range performances mainly include frequency range and intensity range. Those measured in aerodynamic analysis mainly include intraoral pressure, maximum phonation time and rate of airflow. Those measured in acoustic analysis mainly include fundamental frequency, noise-to-harmonic ratios and perturbations. Therefore, by
looking at the numerical values in these analyses, the function of voice can be evaluated objectively. If parameters in these analyses can give good prediction on perceptual rating, their sole use could then objectively reflect the perceptual acceptability of the quality of sound for clinical decision making. However, there has been no single instrumental parameter consistently showing strong correlation with perceptual judgment (Ma & Yiu 2006). This inconclusiveness can be attributed to the large variability of most aerodynamic and acoustic characteristics (Wuyts et al., 2000).

Some researchers hence proposed the use of a combination of instrumental parameters to predict perceptual judgment. They identified combinations of 3 to 7 parameters to achieve concordances with perceptual judgments ranging from 49.9% to 88.0% (Cantarella, Baracca, Pignataro, & Forti, 2001; Giovanni et al., 1996; Ma & Yiu, 2006; Wuyts et al., 2000; Yu & Giovanni, 2003; Yu, Ouaknine, Revis, & Giovanni, 2001; Yu, Revis, Wuyts, Zanaret, & Giovanni, 2002; Yu, Wang, Han, Yang, & Han, 2004). Table 1 is a summary of studies which report the combined use of instrumental parameters to predict perceptual judgment.

Certain studies reported here achieved high predictability of perceptual judgment, which is very encouraging. However, these data were all obtained from the adult population and non-Cantonese population. Only the study by Ma and Yiu (2006) was done in the Cantonese population. These data cannot be applied in Cantonese children due to the anatomical differences of the respiratory and phonation systems between adults and children,
and the cultural and linguistic differences among children in various language regions (Sederholm, McAllister, Sundberg, & Dalkvist, 1992). Therefore, the current study aimed at identifying a minimal set of instrumental parameters to predict the presence of dysphonia in Cantonese school-age children with the following hypotheses: First, there are significant differences in instrumental parameters between children with dysphonia and children with perceptually normal voices; second, pathological voices and normal voices in children can be differentiated using a minimal set of instrumental parameters.

Table 1. Percentage of concordances between perceptual ratings and instrumental parameters reported in the literatures.

<table>
<thead>
<tr>
<th>References</th>
<th>Overall % of concordance</th>
<th>Instrumental parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cantarella et al. (2011)</td>
<td>82.6</td>
<td>Shimmer, glottal efficiency index, maximum phonation time (MPT)</td>
</tr>
<tr>
<td>Giovanni et al. (1996)</td>
<td>66.1</td>
<td>Jitter, glottal leakage, number of harmonics, duration of attack period</td>
</tr>
<tr>
<td>Ma &amp; Yiu (2006)</td>
<td>67.3</td>
<td>Jitter, voice range profile area, peak intraoral pressure, MPT</td>
</tr>
<tr>
<td>Wuys et al. (2000)</td>
<td>49.9</td>
<td>Jitter, highest frequency, lowest intensity, MPT</td>
</tr>
<tr>
<td>Yu &amp; Giovanni (2003)</td>
<td>84.0</td>
<td>Fundamental frequency, signal-to-noise ratio, Lyapunov coefficient, vocal range, estimated subglottic pressure, oral airflow, MPT</td>
</tr>
<tr>
<td>Yu et al. (2001)</td>
<td>86.0</td>
<td>Fundamental frequency, signal-to-noise ratio, Lyapunov coefficient, vocal range, estimated subglottic pressure, MPT</td>
</tr>
<tr>
<td>Yu et al. (2002)</td>
<td>88.0</td>
<td>Signal-to-noise ratio, Lyapunov coefficient, vocal range, estimated subglottic pressure, oral airflow, MPT</td>
</tr>
<tr>
<td>Yu et al. (2004)</td>
<td>79.8</td>
<td>Jitter, shimmer, normalized noise energy, harmonic-to-noise ratio, MPT</td>
</tr>
</tbody>
</table>

Method

Participants

Sixteen children with dysphonia (8 boys and 8 girls) and fourteen children with normal voices (6 boys and 8 girls) participated in this study. Dysphonic children were
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

recruited from Voice Research Laboratory at The University of Hong Kong and from two local primary schools in Hong Kong. They were judged to be perceptually dysphonic by two experienced speech therapists. They aged from 6;02 to 11;02 with a mean age of 8;07 (SD = 1.70) and were diagnosed by Ear, Nose, and Throat surgeons with various types of vocal fold pathologies as summarized in Table 2.

