Acoustical Evaluation of Voice across Different Levels in Vocal Professionals

Sanazeb Pathan* and Yaashna Rajani
Audiologist and Speech Language Pathologist, India

Submission: July 15, 2017; Published: August 09, 2017

*Corresponding author: Sanazeb Pathan, Audiologist and Speech Language Pathologist, India

Introduction

Voice plays a major role in speech & also in communication. Communication skills are important to almost everybody as our society has become service-oriented. Occupational voice health is becoming more important as more people rely on their voices for their work. ‘Vocal professionals’ constitute an ever increasing segment of the population who depend on their voice for a daily living [1]. According to Stemple, Glaze and Gerdman [2], vocal professionals by the very nature of their occupations are at a greater risk of developing voice problems and laryngeal pathologies than general population. The demand on an individual’s vocal system depends on the kind of occupation one is engaged in. Vocal professionals, especially those in the speaking profession require certain qualities in their speech to be successful in their profession such as good pitch range, rhythm, melody, fluency, phrasing, emphasis, modulation & good expiratory air to sustain speech.

Need of the Study

As vocal professionals have profession specific needs, their voice parameters vary depending on their respective professions. Therefore, in order to understand the vocal mechanism of vocal professionals it is very necessary to evaluate the acoustical parameters in voice. There is a need for research that would determine how the speech language pathologists can best be of service to different levels of vocal professionals who are at risk of developing a vocal pathology. Most of the studies among vocal professionals have been conducted on the singers, actors, radio jockeys, teachers, telephone operators etc.; [3-6]. But research on voice of lawyers, doctors, businessmen, clerks, laborers, factory workers etc. is less addressed in literature. Therefore, the acoustic parameters of vocal professionals should be explored so that it can be used by the speech language pathologists in future to evaluate the voice of at risk vocal professionals.

Mostly Indian studies have concentrated more on surveying the risk factors and vocal symptoms in teachers, singers, actors, telephone operators etc.; [7,8]. There have been very few attempts to compare the acoustic parameters of voice of vocal professionals across the levels, however still there is a lacuna in the existing research. The implication of studying the acoustical parameters will be mainly in terms of early identification, evaluation and remediation of at risk vocal professionals who have profession specific vocal demands.

Aim and Objective of the study

The present study aimed to measure the acoustic parameters of voice across different levels of vocal professionals.

Method

The present study was aimed to find out the difference in the voice parameters of vocal professionals across the 4 levels of vocal professionals as classified on the basis of their vocal usage [9].

Participants Of The Study

A total of 120 participants were selected across 4 levels of vocal professionals. Each group consisted of three types of vocal professionals. Total of 10 participants per profession which included 5 males and 5 females were included as shown in the Table 1. The data was collected from radio station Red FM 93.5 Pune, B.V.P hospital, B.V.P Law College, B.V.P Fine Arts College, factories, academic offices, shops, telephone exchange in and around the city of Pune.
Inclusion criteria

a) All participants were within the age range of 20 to 40 years.

b) All participants included in the study were native speakers of Marathi.

Material

A standardized reading passage in Marathi called 'Asabantomoti' developed by Deshmukh [10] was used for the reading task. This passage consisted of 152 words.

Procedure

A detailed case history was taken before the participants were ready to be a part of the present study. All participants signed an informed consent form after they were explained the purpose of the study. A questionnaire was compiled and filled by the clinician for the purpose of obtaining the participants demographic details and information regarding their vocal habits as well as voice use. Voice samples were recorded in a quiet, isolated room with a portable high fidelity Sony digital voice recorder (model ICD-PX312) with an in-built microphone. The sampling frequency of the recorder was 44.1 kHz with the bit rate of 48Kbps. It was hung on the neck of the participant at a distance of 12cm from the mouth.

Participants were asked to perform two tasks, sustained vowel phonation /a/ and reading task. In the first task of sustained phonation the participants were instructed to take a deep breath and then phonate the vowel /a/ for as long as possible for three trials. The best of the three trials was considered as the Maximum Phonation Duration (MPD). The steady state portion of vowel /a/ was selected for acoustic analysis. A 60 second long sample was analyzed.

