Correlation of Perceived Nasality with the Acoustic Measures (One Third Octave Spectral Analysis & Voice Low Tone to High Tone Ratio)

Navya Attuluri* and Pushpavathi M

AIISH, Mysore, India

Submission: June 01, 2017; Published: June 06, 2017

*Corresponding author: Attuluri Navya, AIISH, Mysore, 48, First Floor, Maruthi temple road, TK. Layout, Kuvempunagar, Mysore, India, Tel: 9036377964; Email: navyaaslp@gmail.com

Abstract

Aim: The perceived hypernasality in speech can be evaluated using various qualitative and quantitative methods. The present study aimed to investigate, compare and correlate acoustic parameters one third octave spectra analysis and Voice Low Tone to High Tone Ratio (VLHR) with perceived nasality in children with repaired cleft lip and palate (RCLP) and in age and gender matched typically developing children (TDC).

Methods: The study included 73 children (47 RCLP & 26 TDC) in the age range of four to twelve years. The spontaneous speech and sentences were recorded and analyzed for nasality using standardized perceptual four point rating scale. Based on the severity of perceived nasality children were divided into three groups namely normal, mild, and moderate to severe. The production of the vowel /a/ by participants was subjected to acoustic analysis using VLHR and One third octave spectra analysis using MATLAB software.

Results: The results indicated significant differences in spectral amplitude measured using one third octave spectra analysis at high frequency region for the vowel /a/ between TDC & RCLP. The VLHR measures were not indicating statistically significant difference and relation with the perceived nasality.

Conclusion: One third octave spectra analysis is an effective measure in differentiating the nasality and found to have a significant correlation with perceived nasality in children with RCLP. Hence, one third octave spectra analysis can augment the perceptual evaluation to provide additional information to arrive at a diagnosis.

Keywords: 1/3rd octave analysis; VLHR; Cleft lip and palate; Hypernasality; Perceptual analysis

Abbreviations: VLHR: Voice Low Tone to High Tone Ratio; CLP: cleft lip and palate; RCLP: repaired cleft lip and palate; TDC: typically developing children; LFP: low frequency power; HPF: high frequency power; SSPS: Statistical Package for Social Science; K-S: Kolmogorov-Smirnov;

Introduction

Cleft lip and palate is a congenital disorder resulting in the abnormal closure of the lip and palate. The speech of individuals with cleft lip and palate (CLP) is dominated by the presence of hypernasality. The development of Nasometer by Fletcher and Bishop [1] was a major advancement that facilitated the objective analysis of nasality in speech. Nasometer gives the parameter nasalance which is the ratio of the proportion of the nasal energy in speech to the proportion of nasal and oral energy i.e. Nasalance = nasal/ (nasal+oral) X 100% [2]. Nasometer is extensively used to evaluate the nasality in the speech of individuals with cleft lip and palate.

Hence clinicians often use Nasometer to objectively measure the percentage of the nasal component in the speech of individuals with CLP. Attempts have been made to correlate the nasalance values with the judgments of perceived nasality as perceptual judgment remains the gold standard [3]. The researchers reported the various degree of correlation of nasalance with the perceived nasality [4,5].

The discrepancies among the studies on the correlation of speech based on nasalance with perceived nasality resulted in further investigations to explore the components of speech related to the perceived nasality. The acoustic analysis of speech indicated increased peak amplitudes around first formant region in the speech of individuals with repaired CLP [6-8]. Kataoka [9] investigated the variations in spectral amplitude at one third octave frequency bands. This particular bandwidth was selected...
because it is similar to the critical band of frequencies used by our ear in analyzing speech [10].

Similar attempts were made to develop a quantitative index by evaluating the voice spectrum to analyse the effect of nasal obstruction by Lee, Yang, and Kuo [11]. VLHR is the ratio of low frequency power (LPF) to high frequency power (HPF) of the sound power spectrum. The power is expressed in decibels [11]. The cut-off frequency to divide high and low frequencies was calculated by multiplying fundamental frequency (F₀) with the square root of (4x5). Lee, Wang, Yang, and Kuo [11] conducted a study to measure VLHR with a cut off frequency of 600Hz. They indicated higher VLHR in speech samples of children with hypernasality and found a significant positive correlation (r = 0.76, P < 0.01) of VLHR with nasalance scores.