Table 2. The types of vocal fold pathologies in the dysphonic group.

<table>
<thead>
<tr>
<th>Vocal fold pathology</th>
<th>Number of Dysphonic Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vocal fold nodules</td>
<td>9</td>
</tr>
<tr>
<td>Vocal fold polyp</td>
<td>1</td>
</tr>
<tr>
<td>Posterior glottal chink</td>
<td>1</td>
</tr>
<tr>
<td>Miscellaneous</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>16</strong></td>
</tr>
</tbody>
</table>

The control group was recruited from the student and staff community. They aged from 6;08 to 12;07 with a mean age of 9;06 (SD = 2.00). The mean ages of the two groups were similar (p = .19). All participants, including the dysphonic children and the controls, were native Cantonese speakers, medically healthy, with no auditory problems, speech problems, language problems, history of intubation, severe respiratory and allergies problems.

Additionally, children in the control group had no history of voice disorders and were judged by their caretakers and two experienced speech therapists to have perceptually normal voice.

**Procedures**

The voice functioning of each participant was evaluated using acoustic perturbation analysis, aerodynamics evaluations and voice range profiles. The voice recording sessions for these evaluations were carried out at Voice Research Laboratory at The University of
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

Hong Kong. The recording environment was kept quiet with an ambient noise level less than 40dB, measured by a professional grade sound level meter (Rion NL20). All participants seated upright on a straight-back chair during tasks. Full demonstration was given before the start of each task. Verbal cues and reinforcements were given in supplementary during the tasks to guide and encourage the participants. All verbal inputs and stimuli were in Cantonese. Recordings for acoustic perturbation analysis, aerodynamic analysis and voice range profiles were randomized to counterbalance possible ordering effects of recording sequence on participants’ voice (Shaughnessy & Zechmeister, 2003).

Voice recording for acoustic perturbation analysis.

Voice samples were recorded directly into the Computerized Speech Lab Multidimensional Voice Program (MDVP, KayPENTAX) using a professional grade microphone (Shure Beta 58A). The microphone was placed 10 cm from the participant’s mouth to avoid air burst. Participants were asked to read aloud a sentence /ba₁ ba₁ da₂ ɡɔ₁ ɡɔ₁/ (meaning ‘father hits the elder brother’) for five times at their most comfortable loudness and pitch level. Reading aloud of sentence instead of sustaining phonation of vowels was chosen as the task because samples at discourse level are more representative in reflecting a person’s daily speaking voice (Bele, 2005).

In acoustic perturbation analysis, voice samples used were the middle three trials of the whole sentence /ba₁ ba₁ da₂ ɡɔ₁ ɡɔ₁/. The values of fundamental frequency (F₀) (Hz),
relative average perturbation (RAP) (%), shimmer percent (Shim) (%), and noise-to-harmonic ratio (NHR) were generated by MDVP for each trial. The averaged values for each parameter of the sentence were obtained.

**Voice recording and analysis for aerodynamic evaluations.**

Voice samples were recorded directly into and analyzed by the KayPENTAX Phonatory Aerodynamic System (PAS). KayPENTAX PAS was calibrated according to the manufacturer’s manual before the recording of every participant. Participants were then asked to secure a transducer-connected facemask over the nose and the mouth and then to complete four tasks at their most comfortable pitch and loudness level. The four tasks were maximum phonation of vowel /a/, comfortable phonation of vowel /a/, production of the consonant-vowel syllable string /ipipipipipi/, and production of the sentence /ba₁ ba₂ da₂ gɔ₁ gɔ₁/.

In the first task, participants were instructed to take their deepest breath and then sustain a phonation of vowel /a/ for a maximum duration. Full demonstration was given by the clinician before the participant started. Five trials were completed for each participant. Participants were encouraged verbally to sustain a longer phonation than the previous one after each trial. The maximum phonation times (MPT) (sec) were measured.