The second task was to read a standardized Marathi passage at normal conversational pitch and loudness.

Acoustic Analysis

Acoustic analysis of audio-recorded voice samples was done using the Multi-Dimensional Voice Program Advanced (MDVP-A) of Computerised Speech Lab (CSL) 4500 module (KayPentax, Lincoln Park, NJ). The audio-recorded voice samples were linefed into the CSL module using an XLR female connector jack of 3.5 mm in diameter. A sampling rate of 50,000 Hz and the bit rate of 16Kbps were used for acoustic analysis.

Acoustic Parameters

MDVP Advanced software was used to extract 33 voice-related parameters out of which 8 parameters were selected for the present study. The following 8 parameters of voice were measured for the recorded voice samples of sustained phonation and reading. The acoustic parameters analysed were Mean Fundamental Frequency (mF0), Jitter Percent (jitt), Relative Average Perturbation (RAP), Fundamental Frequency Variation (vF0), Shimmer Percent (shim), Amplitude Perturbation Quotient (APQ), Noise-to-Harmonic Ratio (NHR), Soft Phonation Index (SPI), Maximum Phonation Duration (MPD) and S/z ratio. The normative data developed by Hema, Sangeetha, Pushpavati [11] for the selected parameters were used as Indian standards for the present study with +/- 1 SD criteria to elicit the range.

Results

The parameters were analysed and compared across two tasks; viz sustained phonation task and reading task. Researchers have used sustained vowel phonation task widely for the acoustic analysis [12-15] as it gives you a steady state sample compared to connected sample such as reading or spontaneous speech. But, it is clear that the sustained phonation sample is not enough to represent the multidimensional aspects of connected speech [16,17]. The connected speech sample will have aspects like articulatory rate, precision, intonation, intention etc which may show valuable changes in voice production [15]. Hence the connected speech along with sustained phonation task is found to provide vital information in acoustic parameters. Results of the present study are discussed task wise for eight parameters included in the study.

Task of Sustained Phonation

Descriptive statistics mean and standard deviation were calculated separately for each of the parameters across all four levels. The gender differences are evident across most of the acoustic parameters, which is also reported by the literature [18-22]. The graphs 4, 5, 6, 7, 8, 9, 10 & 11 show the average of acoustic parameters for males and females. The differences in means are graphically shown, but to check whether these differences are statistically significant MANOVA was carried out. Results of MANOVA for sustained phonation task indicates that there is no significant difference between vocal professionals across the levels F (24, 305.13) = 1.23, p=0.05. The studies previously done also conclude that there was no difference in the vocal professionals across sustained phonation task [15]. There

Table 1: Participant Details.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Singers</td>
<td>Actors</td>
<td>Radio operators</td>
<td>Telephone operators</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
</tbody>
</table>

was a significant difference noticed for sustained phonation task across the gender F (8, 105) = 57.28, p<0.01. The gender differences were observed in the parameters of mf0 (p<0.01), jitter (p<0.01), RAP (p<0.01), APQ (p<0.01), NHR (p<0.01) etc. The present study supports the results found in the literature where the gender differences were evident for above parameters [18-23].

**Task of Reading**

Descriptive statistics mean and standard deviation were calculated separately for each of the parameters across all four levels. The gender differences are evident across most of the acoustic parameters as reported by the literature [15,20,21]. Hence the mean of parameters are tabulated separately for males and females across all four levels. The graphs 12, 13, 14, 15, 16, 17, 18 and 19 show the average of acoustic parameters for males and females for the reading task. The differences in means are graphically shown, but to check whether these differences are statistically significant MANOVA was carried out.