Objective measures of speech have always been remained important in augmenting the perceptual evaluation. However, the use of these methods is limited due to lack of adequate published data using appropriate research designs on the sensitivity and specificity of these measures. The easy to use objective measures need to be developed and validated for providing effective clinical and empirical practice. The VLHR and one third octave spectra analysis are new advancements in objective assessment of nasality. This could be an effective alternate or complement the traditional nasalance analysis. Therefore, the present study was taken up with the aim of evaluating the correlation of one third octave spectral analysis and VLHR with the perceived nasality in children with repaired cleft lip and palate.

**Objectives of the study**

a) To perceptually evaluate hypernasality in the speech of children with repaired cleft lip and palate (RCLP) using a standardized rating scale.

b) To investigate and compare the One Third Octave Spectra Analysis and VLHR in children with RCLP and typically developing age and gender matched children.

c) To correlate measures of One Third Octave Spectra Analysis and VLHR with the perceived nasality.

**Method**

The present study considered 73 Kannada speaking children in the age range of four years seven months to twelve years. Among these, forty seven children had misarticulations with RCLP and twenty six were typically developing children. The demographic details are indicated in Table 1.

**Table 1:** Details of the participants included in the study.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Age range</th>
<th>Mean Age</th>
<th>Gender</th>
<th>Severity of Nasality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group I a</td>
<td>4.2 to 12</td>
<td>8.01</td>
<td>Male</td>
<td>Mild</td>
</tr>
<tr>
<td>Group I b</td>
<td>5 to 11</td>
<td>9.4</td>
<td>Female</td>
<td>Moderate to severe</td>
</tr>
<tr>
<td>Group II</td>
<td>6.7 to 12</td>
<td>9.9</td>
<td></td>
<td>Normal nasality</td>
</tr>
</tbody>
</table>

**Participants Selection Criteria**

i. **Inclusion Criteria group I (children with RCLP):** The children with RCLP, cleft palate and repaired soft palate aged between 4 to 12 years with Kannada as their native language. Children with normal cognitive, hearing and mental abilities were considered.

ii. **Inclusion Criteria for Group II (Typically developing children):** Children who passed informal screening for speech and hearing disorders. Children ruled out for disability by administering World Health Organization (WHO) checklist [12].

**Procedure**

**Perceptual Analysis**

A standardized perceptual rating scale developed by Henningsson, Kuehn, Sell, Sweeney, Trost-Cardamone, and Whitehill [13] was used to evaluate perceived hypernasality. The perceptual rating classifies the data onto a 4-point rating scale ranging from 0 through 3, where 0 = within normal limits (WNL), 1 = mild, 2 = moderate, 3 = severe that reflects increasing severity of hypernasality. Three qualified experienced speech language pathologists were considered as judges.

The stimulus used for perceptual evaluation of nasality was audio and video recordings of participant’s spontaneous speech sample recorded with Sony handycam with bearing model no. DCR-SR88. The sample consisted of spontaneous speech (on self-introduction, school, leisure activities and picture description) for the duration of three to five minutes and repetition of oronasal and oral sentences in Kannada language [14]. The judges rated the samples for the severity of the nasality perceived based on standardized perceptual rating scale by Henningsson et al. [13]. The participants rated by the judges as mild were considered as Group Ia. The participants rated as moderate and severe were together considered as Group Ib.

**Acoustic Measures**

**Instructions and Recording:** The participants were demonstrated and instructed to phonate steady state vowel /a/ thrice at a comfortable pitch and loudness in to omni-directional distortion free 1 BALL microphone. The one third octave spectral measures and VLHR were obtained on the middle 500 millisecond section of the sustained phonation of vowel /a/. The PRAAT 8.1 Version was used to select the steady state portion of the vowel and saved for further analysis using MATLAB 7.0 version software.
One-third Octave Spectra Analysis: The speech stimuli were analyzed in 23 one-third octave bands (over a frequency range of 100–16,000 Hz) using a digital filter that was designed to match the ANSI standard (ANSI S1.11-1986). One-third octave spectra analysis was calculated for frequency bands between 100-16,000 Hz on all samples (/æ/; /ɛ/; /ɜ/; /ɒ/). The frequency bands considered for analysis were 39Hz, 50Hz, 630Hz, 793Hz, 1000Hz, 1259Hz, 1587Hz, 2000Hz, 2519Hz, 3174Hz, and 4000Hz.