In the second task, participants were asked to sustain a phonation of vowel /a/ for six seconds and for five trials. The onset to the offset of the phonations of the middle three
trials were selected and analyzed. The mean airflow rates (l/sec) for the three trials were generated by KayPENTAX PAS and the averaged values were obtained.

In the following two tasks (the third and the fourth), a flexible rubber tube was connected to the transducer module in addition to the facemask. Participants were instructed to keep the tube on top of the middle of the tongue while securing the facemask over the nose and the mouth.

In the third task, participants had to produce the consonant-vowel syllable string /pipipipipipipi/ in one continuous breath with equal stress on each syllable for five trials. The most stable three successive consonant-vowel syllables /pi/ of the middle three trials were selected and analyzed. The peak intraoral pressures (cm H$_2$O) for the three trials were calculated and the averaged values were obtained.

In the fourth task, participants had to read aloud the sentence /ba1 ba1 da2 gɔ1 gɔ1/ for five times. The signals for the second /ba1/ of the middle three trials were selected and analyzed. The peak intraoral pressures (cm H$_2$O) for the three trials were calculated and the averaged values were obtained.

**Voice recording for voice range profile and speech range profile.**

Voice samples were recorded directly into Swell’s Phog 2.0 using a headset microphone (AKG c420) placed 5 cm away from the left corner of the mouth at a depression angle of 45°. Swell’s Phog 2.0 was calibrated according to the manufacturer’s manual.
before the recording of every participant. Participants performed pitch gliding as the warm up exercise before the recording started. To obtain the voice range profile, they were guided to say /a/ at the softest and loudest intensity across the participant’s maximum phonation frequency range. Hand signals were given to encourage the client to achieve maximum and minimum intensities and frequencies (Coleman, 1993). To obtain the speech range profile, they were asked to read aloud a monologue “Speaking is Not Easy” (Appendix) at their most comfortable loudness and pitch level. The voice range parameters obtained for each voice range profile included the highest and lowest frequencies (Hz), frequency range (semitone), the highest and lowest intensity levels (dB), intensity range (dB), and total area (dBST).

**Reliability measurements**

For the evaluation of the inter-rater agreement of the study, 20% of participants (three dysphonic and three control participants) were randomly selected and re-analyzed by another rater. Pearson correlation coefficients of the two sets of analyses were calculated.

**Statistical analyses**

In order to identify a minimal set of parameters to predict the presence of dysphonia, all data obtained in the evaluation of voice functioning (including acoustic perturbation analysis, aerodynamic evaluations, and voice range profiles) were analyzed in two steps. First, data of the control group and the dysphonic group in each parameter were compared using independent t test to identify in which parameters did the two groups differ from each
other significantly. As more than one statistical tests were done on some sets of data, Bonferroni adjustment was carried out to minimize potential type I error. Therefore, the alpha level for each evaluation was adjusted by dividing .05 by the number of statistical tests used for that particular set of data. The adjusted alpha levels were .013 (.05/4) for acoustic perturbation parameters, .05 (.05/1) for aerodynamic parameters, and .007 (.05/7) for parameters in voice range profiles. Second, among the parameters identified in Step 1, the most powerful predictors were selected using step-wise discriminant function analysis. The predicting power of these predictors in categorizing samples back into the control group and the dysphonic group was also evaluated and was expressed as a discriminant function.

Results

Voice functioning analyses

Performance of the control group and the dysphonic group in voice functioning analyses were compared using parametric statistics as Kolmogorov-Smirnov test showed that values of all parameters were normally distributed (p > 0.05).

Acoustic perturbation analysis.

The dysphonic group demonstrated significantly higher relative average perturbation (RAP) value (p < .001) and shimmer percent (Shim) value (p < .001) than the control group. Table 3 summarized the mean values of the acoustic perturbation parameters of the two groups.
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

