INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

Spectral Enhancement of Cleft Lip and Palate Speech

Vikram C. M, Nagaraj Adiga, and S. R. Mahadeva Prasanna

Departmen of Electronics and Electrical Engineering
Indian Institute of Technology Guwahati, Guwahati-781039, India
{cmvikram, nagaraj, prasanna}@iitg.ernet.in

Abstract

The quality of cleft lip and palate (CLP) speech is affected due to hyper-nasality and mis-articulation. Surgery and speech therapy are required to correct the structural and functional defects of CLP, which will result in an enhanced speech signal. The quality of the enhanced speech is perceptually evaluated by speech-language pathologists and results are highly biased. In this work, a signal processing based two stage speech enhancement method is proposed to get the perceptual benchmark to compare the signal after the surgery / therapy. In the first stage, CLP speech is enhanced by suppressing the nasal formant and in the second stage, spectral peak-valley enhancement is carried out to reduce the hyper-nasality associated with the CLP speech. The evaluation results show that the perceptual quality of CLP speech signal is improved after enhancement in both stages. Further, the improvement in the quality of the enhanced signal is compared with the speech signal after palatal prosthesis / surgery. The perceptual evaluation results show that the enhanced speech signals are better than the speech after prosthesis / surgery.

Index Terms: Cleft Lip and Palate (CLP) speech, formant enhancement, hyper-nasality, speech therapy.

1. Introduction

Cleft lip and palate (CLP) is a craniofacial abnormality and a congenital disorder. The structural correction of CLP may not lead to functional correction of velo-pharyngeal valve. The presence of an oro-nasal fistula in hard / soft palate may also lead to oro-nasal coupling [1]. Hence, the presence of cleft in the palate or velo-pharyngeal dysfunction (VPD) or oro-nasal fistula or a combination of all these in CLP may lead to hypernasality and mis-articulation, which will result in unintelligible speech. Improvement in speech quality can be achieved by clinical methods such as surgery, application of aids such as palatal prosthesis, and speech therapy [2, 3, 4]. Improvement in speech quality, after surgery / therapy is often evaluated by expert speech-language pathologists (SLPs), which may give biased assessment results. However, selection of parameters and development of rating scale are the challenging issue involved in perceptual evaluation [5, 6]. Speech therapy is a long-term process, the same SLP may not be available to evaluate the improvement in speech quality for every-time. In spite of all these drawbacks, the perceptual evaluation is most commonly used by SLPs because of its simplicity. To improve the perceptual evaluation results, a perceptual benchmark is required to evaluate the speech quality after surgery / therapy.

The presence of oro-nasal coupling in CLP enables the passage of periodic glottal flow through nasal cavity which will result in hyper-nasal speech. Hence, voiced sounds, especially vowels are rich in hyper-nasality information. As the glottal flow passes through the nasal cavity, nasal formants around 250 Hz (P_0) and 1000 Hz (P_1) are added with vowel formants [7, 8, 9]. In addition to P_0 and P_1, a decrease in the strength of formants also observed. Increase in the level of dip or valley between F_2 and F_3 is also considered as an important cue for the nasality in [9, 10]. Due to the effect of addition of nasal formants and decrease in the formant strength, an increase in the spectral flatness over mid band frequencies i.e. 1000-2500 Hz is observed [11, 12, 13].

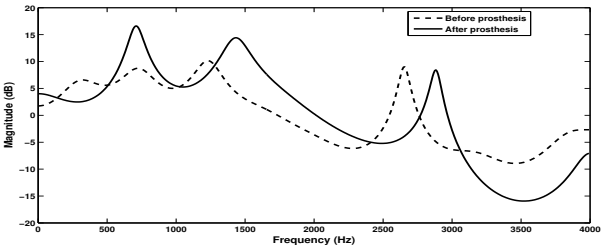

Figure 1: Linear prediction (LP) spectra of CLP speech before and after palatal prosthesis application. The figure shows that after the application of palatal prosthesis, nasal formant around 250 Hz is suppressed and peak-to-valley ratio get increased.

The resonant structure of CLP speech can be corrected by clinical methods such as surgery, palatal prosthesis, and speech therapy. Fig. 1 shows the linear prediction (LP) spectra of a segment of vowel /a/ computed before and after the application palatal prosthesis. After the application of prosthesis, there is a significant improvement in the resonant structure in terms of suppression of nasal formant around 250 Hz and increase in the formant strength is observed. Such improvement in the resonant structure from clinical methods (Fig. 1) can also be achieved by modifying the speech spectrum using signal processing methods. Suppression of formant peaks and decrease in the spectral peak-to-valley contrast is also observed in speech degradation under noisy environments and text-to-speech synthesizers. As a solution, the formant enhancement in terms of increasing formant amplitude and sharpening of formant peaks is carried out to enhance the degraded speech quality [14]. The extra nasal formant can be detected by group delay and LP methods [15, 16]. Motivated by the improvement of resonance structure achieved from clinical methods, the paper aims to develop a signal processing based algorithm for CLP speech enhancement. In particular, removal of nasal formants, enhancement of formant peaks, and suppression of valleys of the spectrum is carried out.