Voice Low Tone to High Tone Ratio (VLHR): The low-frequency power section (LPF) is defined as the summation of the power from 50 Hz to 600Hz and high frequency power section (HPF) also can be expressed as addition of power from 600Hz to 8063Hz. Voice low tone to high tone ratio was obtained from 10 x log10 (LPF/HPF) [15].

Statistical Analysis

The data obtained by all these measures was subjected to appropriate statistical analysis using Statistical Package for Social Science, Version 17.0 (SPSS). The normality of the data across the groups was analyzed using Kolmogorov-Smirnov (K-S) test. Multivariate analysis (MANOVA) was administered to differentiate the three groups across all the objective measures (VLHR and one third octave spectra analysis). Post hoc multiple comparison was carried out using Duncan’s test followed by MANOVA. Descriptive statistics was used to group the data based on perceptual rating assigned. Cronbach’s Alpha coefficient and Spearman rank order correlation were used to analyze the reliability and correlation respectively.

Results and Discussion

Perceptual evaluation of hypernasality

Table 2: Degree of hypernasality in participants across the stimuli.

<table>
<thead>
<tr>
<th>Degree of Hypernasality</th>
<th>SS</th>
<th>OS</th>
<th>ONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>01</td>
<td>02</td>
<td>01</td>
</tr>
<tr>
<td>Mild</td>
<td>23</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>Moderate</td>
<td>15</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>Severe</td>
<td>08</td>
<td>07</td>
<td>07</td>
</tr>
</tbody>
</table>

The perceived hypernasality was rated based on a four point rating scale for spontaneous speech, oral and oronasal sentences for forty seven children with RCLP. Using rating scale the stimuli were rated by three judges and a consensus agreement by any two out of three judges on a stimulus parameter was obtained to group the participants based on the stimuli. Twenty three were rated as mild hypernasal, sixteen were rated as moderate hypernasal, and eight were rated as severe hypernasal. Table 2 depicts the distribution of the participants based on the nasality across the stimulus.

Inter and intra judge reliability measures

The inter judge reliability of perceptual evaluation of hypernasality was performed using Cronbach’s alpha coefficient for the entire sample across the groups with RCLP. The coefficients are 0.72, 0.79 and 0.83 for spontaneous speech, oral sentences and oronasal sentences respectively. The intra judge reliability of perceptual evaluation of hypernasality was performed for 25% of the entire sample across the groups. The intra judge reliability of perceptual evaluation of hypernasality was 0.72 to 0.92 across the stimuli as shown in Table 3. High reliability ratings were obtained for oral sentences and similar alpha coefficients were obtained for oronasal and spontaneous speech.

Table 3: Intra - judge reliability measures for perceptual evaluations.

<table>
<thead>
<tr>
<th>Judges</th>
<th>SS</th>
<th>OS</th>
<th>ONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>J1</td>
<td>0.92</td>
<td>0.88</td>
<td>0.72</td>
</tr>
<tr>
<td>J2</td>
<td>0.78</td>
<td>0.80</td>
<td>0.72</td>
</tr>
<tr>
<td>J3</td>
<td>0.87</td>
<td>0.83</td>
<td>0.79</td>
</tr>
</tbody>
</table>


Discussion

In clinical investigations of hypernasality, perceptual evaluation is considered as gold standard along with the objective measures [3]. Based on perceptual judgment, the children were divided in to normal, mild and moderate to severe hypernasal groups. The intra and inter judge reliability measures denote the relationship between all the ratings given by various judges [16]. In the present study, inter - judge reliability of perceptual rating was almost similar for oral (0.79) and oronasal (0.83) sentences followed by spontaneous speech (0.72). The results indicated low reliability for spontaneous speech than oral and oronasal sentences. This can be due to high fluctuations in acoustic properties of frequency and amplitude during spontaneous speech. The constant change in these properties leads to difficulty to judge the nasality in speech. Similar reliability ratings across the stimulus were reported by Vogel et al. [17] who indicated good inter rater agreement between the judges with an overall score of 0.78 for various oral and nasal passages. The differences in reliability measures with the stimulus variations were also documented by Watterson et al. [18].

