Spectral and Cepstral Measures of Vocal Fatigue in Indian Heavy Metal Vocalists

S. V. Narasimhan, Mahesh Soumya
Department of Speech and Hearing, JSS Institute of Speech and Hearing, Mysuru, Karnataka, India

Abstract

Introduction: Heavy metal vocalists tend to overuse or differently use their vocal systems for singing, and this could have a negative impact on their voice quality. However, there is limited research concerning the objective measures of vocal changes caused due to vocal fatigue in Indian heavy metal vocalists. Thus, the objectives of the study were to document the differences in the cepstral and spectral parameters of voice before and after vocal fatigue in Indian heavy metal vocalists and to investigate the correlation between the phonatory habits and the cepstral and spectral parameters in postfatigue voice of Indian heavy metal vocalists. Methods: Phonation samples were collected before the vocal performance (pretest) and after the vocal performance (posttest) from 16 vocalists belonging to Indian heavy metal genre. The phonation samples were analyzed, and the spectral and cepstral parameters were extracted in both the conditions. Results: Out of spectral parameters, only H1-H2 was significantly different between the pre- and postfatigue voices. Both the cepstral parameters were noted to be significantly lower after vocal fatigue. Moderate and strong correlations were observed between the phonatory habits and the objective parameters of postfatigue voice in heavy metal vocalists. Conclusion: The cepstral measures of voice were more effective in identifying the vocal fatigue-related voice changes in heavy metal vocalists.

Keywords: Cepstral parameters, heavy metal vocalists, singers, spectral parameters, vocal fatigue, vocal loading

INTRODUCTION

Vocal fatigue refers to a condition associated with the loss of voice as a result of increased phonatory effort or vocal loading.[1] Vocal fatigue has been traditionally defined in terms of the symptoms such as increased vocal effort leading to a continuing sensation of vocal tiredness, excessive usage of voice, reduced voice control, laryngeal discomfort, and loss of voice that worsens with vocal loading and frequently advances with time as experienced by the voice user.[2-7] The changes in voice caused due to prolonged or excess demand on voice is known as vocal loading.[8,9]

Studies have also investigated the effects of vocal loading on the acoustic characteristics of voice and have reported various acoustic changes such as increase in fundamental frequency,[5,10] increase in sound pressure level and shimmer,[11] increased jitter,[10] decreased jitter,[5] or no essential change in jitter[12] due to vocal loading. Investigations on the short-term effect of vocal loading and also the effect of vocal rest on voice have reported an increase in jitter, shimmer, and fundamental frequency after the performance, and these parameters showed recovery back to baseline after 17–18 h of voice rest.[13]

Through the years, while it has been well documented that there exist characteristic acoustic differences in the singer’s voice, the cepstral characteristics of voice in singers have not been researched on much. The cepstral analysis is one popular tool used for voice analysis. Cepstrum is defined as a discrete Fourier transform of the logarithm power spectrum. Cepstral measures indicate the extent of harmonic organization in voice, i.e., a prominent cepstral peak reflects a spectrum having a rich harmonic structure. Research on the cepstral analysis of the singer’s voice has shown that trained singers showed significantly higher Smoothed cepstral peak prominence (CPPS) values compared...
Researchers have also reported higher cepstral peak prominence (CPP) and CPPS values among singers and have concluded that the voice of Carnatic singers had better harmonic structure and organization compared to the voice of nonsingers.[16,18] Apart from the traditional acoustic measures, spectral analysis of voice offers a visual representation of the various peaks and formants and also valuable formant and harmonically related information.[19] Various studies have also used spectral measures of voice to document the characteristics of the singer’s voice.[20-22] Past studies have found the presence of the singer’s formant that refers to the concentration of energy around 2000–3000 Hz.[23] in the singer’s voice.