Further, the paper is organized as follows, section 2 describes the analysis of CLP speech. The enhancement procedure for the CLP speech is discussed in section 3. The evaluations and the clinical application of proposed method is dis-

http://dx.doi.org/10.21437/Interspeech.2016-842

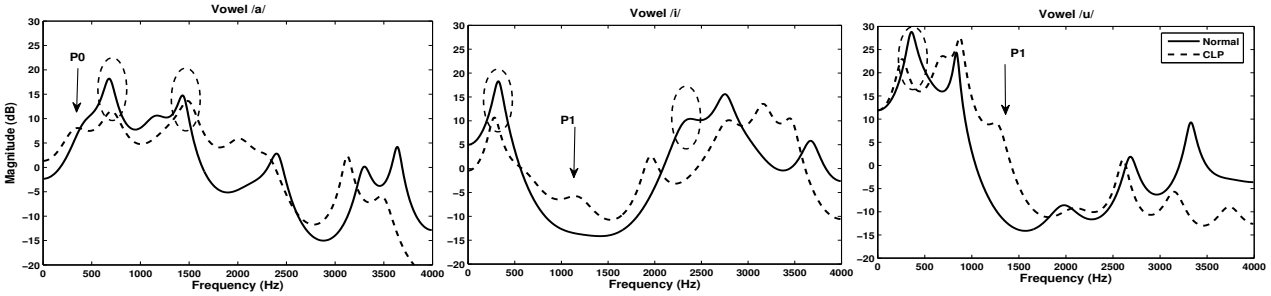

Figure 2: LP spectra of vowels $/a/$, $/i/$, and $/u/$ for normal and CLP speech. The circled parts in the figure shows the reduction in peak-valley ratio in CLP. The nasal formants P_0 and P_1 are well evident in /a/ and /i/, respectively.

cussed in section 4. Finally, section 5 summarizes the work and gives the scope for future work.

2. Analysis of CLP speech

2.1. Database

CLP speech samples are obtained from All Indian Institute of Speech and Hearing (AIISH), Mysore, India [17]. In this work, speech samples from Kannada speaking persons of age range 18-30 are recorded under clean room conditions (Kannada is a Dravidian language, spoken in the southern part of India). Description of the database used in this work is given in Table 1. The word and sentence level stimuli specified by SLPs, which are rich in vowels, stops, and fricatives are considered for recording. The database consists of both, word and sentence level recordings from 10 normal and 21 CLP subjects. Among 21 CLP subjects, 10 repaired, 10 unrepaired, 2 pre and post palatal prosthesis, and 1 pre and post primary palatal surgery are considered. All CLP speech samples in the current database are perceptually evaluated and presence of hyper-nasality is reported.

Table 1: DATABASE

Sl. No	Catagory	No. of subjects	No. of speech samples
1	Normal	10	20
2	Repaired	10	20
3	Unrepaired	08	16
4	Palatal prosthesis	02	20 (pre) and 20 (post)
5	Surgery	01	10 (pre) and 10 (post)

2.2. Spectral Analysis

The LP based spectral analysis of normal and CLP speech are carried out to analyze the effect of nasal coupling [18]. In conventional speech analysis, for an 8 kHz sampled speech signal, LP order order of 8-12 is preferred to capture the F_1, F_2, F_3, and F_4 information. However, as mentioned in [8], for nasalized speech, in order to capture P_0 and P_1 information, LP spectrum of order 16 (for $f_s = 8\ kHz$) is computed. A 25 ms segment taken from the sustained portion of vowel is used to compute the LP spectrum. The Fig. 2 shows the LP magnitude spectra computed for the vowels /a/, /i/, and /u/ (low, mid, and high vowels) of normal and CLP speech. Presence of an extra nasal formant (P_0) near 250 Hz in /a/ and extra nasal formant (P_1) near 1000 Hz in /i/ and /u/ is noticed in Fig. 2. Reduction in amplitude of formants and increase in the spectral flatness is also observed.

The formant peaks and valleys between the peaks are severely affected due to nasal coupling. So the resonance structure is further analyzed in-terms of peak-to-valley ratio. Let A_1 and A_2 be amplitude of F_1 and F_2, respectively. V_1 and V_2 be the amplitude of valley between F_1 and F_2, F_2 and F_3, respectively. The peak-to-valley ratios computed for the vowel /a/ of normal and CLP speech samples from the current database, which are as shown in Fig. 3. In order to avoid the detection of P_0 as F_1, only peaks above 300 Hz are considered (for vowel /a/ F_1 >600 Hz). For normal speech, the ratio of A_1 to V_1 is denoted as $AN_1 - VN_1$, A_2 to V_1 as $AN_2 - VN_1$, and A_2 to V_2 as $AN_2 - VN_2$. Similarly, for CLP speech the peak-to-valley ratios are denoted as $AC_1 - VC_1$, $AC_2 - VC_1$, and $AC_2 - VC_2$. From Fig. 3, it can be noticed that compared to normal speech, peak-to-valley ratio decreases in CLP speech.