Results of MANOVA for reading task indicates there is a significant difference across the levels F (24, 305.13) = 3.74, p<0.01. The studies previously done conclude that there was a significant difference in acoustic parameters among the vocal professionals for the reading task [20,21,23,24]. There was a significant difference noticed between the genders F (8, 105) = 49.02, p<0.01. The gender differences were observed in the parameters of mf0 (p<0.01), jitter (p<0.01), RAP (p<0.01), APQ (p<0.01), NHR (p<0.01) etc. The present study supports the results found in the literature where the gender differences were evident for above parameters [18-23].

**Table 2**: Results of tests of between subject effects for reading task across levels.

<table>
<thead>
<tr>
<th>Source</th>
<th>Dependent variable</th>
<th>df</th>
<th>Mean Square</th>
<th>F value</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Levels</td>
<td>mf0</td>
<td>3</td>
<td>325.11</td>
<td>4.90</td>
<td>0.00**</td>
</tr>
<tr>
<td></td>
<td>Jitt</td>
<td>3</td>
<td>3.53</td>
<td>5.29</td>
<td>0.00**</td>
</tr>
<tr>
<td></td>
<td>RAP</td>
<td>3</td>
<td>0.97</td>
<td>4.68</td>
<td>0.00**</td>
</tr>
<tr>
<td></td>
<td>vf0</td>
<td>3</td>
<td>239.91</td>
<td>8.73</td>
<td>0.00**</td>
</tr>
<tr>
<td></td>
<td>Shim</td>
<td>3</td>
<td>4.99</td>
<td>1.11</td>
<td>0.34</td>
</tr>
<tr>
<td></td>
<td>APQ</td>
<td>3</td>
<td>13.57</td>
<td>1.74</td>
<td>0.16</td>
</tr>
<tr>
<td></td>
<td>NHR</td>
<td>3</td>
<td>0.02</td>
<td>6.41</td>
<td>0.00**</td>
</tr>
<tr>
<td></td>
<td>SPI</td>
<td>3</td>
<td>68.75</td>
<td>3.23</td>
<td>0.02*</td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level

As seen in Table 2, a one way between subjects multivariate analysis of variance was carried out to assess the difference between the levels across each parameter. For reading task, fundamental frequency information measures (mf0), F(3)= 4.90, p<0.01, short term and long term frequency perturbation parameters like (jitt), F(3)= 5.29, p<0.01, (RAP), F(3)= 4.68, p<0.01, (vf0), F(3)= 8.73, p<0.01, noise related parameters (NHR), F(3)= 6.41, p<0.01, and (SPI), F(3)= 3.23, p<0.05, showed statistically significant difference across levels. Bonferroni’s pair wise multiple comparisons across levels were carried out Table 2.

**Table 3**: Post hoc analysis of mean fundamental frequency (mf0) for reading task across levels.

<table>
<thead>
<tr>
<th>mf0</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>20.76**</td>
<td>16.83</td>
<td>22.97**</td>
<td>2.00</td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level

For the parameter mf0 there was a statistically significant difference noticed between the levels I & II (p<0.01) as well as levels I & IV (p<0.01). In the reading task there are patterned variations of pitch over linguistic units of differing length, such inflections may lead to change in the mean fundamental frequency for the reading task. Inter speaker variability also contributes to such differences in mean fundamental frequency. These findings are reported in literature [4,19,20,23,25,26]. In case of elite vocal performers higher mf0 is present as a consequence of the training they undergo. This finding supports the studies done by Sheela [27] where the results indicated that trained vocal professionals had significantly different acoustic parameters as compared to the untrained professionals (Table 3).

For the parameter jitter (p<0.01) & RAP (p<0.01) there was a statistically significant difference noticed between the levels I & IV as well as between levels II & IV. In the reading task, there are abrupt transitions which occur during voicing onset and offset. There is a certain degree of apparently random variability of the cycle to cycle duration. Age, gender, oral and nasal sounds present in the reading text, are all contributing to the variations in the jitter across levels [20,23,28-31]. This difference in the jitter can also be attributed to the difference noticed in the mean fundamental frequency across levels. Literature indicates that jitter is influenced by mean f0 of the speaker’s phonation [32]. Although there was significant difference across the levels, still the jitter values were within normal limits. Jitter variations are...
also observed in elite vocal performers who smoke and consume alcohol [23,30].