Studies on spectral analysis have also identified the harmonic amplitude differences or the parameters of spectral tilt as the major features influenced by the quality of voice.[24,25] Parameters of spectral tilt or the harmonic amplitude differences are determined by comparing the amplitude of the various harmonics and formant frequencies. Harshan and Narasimhan studied the spectral parameters of voice in various registers of singing among classical Carnatic singers and reported that the spectral parameters were sensitive to the changes in vocal characteristics in various registers of singing in Carnatic singers.[26] Even though the parameters of spectral tilt or amplitude differences of voice provide a quantifiable acoustic index of the vocal fold adduction,[27] only a few studies have used these measurements to assess the voice characteristics in singers.

In the existing literature, the spectral and cepstral characteristics of voice in singers have been documented primarily from the Western classical tradition or Carnatic music singers. Other than these groups, attempts to investigate vocal aspects in vocalists of various genres of singing – such as heavy metal singing or death metal singing – have been limited.

Heavy metal is a genre of rock music that came into existence around the 1970s. Heavy metal vocalists usually overdrive their voices, tend to sing longer sustained notes, and produce heavy vibrato for projecting their voice. Heavy metal genre is characterized by a combination of vocal settings such as tensed vocal folds and vocal tract, elevated laryngeal position, pharyngeal constriction, wide vocal range, and use of various quality of voice such as harsh, falsetto, whisper, or creaky voice.[27] Most of the heavy metal vocalists tend to overuse or differently use their vocal systems for singing, resulting in varying degrees of vocal strain[28] and could have an adverse impact on overall voice quality.[29] With the increasing popularity of the metal singing genre,[30] there is a need to investigate the effect of vocal fatigue on the spectral and cepstral measures of voice in Indian heavy metal singers.

From the literature review, it is apparent that the cepstral and spectral parameters provide a robust and quantifiable tool for documenting the voice changes in singers due to vocal fatigue. Even though the correlation between the phonatory habits (as vocal fatigue frequently is a subjective phenomenon) and the objective parameters (cepstral and spectral parameters) of voice can enable the voice pathologist to be more effective in documenting the changes in singers’ voice secondary to vocal fatigue, only a few studies have examined the association between the cepstral and spectral measures. Hence, the objective of this study was to document the differences in the cepstral and spectral parameters of voice in Indian heavy metal vocalists before and after vocal fatigue. The second objective was to investigate the correlation between the phonatory habits and the cepstral and spectral parameters in postfatigue voice of Indian heavy metal vocalists.

**Methods**

**Participants**

The study included 16 male heavy metal vocalists (extreme metal screamers) of Indian origin, who were professionally trained for at least 5 years. The mean age of the participants was 30.09 ± 1.72 years (range: 27–32 years). Participants were vocalists in a group of Indian heavy metal band and were recruited by the first author from a music recording studio at Bangalore. Participants were assessed by a qualified otolaryngologist, and laryngoscopic evaluation revealed vocal folds to be structurally and functionally normal in all the participants. A questionnaire in English [Appendix 1] was administered on all the participants and the details on their phonatory habits, previous history of voice problems, tobacco usage, voice therapy, or surgical or medical intervention of all the participants. The responses from the questionnaire were examined, and the participants having any recent history or current complaints of respiratory tract infections, laryngeal surgeries, tobacco usage, hearing loss, or family history of voice disorders during recording were not involved in the present study. The number of participants was determined based on the statistical power analyses using G-Power (ver. 3)[31] using the mean and standard deviation values of the a spectral parameter obtained in a pilot study using five Indian heavy metal vocalists and five nonsingers. Convenience and purposive sampling were employed in selecting the participants. All the participants were explained the purpose of the voice recording, and informed consent was obtained.