The intra - judge reliability for spontaneous speech, oral sentences, and oronasal sentences ranged from 0.78 to 0.92, 0.80 to 0.88, and 0.72 to 0.79 respectively. The variations in intra judge reliability can be attributed to the difficulty to rate standard speech sample, for a given speaker, as the listener will determine different degrees of hypernasality for virtuous points in the rating scale, as this is an arbitrary image, there will be individual variations. The findings of Vogel et al. [17] indicated
intra rater reliability was ranging from 0.66 to 0.91 for passages with varying proportion of nasal phonemes. The reduced reliability ratings can be attributed to the difficulty in judging hypernasality as speech is a multidimensional task [18]. Another study by Tsai [19] reported intra rater reliability of two judges for the spontaneous speech was 0.74 and 0.90 and inter-rater reliability was 0.91. The differences across the studies were attributed to the methodological variations. The ratings of the study by Tsai [19] were based on two judges and the rating scale used was visual analogue scale ranging from 0mm indicating “no nasal resonance” to 100mm representing “the most nasal resonance”.

The intra judge reliability measures in the present study are similar to findings of Kataoka et al. [20] who reported intra judge reliability of hypernasality ratings by experienced listeners and graduate students ranged from 0.77 to 0.88 and 0.70 to 0.89 respectively. These results are in close approximation with the present study and this can be due to the similarity in the procedure followed for perceptual evaluation using an equal appearing interval rating scale.

One third octave spectra analysis of vowel /a/

A. Mean of one third octave spectra analysis vowel /a/ across the groups: Overall the groups exhibited increased energy concentration at 1000Hz, 1259Hz, and 1587Hz than in the other frequency regions computed. However, in general, minimal differences in energy concentration across the groups observed for all the frequencies. However, the increased spectral amplitude at mid and high frequencies in TDC than RCLP except for 2000Hz and 2519Hz as shown in Figure 1. MANOVA results indicated significant differences in one third octave spectral measures for the vowel /a/ in 1000 Hz, 1587 Hz, and 4000Hz across the groups at p > 0.05 level of significance. The post hoc analysis revealed significant differences in energy concentration at 1000 Hz and 1587 Hz for the vowel /a/ between mild hypernasal and TDC.

B. Correlating the one third octave spectra analysis of /a/ with perceived nasality: The correlation coefficients were significant at p < 0.05 for vowel /a/ at 3174Hz (-0.23) and 4000Hz -0.30). The significant correlation coefficients indicate that the perception of nasality in the speech can have a moderate influence on the measures of one third octave spectral energy.

VLHR measures for vowel /a/

Table 4: Mean and SD of VLHR.

<table>
<thead>
<tr>
<th>Participants</th>
<th>Amplitude (dB) of /a/</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>Group I a</td>
<td>7.21</td>
</tr>
<tr>
<td>Group I b</td>
<td>6.13</td>
</tr>
<tr>
<td>Group II</td>
<td>5.44</td>
</tr>
</tbody>
</table>

A. The mean of VLHR for vowels /a/ across the groups: Table 4 depicts the results for mean and standard deviation of VLHR for the vowel /a/ across the groups. The increased VLHR measures were observed for children with RCLP (Ia & Ib) than TDC. The increased VLHR is exhibited by mild hypernasal followed by moderate to severe hypernasal and control group for VLHR of /a/. MANOVA was done to find the statistically significant differences across the three groups over the dependent variables. The MANOVA results indicated no significant differences in VLHR measures across the groups /a/ = {F (2,70) = 2.82, p > 0.05}.

B. Correlating the VLHR measures with perceived nasality: The relation between perceived nasality exhibited by children with RCLP and TDC with VLHR measures for the vowel /a/ were evaluated using Spearman Correlation Coefficient. The results indicated correlation coefficient of 0.084 (p > 0.05). The VLHR measures were not significantly correlated with perceived nasality. This indicates a poor correlation between VLHR and perceived nasality.

Discussion

Spectral analysis was used to explore the acoustic properties of speech in individuals with cleft lip and palate [21]. In the present study, acoustic analysis of speech using one-third octave spectra analysis and VLHR measures were investigated on children with RCLP and control group. The review of literature indicated that nasalization cannot be measured accurately by using formant analysis alone, specifically in the presence of high fundamental frequency [20]. The shape of the entire region of the spectral envelope is important for vowel perception rather than the frequency and amplitude of the spectral peaks. Therefore, 1/3rd octave spectral analysis evaluates overall spectral envelope to have a theoretical advantage in analyzing hypernasal vowels.