Table 3. Mean and standard deviations of acoustic perturbation parameters of the control and dysphonic group.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Control group (n = 14)</th>
<th>Dysphonic group (n = 16)</th>
<th>Independent t Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>(SD)</td>
<td>Mean (SD)</td>
<td>t</td>
</tr>
<tr>
<td>Fundamental frequency (Hz)</td>
<td>263.20 (23.67)</td>
<td>241.00 (34.26)</td>
<td>2.03</td>
</tr>
<tr>
<td>Relative average perturbation (%)</td>
<td>0.92 (0.23)</td>
<td>1.68 (0.65)</td>
<td>-4.36</td>
</tr>
<tr>
<td>Shimmer percent (%)</td>
<td>5.14 (0.85)</td>
<td>8.46 (2.27)</td>
<td>-5.43</td>
</tr>
<tr>
<td>Noise-to-harmonic ratio</td>
<td>0.19 (0.032)</td>
<td>0.21 (0.030)</td>
<td>-1.80</td>
</tr>
</tbody>
</table>

*Significant at .013 level (two-tailed).

Aerodynamic evaluations.

The dysphonic group demonstrated significantly shorter maximum phonation time (MPT) (p = .004) and higher mean airflow rate (p = .018) than the control group. Table 4 summarized the mean values of the aerodynamic parameters of the two groups.

Table 4. Mean and standard deviations of aerodynamic parameters of the control and dysphonic group.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Control group (n = 14)</th>
<th>Dysphonic group (n = 16)</th>
<th>Independent t Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>(SD)</td>
<td>Mean (SD)</td>
<td>t</td>
</tr>
<tr>
<td>Maximum phonation time (sec)</td>
<td>19.17 (4.15)</td>
<td>13.67 (5.28)</td>
<td>3.14</td>
</tr>
<tr>
<td>Mean expiratory airflow rates (l/s)</td>
<td>0.081 (0.049)</td>
<td>0.13 (0.60)</td>
<td>-2.51</td>
</tr>
<tr>
<td>Peak intraoral pressures (cm H₂O)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1) Consonant-vowel syllable string</td>
<td>10.19 (1.76)</td>
<td>11.11 (3.39)</td>
<td>-0.95</td>
</tr>
<tr>
<td>2) Sentence</td>
<td>9.59 (2.28)</td>
<td>11.03 (2.68)</td>
<td>-0.95</td>
</tr>
</tbody>
</table>

*Significant at .05 level (two-tailed).

Voice range profiles.

The dysphonic group demonstrated significantly smaller profile area (p < .001) and intensity range (p = .002) in the voice range profile than the control group. However, the two groups performed similarly in speech range profile and any differences present were insignificant (p > .007). Table 5 summarized the mean values of parameters in voice range profiles of the two groups.
Table 5. Mean and standard deviations of voice range profile parameters of the control and dysphonic group.

| Parameters                     | Control group  
|                               | (n = 14) | Dysphonic group  
<table>
<thead>
<tr>
<th></th>
<th>(n = 16)</th>
<th>Independent t Tests</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td><strong>Voice range profile</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency parameters (Hz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest frequency</td>
<td>1493.59</td>
<td>1200.00</td>
</tr>
<tr>
<td>Lowest frequency</td>
<td>154.51</td>
<td>163.23</td>
</tr>
<tr>
<td>Frequency range(^a)</td>
<td>39.10</td>
<td>33.31</td>
</tr>
<tr>
<td>Intensity parameters (dB)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum intensity</td>
<td>104.57</td>
<td>98.94</td>
</tr>
<tr>
<td>Minimum intensity</td>
<td>53.57</td>
<td>54.50</td>
</tr>
<tr>
<td>Intensity range</td>
<td>51.00</td>
<td>44.44</td>
</tr>
<tr>
<td>Profile area (dBST)</td>
<td>1294.79</td>
<td>912.06</td>
</tr>
<tr>
<td><strong>Speech range profile</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency parameters (Hz)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest frequency</td>
<td>433.20</td>
<td>376.62</td>
</tr>
<tr>
<td>Lowest frequency</td>
<td>155.79</td>
<td>168.12</td>
</tr>
<tr>
<td>Frequency range(^a)</td>
<td>17.28</td>
<td>13.94</td>
</tr>
<tr>
<td>Intensity parameters (dB)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maximum intensity</td>
<td>83.86</td>
<td>84.81</td>
</tr>
<tr>
<td>Minimum intensity</td>
<td>59.64</td>
<td>61.63</td>
</tr>
<tr>
<td>Intensity range</td>
<td>24.21</td>
<td>23.93</td>
</tr>
<tr>
<td>Profile area (dBST)</td>
<td>252.86</td>
<td>204.13</td>
</tr>
</tbody>
</table>

*Significant at .007 level (two-tailed).