**Procedure**

The voice recording was carried out in a soundproof music recording studio in two conditions. Pretest (prefatigue condition) was 30 min before the vocal performance of the day started and was after 15 min of warm-up exercises. Posttest (postfatigue condition) was within 30 min after the completion of vocal performance and before the cooldown exercises started. The participants did not consume water or any type drink between the two recordings. Vocal performance included the jam sessions comprising the rehearsal of the album songs to make performance-ready in the soundproof music recording studio. The jam sessions included 90–120 min of continuous singing or reading sheet music at high pitches and loud volumes with the background music on electric guitar
and keyboards. The background noise levels in the recording studio during the performance were around 95–100 decibels.

During the voice recordings in both the conditions (i.e., pretest and posttest conditions), the participants were instructed to phonate the vowel /a/ at their habitual pitch and comfortable loudness levels thrice. A dynamic unidirectional microphone with a frequency response of 60–12,000 Hz was used to record all the voice samples. The participant’s mouth-to-microphone distance was 10 cm, and the voice was recorded onto the Praat software (version 6.0.56).[32] The samples were digitized at a sampling frequency of 44.1 kHz and 16 bits/sample quantization.

Analysis
Out of the three recordings obtained, one recording was chosen randomly for analysis. A steady portion of 6–8 ms in the middle portion of vowel /a/ was selected for analysis. The cepstral analysis was carried out using a speech tool, and the values of the cepstral parameters (CPP and CPPS) were obtained using the Hillenbrand algorithm for all the phonation samples obtained on both the pretest and posttest conditions. The recorded samples were transformed into the frequency domain, and the spectral slices of the samples were extracted. The spectral amplitudes of H1 (first harmonic), H2 (second harmonic), A1 (harmonic in the first formant frequency region), A2 (harmonic in the second formant frequency region), and A3 (harmonic in the third formant frequency region) were identified from the spectral slice and the spectral tilt parameters, namely H1-H2, H1-A1, H1-A2, and H1-A3, were then calculated, and the results were tabulated for the samples recorded in both conditions.

Results
The tabulated data were analyzed statistically using the Statistical Package for the Social Sciences (SPSS version 24) SPSS software (version 24, SPSS Inc., Illinois, USA). The results of the descriptive and inferential statistics for the spectral and cepstral parameters of voice between pretest and posttest conditions are presented in Graph 1 and Table 1. The results showed that the mean values of all the spectral parameters were higher in posttest compared to pretest condition. The mean values of the cepstral parameters (CPP and CPPS) were noted to be lower in posttest compared to pretest condition. Standard deviation values were noted to be higher for both cepstral and spectral measures in posttest condition, except for H1-A1 and CPP.

Shapiro–Wilk test ($P > 0.05$) was carried out to check the distribution of data across both conditions. The results showed that the data were normally distributed across both conditions for all the spectral and cepstral parameters. A paired sample $t$-test at a significance level of 0.05 (95% confidence interval) was carried out to see the effect of the independent variable (pretest and posttest conditions) on every dependent variable (spectral and cepstral parameters). The results showed that there were no significant differences in any of the spectral parameters except H1-H2 between the two conditions. Both the cepstral parameters, i.e., CPP and CPPS, showed statistically significant differences between both conditions. Thus, it was inferred that the mean values of H1-H2 were significantly higher, and the mean values of CPP and CPPS were significantly lower postvocal fatigue.

Table 2 presents the results of Spearman’s correlation coefficient carried out to check for the strength of the correlation between the phonatory habits and both the spectral and cepstral parameters of postfatigue voice in metal singers. Out of five phonatory habits, the hours of voice usage and the amount of water intake were considered for correlation with all the objective parameters of voice as all the other questions related to phonatory habits had similar responses from all the participants. The results showed that there was a strong correlation between the hours of voice usage and the two spectral parameters, namely H1-H2 and H1-A3. However, the cepstral parameters, i.e., CPP and CPPS, showed a strong and moderate negative correlation with the hours of voice usage. It was also noted that all the spectral parameters showed a strong negative correlation with the amount of water intake except H1-A1 that showed a moderate negative correlation. The cepstral parameter CPP showed a strong positive correlation and CPPS showed a moderate positive correlation with the amount of water intake. All the correlation coefficient values were found to be significant at an alpha level of 0.05.