The present study shows significant differences in the jitter values of elite vocal performers when compared to the other three levels of vocal professionals. Also, in professional voice users there is a common vocal symptom of vocal fatigue which may lead to variations in the jitter. Heterogeneity in the levels with respect to age and professions also might have contributed to such variations in the jitter values as reported in the literature [22,33]. (Tables 4 & 5).

**Table 4:** Post hoc analysis of jitter (jitt) for reading task across levels

<table>
<thead>
<tr>
<th>jitt</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>0.03</td>
<td>0.49</td>
<td>0.69**</td>
<td></td>
</tr>
<tr>
<td>Level II</td>
<td></td>
<td>0.45</td>
<td>0.66**</td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td></td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level**

**Table 5:** Post hoc analysis of relative average perturbation (RAP) for reading task across levels

<table>
<thead>
<tr>
<th>RAP</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>0.00</td>
<td>0.23</td>
<td>0.35**</td>
<td></td>
</tr>
<tr>
<td>Level II</td>
<td></td>
<td>0.24</td>
<td>0.36**</td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td></td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level**

For the parameter vf0 there was a statistically significant difference noticed between the levels I & II (p<0.01), levels I & III (p<0.01) as well as levels I & IV (p<0.01). vf0 changes during phonation as well as reading tasks. This is specifically noticed because there are changes in the mf0 for different tasks. A variety of factors influence this change in reading such as the number of words spoken, amount of intonation used and amount of air pressure varied as reported in the literature [15,20].

Although there were significant differences in the vf0 between the levels the variations were within the normal limits. The difference across levels can be attributed to the mean age (mentioned in the method) differences observed in the present study. It has also been reported that there are age dependent variation in vf0 [22].

Table 6: Post hoc analysis of variation in fundamental frequency perturbation (vf0) for reading task across levels.

<table>
<thead>
<tr>
<th>vf0</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>4.38**</td>
<td>3.96*</td>
<td>6.83**</td>
<td></td>
</tr>
<tr>
<td>Level II</td>
<td></td>
<td>-0.42</td>
<td>2.44</td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td></td>
<td>2.86</td>
<td></td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level**

For the parameter NHR there was a statistically significant difference noticed between levels I & II (p<0.01), levels I & III (p<0.01) as well as levels I & IV (p<0.01). v0 changes during phonation as well as reading tasks. This is specifically noticed because there are changes in the mf0 for different tasks. A variety of factors influence this change in reading such as the number of words spoken, amount of intonation used and amount of air pressure varied as reported in the literature [15,20]. Although there were significant differences in the vf0 between the levels the variations were within the normal limits. The difference across levels can be attributed to the mean age (mentioned in the method) differences observed in the present study. It has also been reported that there are age dependent variation in vf0 [22].

For the parameter NHR there was a statistically significant difference noticed between levels I & III (p<0.01) and also between levels I & IV (p<0.01). Higher MPD gives rise to higher respiratory support and lesser effort on laryngeal valving during vocal fl. closure which in turn results in higher number of harmonics than noise [24]. NHR may have varied across levels due to the symptom of GERD which occurs due to late meals (67%) [26]. Literature has also stated that noise related parameters can have a nonlinear trend of variations when the voice samples were recorded versus when live voice samples are used [29]. Authors have found that there were significant differences in the noise related parameters in recorded voice as compared to live voice because digitization and sampling frequency might have altered the signal to noise ratio of the sample which may in turn must have led to the variability in NHR.
Literature has postulated that variations in the length of the vocal tract, resonating quality of voice, effect of training, vocal loading per day etc. may contribute towards the variations in f0 [6,23,27,28,34,35,37]. In the present study vF0 is significantly different in elite vocal performers as compared to the other three levels. This must be because vocal professionals in level I are trained for their profession as compared to the other three levels as discussed in the literature by Boominathan, Rajendran, Nagarajan, Seethapathy & Gnanasekar [7] (Table 7).