Another advantage of using one third octave spectra analysis is that the 1/3rd octave bandwidth matches with the...
Critical band analyzed by ear for the perception of speech [10]. Hence, it can be postulated to correlate well with the perceptual analysis of nasalization. One third octave spectral energy across the frequencies ranging from 100 Hz to 16000 Hz was calculated in the present study. However, literature reported [22] that significant change in spectral amplitude of speech with hypernasality was evident in the frequency bands from 396 Hz to 4000 Hz. Hence, in the present study we have restricted evaluation of spectral amplitudes between 396 Hz to 4000 Hz for the final analysis.

The results of the present study indicated diversifying outcomes across the stimulus for all the groups. There was no specific trend observed to conclude on the effect of hypernasality in the spectral amplitude of vowel /a/. Among the groups, the differences in spectral amplitudes of TDC were relatively low in children with mild hypernasality than moderate to severe hypernasality. There is an increase in the spectral amplitude around low and mid frequencies and reduction in high frequencies with the increase in perceived nasality. This can be attributed to the increase in the velopharyngeal gap which can lead to increased perception of nasality. The increase in the cross sectional area of the velopharyngeal opening can lead to shift in the frequency of the first formant and increased formant bandwidth. The energy concentrated at particular frequencies are indicated as formants. In the first formant region, a pole zero pair is added and the gap between the pole and zero increases with respect to velopharyngeal gap. In this gap, an additional pole is added indicating spectral prominence with the increased VP gap. The results are in agreement with the findings of Vogel et al. [17] who reported higher spectral amplitude at low and mid frequency bands from 476 Hz to 1200 Hz in hypernasal speakers. They also stated that the significant differences in one third octave spectra analysis were only found across severe hypernasal and control groups.

The additional spectral peaks around first formant (F1) were only noticed in moderate to severe hypernasal group and the absence of these peaks indicate reduced hypernasality in the mild hypernasal group. These researchers reported that participants with hypernasal speech were exhibiting increased spectral amplitude between first and second formants around 1 kHz and decreased between second and third formants.

The spectral change over the duration of the vowel was considered as the coexisting speech characteristics that influenced the percentage of hypernasality perceived. Hence, another acoustic measure based on spectral energy was considered for the investigation in VLHR. The results of the present study also indicated increased VLHR measure for vowel /a/ in children with RCLP than control group. The reduced VLHR in control group can be attributed to increased spectral energy at high frequency regions than hypernasal speakers, due to the presence of anti formats toward high frequency regions in hypernasal speech. The reduced spectral energy between F2 and F3 was also reported by Yoshida et al. [23] and Vogel et al. [17]. A study done by Lee et al. [24] indicated decreased high frequency energy (anti-resonance) than low frequencies for nasal voices differentiating significantly from the acoustic characteristics of speech of healthy controls. However, in the present study, the difference in spectral amplitude across the groups was not statistically significant.

The correlation analysis in the present study also indicated no significant relation between the VLHR measures with the perceived nasality. The VLHR measures are based on the sum of the amplitudes in the spectrum. The spectral amplitudes can also be attributed to variations in frequency domain characteristics of voice in nasalized speech, such as a reduction in the intensity of the first formant, the presence of extra resonances, and increased bandwidth of formants [25]. The formants can vary with respect to the position of articulators, particularly tongue [26]. The results are in accordance with the findings of Vogel et al. [17] who also reported no significant differences in the VLHR measures in the children with hypernasality and typically developing children in the age range of 4 years to 12 years, using cut off frequency 600 Hz. In the contrary, few studies measuring VLHR in the adult population with hyponasality [11] showed a significant difference between hyponasal and control groups. The differences in the studies can be attributed to the methodological differences with respect to the subject selection, cut off frequency and the procedure of measuring VLHR [27].

Conclusion

The present study evaluated and correlated measures of hypernasality in children with RCLP and TDC based on perceptual rating scale and objective measures (one third octave spectra analysis & VLHR). In one third octave spectra analysis, the children with hypernasality exhibited significantly less spectral energy at high frequencies than the control group. The increased VLHR was observed in children with hypernasal groups than the control group, however, there were no significant differences noticed across the groups. These are easy to use objective measures which can augment the perceptual evaluation along with other quantitative measures for diagnosis and to evaluate the efficacy of various treatment techniques.

References