Discussion
The study investigated the cepstral and spectral parameters of...
Narasimhan and Mahesh: Vocal fatigue in Indian heavy metal vocalists

The study results also revealed significant differences in spectral and cepstral parameters (CPP and CPPS) of voice between the prefatigue and postfatigue conditions in Indian heavy metal vocalist. Spectral measures were significantly higher and cepstral parameters were significantly lower in postfatigue voice compared to the prefatigue voice. These changes might be expected as the heavy metal genre is characterized by a combination of vocal settings such as tensed vocal folds and vocal tract, elevated laryngeal position, pharyngeal constriction, wide vocal range, and use of various quality of voice such as harsh, falsetto, whisper, or creaky voice[27] that could result in vocal strain[28] and have an adverse impact on overall voice quality[29] following vocal fatigue.

It was also noted from the results of the present study that out of four spectral parameters analyzed, only H1-H2 showed a significant difference between the prefatigue and postfatigue voices. Nevertheless, both the cepstral parameters, i.e., CPP and CPPS, revealed significant differences between the prefatigue and postfatigue voices. Thus, the results suggest that the cepstral measures of voice were more effective in identifying the changes in the singer’s voice caused due to vocal fatigue. Earlier studies on dysphonic voice have also reported similar findings and have considered the cepstral measures of voice to better predict the severity of dysphonia than spectral measures.[30,31] Therefore, the results of the study are in accordance with the results reported by the past studies on the analysis of voice in patients with voice disorders.

The study results also revealed that there was a strong correlation between the phonatory habits and the objective parameters of voice in posttest condition. The spectral parameter showed a positive correlation with the hours of voice usage and negative correlations with the amount of water intake. Therefore, the hours of voice usage and spectral parameters show a direct relationship, whereas the amount of water intake and the spectral parameters show an inverse relationship. Thus, it was inferred that the increased hours of voice usage and reduced amount of water intake might result in higher values of spectral parameters. However, the cepstral parameters, namely CPP and CPPS, showed a negative correlation with the hours of voice usage and positive correlations with the amount of water intake. Therefore, it was concluded that the increased hours of voice usage and reduced amount of water intake might result in reduced values of cepstral parameters indicating a voice with the poor harmonic organization.

The present study also has a few limitations. First, the study considered only the male vocalists belonging to Indian heavy metal genre, and thus, the study failed to comment on the

---

Table 2: Spearman’s correlation coefficient between spectral and cepstral measures of voice

<table>
<thead>
<tr>
<th>Parameters</th>
<th>H1-H2</th>
<th>H1-A1</th>
<th>H1-A2</th>
<th>H1-A3</th>
<th>CPP</th>
<th>CPPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hours of voice use</td>
<td>0.84</td>
<td>0.56</td>
<td>0.61</td>
<td>0.83</td>
<td>−0.83</td>
<td>−0.61</td>
</tr>
<tr>
<td>Water intake</td>
<td>−0.71</td>
<td>−0.54</td>
<td>−0.76</td>
<td>−0.84</td>
<td>0.74</td>
<td>0.69</td>
</tr>
</tbody>
</table>

*Significant at an alpha level of 0.05 (95% confidence interval). CPPS: Cepstral peak prominence smoothed; CPP: Cepstral peak prominence

---

The study results also revealed that both the cepstral parameters were significantly lower after vocal fatigue. As cepstrum is the measure of the degree of harmonic organization in the vocal signal, reduced values of the cepstral parameters indicate the presence of noise in voice signal.[32,33] Thus, significantly lower CPP and CPPS values observed after the vocal fatigue in this study may be attributed to the abnormality in the vibratory nature of the vocal folds leading to aperiodicity in vocal fold vibration resulting in increased noise component and poor harmonic structure at the level of the glottis due to vocal fatigue. Therefore, the present study results confirm that the spectral measure, especially H1-H2 and cepstral parameters of voice, can be used as efficient tools to study the voice changes due to vocal fatigue in singers.