\(^a\) Frequency range was measured in semitone.

**Reliability measurements**

Pearson’s correlation coefficient was used to evaluate the inter-rater reliability in measurements in acoustic perturbation analysis, aerodynamic evaluations and voice range profiles. The correlation coefficient ranged from .978 (p < .001) to 1.00 (p < .001).
Discriminant function analysis

Discriminant function analysis (DFA) with step-wise entry method was used to select the most powerful predictors and to evaluate their predicting power in categorizing participants back into the control group and the dysphonic group. Five instrumental parameters in which the two groups differed significantly were put into DFA. They included relative average perturbation (RAP), shimmer percent (Shim), maximum phonation time (MPT), mean airflow rate, and profile area of voice range profile ($A_{\text{max}}$). The intensity range was not put into DFA because its influence was embedded in $A_{\text{max}}$ and so the independence of variables used in DFA would be better maintained. Shim and MPT were selected as the most powerful predictors. Discriminant function coefficients which reflected the weights of Shim and MPT in predicting the presence of dysphonia were generated by DFA. A discriminant function was formed:

$$\text{Discriminant score (D-score)} = 0.506\text{Shim} – 0.131\text{MPT} – 1.373$$

D-score for each participant was calculated and the scores ranged from -2.13 to 3.03 with an optimal cutting score 0. Children with D-scores below 0 were classified as normal and children with D-scores above zero were classified as dysphonic. This classification achieved an overall predicting accuracy of 93.33% (28 out of 30 participants). Table 6 summarized the predicting accuracies of Shim and MPT in classifying participants into the two groups.
THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

Table 6. Number of participants (percentage) predicted by shimmer percent and maximum phonation time into the control and the dysphonic group by step-wise discriminant function analysis.

<table>
<thead>
<tr>
<th>Perceptual grouping</th>
<th>Normal participants (percentage) predicted by Shim and MPT</th>
<th>Dysphonic participants (percentage) predicted by Shim and MPT</th>
<th>Total participants (percentage) predicted by Shim and MPT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>14 (100)</td>
<td>0 (0)</td>
<td>14 (100)</td>
</tr>
<tr>
<td>Dysphonic</td>
<td>2 (12.5)</td>
<td>14 (87.5)</td>
<td>16 (100)</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>14</td>
<td>30 (100)</td>
</tr>
</tbody>
</table>

Notes. The overall predicting accuracy was 93.3%. The bolded figures represented the correctly predicted number of participants and the corresponding percentage.

Discussion

The investigation of the use of multiparametric approach in the evaluation of voice quality has been a rapidly growing trend in the recent two decades because of the complementation of the approach’s objectivity to the subjectivity of the traditional perceptual evaluation approach and the advancement of technology. Related research has been focusing on the adult population and the results were thrilling (Cantarella et al., 2001; Giovanni et al., 1996; Ma & Yiu, 2006; Wuyts et al., 2000; Yu & Giovanni, 2003; Yu et al., 2001; Yu et al., 2002; Yu et al., 2004). In order to investigate the feasibility of applying the multiparametric approach in the child population, the current study aimed at identifying a minimal set of instrumental parameters in predicting the presence of dysphonia for Cantonese school-age children.

The first hypothesis of the study was that there were significant differences in instrumental parameters between children with dysphonia and children with perceptually normal voices. This hypothesis was testified and positive results were found. Among the parameters extracted from MDVP, KayPENTAX PAS, and Swell’s Phog 2.0, values of six
parameters were significantly different between the control group and the dysphonic group.

The values of these parameters were agreed by two raters with Pearson’s correlation coefficients ranged from .978 ($p < .001$) to 1.00 ($p < .001$).

Two of the six were acoustic parameters, including relative average perturbation (RAP) and shimmer percent (Shim). A number of studies reviewed earlier reported similar findings. Among them, Yu et al. (2004) also identified RAP and Shim as two sensitive parameters for differentiating dysphonic and normal voices. RAP and Shim evaluate the period-to-period variations of pitch and amplitude respectively (Kent & Ball, 2000; Yiu, Worrall, Longland, & Mitchell, 2000). The dysphonic group demonstrated significantly higher RAP and Shim values than the control group. This indicated that the dysphonic group had less stable pitch and amplitude during voicing which could then contribute to perceived hoarseness (Martin & Lockheart, 2000).