### Table 7: Post hoc analysis of noise to harmonic ratio (NHR) for reading task across levels.

<table>
<thead>
<tr>
<th>NHR</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td>0.03</td>
<td>0.05**</td>
<td>0.06**</td>
<td></td>
</tr>
<tr>
<td>Level II</td>
<td>0.01</td>
<td>0.02</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td>0.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level

For the parameter SPI there was a statistically significant difference noticed between the levels I & IV (p<0.01). SPI is an evaluation of the poorness of high frequency components that may be an indication of loosely adducted vocal folds during phonation. As it provides information related to glottic closure, SPI value and asthenic voice quality are significantly related [38]. Due to the vocal fatigue there could be a possibility of weakness in voice which must have led to incompletely adducted vocal folds. This partial adduction of vocal folds must have caused variations in SPI in the first two levels of vocal professionals.

Our study supports the previously done studies [6,39,40] where the results indicated that due to varied vocal demands there are alterations in the vocal tracts which lead to increased spectral energy in higher part of spectrum which in turn caused incomplete vocal fold adduction. Effect of training was also attributed to the variations in SPI especially in elite vocal performers [41] (Table 8).

### Table 8: Post hoc analysis of soft phonation index (SPI) for reading task across levels.

<table>
<thead>
<tr>
<th>SPI</th>
<th>Level I</th>
<th>Level II</th>
<th>Level III</th>
<th>Level IV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level I</td>
<td></td>
<td>-3.70**</td>
<td>-1.73</td>
<td>-1.87</td>
</tr>
<tr>
<td>Level II</td>
<td>1.96</td>
<td></td>
<td>1.82</td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td>-0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**significant difference at 0.01 level

For maximum phonation duration the results reveal that mean values for level I was comparatively greater than the other levels. When we compare the means for MPD across gender the results reveal that MPD was higher for males as compared to females (Figure 1). There is a difference in the mean of MPD across levels and between genders. Therefore to find out whether the difference was statistically significant MANOVA was carried out.

### Figure 1: Mean values of maximum phonation duration (MPD) between genders.

Results of MANOVA for maximum phonation duration reveal that there is a significant difference between the levels F (3) = 9.879, p<0.01 and also between the genders F (1) = 88.73, p<0.01. Bonferroni’s pair wise multiple comparisons across levels were carried out. Parameter wise results are summarized in the Table 18. The results indicate that there was a significant difference in MPD between the level I & II (p<0.01), level I & III (p<0.01) and level I & IV (p<0.01). This difference was significant as the vocal professionals in the level I (actors, singers) were trained for their voice as compared to the other three levels. Literature also reports evidences of effects of training increases maximum phonation duration [3,27]. Also, there was a gender wise statistically significant difference observed in which the males were able to sustain phonation for longer duration than females as reported the literature [42,43,44-60].

### a) S/z ratio:

For the task of s/z ratio the results reveal that there was no difference in the mean values across all the levels as well as genders, as shown in the Figure 2. Results of MANOVA for s/z ratio reveal that there is no significant difference between the levels F (3) = 0.950, p=0.05 and also across gender F (3) = 0.06, p=0.05. The reason for obtaining no significant differences could be because participants having normal voice were selected for the study. According to Indian standards an average normal speaker is expected to sustain both voiceless /s/ and the voiced /z/ for approximately equal duration’s resulting in a ratio of one [24]. Acoustic analysis of voice seems to be an effective method for voice evaluation in vocal professionals who are at risk for developing any vocal pathology. Therefore prevention and early
detection of these voice disorders are of particular clinical importance. This study is a preliminary step to commence profession specific voice evaluation, as there are various professional vocal demands across different levels [60-122].

Figure 2: s/z ratio.

References

Global Journal of Otolaryngology

39. Froseheles E (1939) Twentieth century speech and voice correction. Philosophical Library, New York, USA.
57. Colton RH, Casper, JK (1996) Understanding Voice Problems: A Physiological Perspective for Diagnosis and Treatment (2nd edn.), Williams & Wilkins, Maryland, USA.