Graph 1: Mean and standard deviation of spectral and cepstral parameters in condition 1 and condition 2.
gender-linked differences. Second, as the responses from the questionnaire were only used to exclude the participants having the history or the presence of voice disorders, respiratory problems, hormonal imbalances, and other associated problems from the study, and as no specific questionnaire on Indian heavy metal vocalist was readily available that could be used for selection of the participants in the present study, the study did not use any standardized questionnaire. The questionnaire was constructed using the questions asked during the case history for voice-disordered patients that included details on the phonatory habits, work habits, previous history and presence of voice problems, hormonal imbalances, respiratory problems, tobacco usage, voice therapy, or surgical or medical intervention. Finally, even though the study reported a moderate and strong correlation between the phonatory habits and the objective parameters of voice heavy metal vocalists, considering the subjective ratings of vocal fatigue would have provided a better insight on the correlation between the subjective and objective measures and better quantifying vocal fatigue in Indian heavy metal vocalists. Nevertheless, the results of this study confirm that the spectral measure, especially H1-H2, and cepstral parameters of voice can be used as efficient tools to investigate the voice changes due to vocal fatigue in singers.

**Conclusion**

The study documented the differences in the cepstral and spectral parameters of voice in Indian heavy metal vocalists before and after vocal fatigue. The study also investigated the strength of correlation between the phonatory habits and the cepstral and spectral parameters in postfatigue voice of Indian heavy metal vocalists. Voice recording was collected before (pretest condition) and after the vocal performance (posttest condition) from Indian heavy metal vocalists. The phonation samples were analyzed, and the spectral and cepstral parameters were extracted. The results revealed that the cepstral measures of voice were more effective in identifying the vocal fatigue-related voice changes in heavy metal vocalists. Moderate and strong correlations were noted between the cepstral and spectral parameters of voice in Indian heavy metal vocalists. Future studies relating the subjective and objective measures of vocal fatigue can provide a better insight into the quantifying vocal fatigue in Indian heavy metal vocalists.

**Financial support and sponsorship**

Nil.

**Conflicts of interest**

There are no conflicts of interest.

**References**

Narasimhan and Mahesh: Vocal fatigue in Indian heavy metal vocalists


APPENDIX

Appendix 1: Questionnaire used in the study

<table>
<thead>
<tr>
<th>Serial Number</th>
<th>Questionnaire</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Do you any complaints or history of voice problems? If yes, please describe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Do you find any change in speaking voice, e.g., loss of strength, hoarse, rattle, breathy?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Do you feel like there is a lump in the throat?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Have you experienced change in voice that has lasted longer than 2 weeks?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>Do you experience tension/stress while talking?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Do you experience vocal fatigue - feeling tired or aches during or after continuous voice use?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Do you have to indulge in talking for extended periods of time during work? If yes, please describe</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>Are you exposed to dusty environment? If yes, for how long?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Are you exposed to chemical fumes/substances?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Do you spend more time in air conditioning?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Do you have any complaint or history of respiratory problem, gastroesophageal reflux/gastritis, hormonal problem, hearing problem, neurological problem</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Have you previously consulted and/or received treatment, surgery, or voice therapy from any health-care professional for your voice?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>How long (h) do you use your voice during the day for singing?</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1–2 h</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>2–4 h</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4–6 h</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>&gt;6 h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>Do you often raise your voice (yell, scream, shout) during singing?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>Do you carry out warm-up and cooldown exercises before and after singing session?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>Do you drink less than six cups or less than 1.5 L of water during the day?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Do you regularly use any kind of medication that negatively affects your voice?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Do you indulge in smoking?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Do you indulge in alcohol consumption?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Do you find your voice to be weak or loss of voice, after using singing for a long time?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>