Two of the six were aerodynamic parameters, including maximum phonation time (MPT) and mean airflow rate during comfortable phonation. These two parameters were also identified by Yu & Giovanni (2003) and Yu et al. (2002). The task of sustaining a maximum phonation assesses one’s efficiency in managing adequate air supply during phonation (Weinrich, Brehm, Knudsen, McBride, & Hughes, 2013). A short MPT can be attributed by a large amount of air escape. It is related to the mean airflow rate which indicates the rate of air passes through the glottal region during phonation (Weinrich et al.
A high airflow rate indicates the presence of glottal incompetence (Weinrich et al. 2013). The dysphonic group demonstrated significantly shorter MPT and higher airflow rate than the control group suggesting that children with dysphonia had lower efficiency in managing airflow during phonation and incomplete glottal closure could be one of the attributors. However, unlike some literatures, like Ma and Yiu (2006), the intraoral pressures of the two groups were not significantly different. Intraoral pressure is an estimation of subglottal pressure which is the amount of pressure directly below vocal folds during phonation (Weinrich, Salz, & Hughes, 2005). Higher subglottal pressure is needed to set vocal folds with increased stiffness into vibration (Weinrich et al. 2005). Therefore, similar intraoral pressures of the two groups suggested that any increment in stiffness in vocal folds due to vocal pathologies was not significant for the dysphonic participants in the study.

The remaining two parameters were the profile area and the intensity range of voice range profile. Voice range parameters were identified by a number of literatures as sensitive. The study of Yu et al. (2002) found voice range measurement as the most important parameter in correlating with perceptual judgment of dysphonia. The voice range profile reveals one’s maximum intensity range across all his or her phonation frequency (Kent, Kent, & Rosenbeck, 1987). Possible ranges of frequency and intensity to be produced are determined by one’s ability in controlling the tension of the vocal folds and in generating and maintaining the required subglottal pressure (Sapienza & Hoffman-Ruddy, 2009).
dysphonic group demonstrated significantly narrower intensity range and smaller voice range
profile area than the control group indicating the possibility of their vocal pathologies
limiting their maximum potential in voicing with various pitches and intensities.

The second hypothesis of the study was that pathological voices and normal voices in
children could be differentiated using a minimal set of instrumental parameters. Positive
results were also found for this hypothesis. Among the six parameters discussed above,
Shim and MPT were found to be the most powerful predictors of dysphonia in children and
they formed a minimal set of instrumental parameters which successfully differentiated
93.33% of participants with normal or dysphonic voices. The prediction of the presence of
dysphonia using these two parameters was expressed by the discriminant function: 
\[ D\text{-score} = 0.506\text{Shim} - 0.131\text{MPT} - 1.373. \]
In other words, the presence of dysphonia in children can be screened simply by obtaining children’s Shim and MPT values and calculate the D-scores for them. A positive D-score indicates the presence of dysphonia and vice versa.

Furthermore, discriminant function analysis generated a larger discriminant coefficient for
Shim (0.506) than MPT (-0.131) indicating the relative importance of Shim in classifying
normal and dysphonic voices. Also, the opposite sides of the coefficients suggested the
opposite indicating natures of the two. In other words, higher Shim and lower MPT values
indicated higher probability of a child being dysphonic and lower Shim and higher MPT
values indicated higher probability of a child being normal.
The instrumental parameters identified in this study agreed with many of the literatures reviewed previously, especially in Cantarella et al. (2011) who also identified Shim and MPT as two of their three indicators in predicting 82.6% of dysphonic severities in adults. This agreement was not surprising as children had essentially the same physiologies in respiration and phonation when compared to adults though their anatomical structures were constantly maturing. It seemed more surprising that a combination of only two instrumental parameters was able to classify 93.33% of children with and without dysphonia while previous studies only achieved 88% of concordance (Yu et al. 2002) at maximum. However, this high accuracy was reasonable when given a second thought as the current study classified children with and without dysphonia only while previous studies dealt with a lot more variations when classifying adults into various dysphonic severity groups, namely, normal, mildly, moderately, and severely dysphonic.

**Clinical implications**

In clinical setting, dysphonic and non-dysphonic voices were less perceptually distinctive in children than in adults. As observed in this study, certain degrees of breathiness were present in many children. This was supported by Patel, Dixon, Richmond, and Donohue (2012) as they found that, in the population with normal voices, only 14.29% of children (compared to 30.36% in adults) had complete glottal closure during vocal fold adduction and 80.36% (compared to 39.29% in adults) of them had longer open phase during
phonation. Therefore, a certain degree of air leakage leading to perceptual breathiness could be normal for children due to their phonation configurations and this increased the difficulty in perceptually distinguishing those with pathological breathiness from those with acceptable breathiness. Hence, in the cases with perceptual ambiguity, objective parameters identified in this study could give some valuable information when making diagnosis.

The use of instrumental measures was found to be effective, objective and convenient. However, cautions have to be made. As reviewed earlier, despite the successfulness of each individual research, results of studies on multiparametric approach in evaluating adult voice quality were not yet consistent. Similar studies on the child population were even fewer for comparison. Therefore, it is reminded that over reliance on objective parameters for clinical judgment should be avoided before further evidence showing high consistency over the association between perceptual evaluation and instrumental parameters is published.

Limitations and future research directions

The current study successfully identified a combination of two instrumental parameters in predicting the presence of dysphonia in Cantonese school-age children. However, a few factors limited the power of the study. One is the sample size. Involving only 30 participants in total was not ideal as a representative data base. Second is the range of dysphonic severity and types of vocal fold pathologies. The dysphonic group in the study included no severely dysphonic subjects and most of the dysphonic participants were
diagnosed with vocal fold nodules. These also reduced the representativeness of the study.

In spite of the above limitations, the current study was still successful. Further effort is thus worth devoting in the future to maximize the sample size and diversities of dysphonic severity and pathologies. This allows further investigation of the feasibility of using a multiparametric approach in differentiating voices with and without dysphonia and voices with various dysphonic severities in Cantonese school-age children.

**Conclusion**

The present research demonstrated that multiparametric approach is feasible as a complementary use to perceptual judgment in identifying children with normal and dysphonic voices. Significant differences in six instrumental parameters were found between children with dysphonia and children with perceptually normal voices. Two of them formed a minimal set of instrumental parameters to differentiate pathological voices and normal voices in Cantonese school-age children with 93.33% accuracy.

**Acknowledgement**

I would like to express my gratitude to my supervisor, Dr. Estella Ma, for her valuable guidance and time devoted into the study. I would like to acknowledge Dr. Raymond Tsang who made diagnoses for the vocal fold pathologies and Ms. Soby So who evaluated the voice samples in this study. I would also like to express my appreciation to all the participants and their parents for supporting the study.
References


THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

Logopaedica, 48, 175-185.


Phoniatrica Et Logopaedica, 41(2-3), 89-144.


Group.


evaluation of voice quality: Review, tutorial, and a framework for future research.


Lee, L., Stemple, J. C., Glaze, L., & Kelchner, L. N. (2004). Quick screen for voice and

supplementary documents for identifying pediatric voice disorders. Language, Speech,

& Hearing Services in Schools, 35(4), 308-319.


20(3), 380-390.


Seminars in Speech and Language, 28(4), 343-350.


THE EVALUATION OF DYSPHONIA: A MULTIPARAMETRIC APPROACH

(Unpublished dissertation.) The University of Hong Kong, Hong Kong.


Voice Handicap Index (pVHI): A new tool for evaluating pediatric dysphonia.

*International Journal of Pediatric Otorhinolaryngology, 71*(1), 77-82.
Appendix

Reading material for speech range profile (taken from local primary school textbook)

說話不簡單

上課了，山羊老師要教大家說話。小牛和小馬聽了，都覺得好笑，心裡想，誰不會說話呢？

山羊老師請同學們出來練習說話。小牛第一個舉手要說笑話。他越說越快，越說聲音越小，還沒說到一半，就笑個不停。

小馬出來給大家講故事，同學們都專心聆聽。可是，小馬前言不搭後語，大家越聽越不明白。

小牛和小馬終於知道，說話真不簡單，也要好好學習。