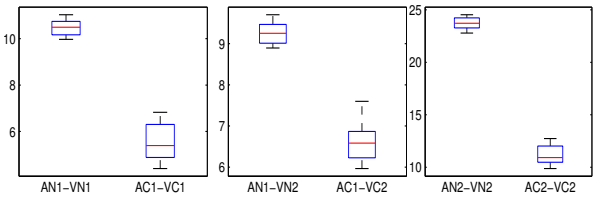

Figure 3: Peak to valley ratios for normal and CLP speech.

The formant structure (F_1, F_2, and F_3) of normal /a/ is (710 Hz, 1100 Hz, and 2540 Hz), /i/ is (280 Hz, 2250 Hz, and 2890 Hz) and /u/ is (310 Hz, 870 Hz, and 2250 Hz). Since, in vowel /u/ P_0 and F_1, P_1, and F_2 are very close together, it is difficult to analyze by LP method. As mentioned in the literature [8, 16], P_0 in /a/ and P_1 in /i/ are highly contribute for the nasality. The F_1 and P_0 in /a/, P_1 and F_1 in /i/ are well separated, LP method can be used to model them. Hence, the utterances containing only vowels /a/ and /i/ are considered for the enhancement.

3. Spectral enhancement of CLP Speech

The LP analysis based spectral enhancement is widely used in speech enhancement applications [14]. Sharpening of formant peaks by increasing the amplitude of format peaks and decreasing the amplitude of valleys are carried out. Similar to the enhancement of degraded speech, enhancement of CLP speech also requires spectral enhancement. Hence, in this work LP based spectral enhancement method is proposed. The CLP speech enhancement algorithm consists of

1. Removal of nasal formants P_0 in /a/ and P_1 in /i/.
2. Spectral enhancement by peak enhancement and valley suppression.

The procedure of CLP speech enhancement is explained in Algorithm 1. The enhancement process is carried out on the LP

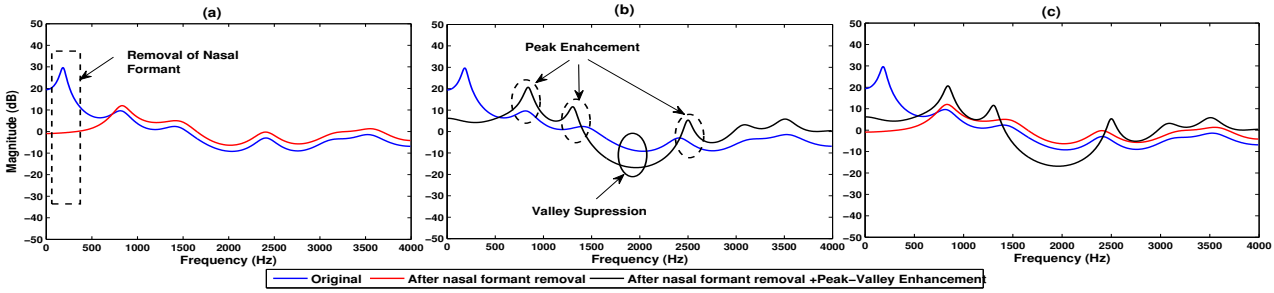

Figure 4: Two stage CLP speech enhancement process. (a) After nasal formant removal (stage-1), (b) After nasal formant removal and formant enhancement (stage-2), and (c) Combination of 2 stages

Algorithm 1 Spectral Enhancement

Step 1: Compute LP coefficients with LP order N=16.

$$\boldsymbol{H(z) = \frac{1}{\prod_{i=0}^{N-1}(1 - a_i z^{-1})}} \quad (1)$$

Where, $\boldsymbol{a_i}$ for $\boldsymbol{i = 0, 1, 2, .., N-1}$ are the roots of LP polynomial. Step 2: Compute LP residual.
Step 3: Removal of nasal formant

if $\boldsymbol{f_{a_1}}$ < 350 Hz **then**
  if $\boldsymbol{f_{a_2}}$ > 600 Hz & $\boldsymbol{f_{a_3}}$ >1800 Hz **then**
    Remove $\boldsymbol{a_i}$ near 1000 Hz
    Go to Step 4
  else
    Remove $\boldsymbol{a_i}$ near 250 Hz
    Go to Step 4
  end if
else
  Go to Step 5
end if

Step 4: Estimate LP spectrum $\boldsymbol{S(f)}$
Step 5: Locate formants $\boldsymbol{F_1, F_2, F_3,}$ and $\boldsymbol{F_4}$ and compute $\boldsymbol{A_1, A_2, A_3,}$ and $\boldsymbol{A_4}$
Step 6: Locate valleys $\boldsymbol{D_1, D_2,}$ and $\boldsymbol{D_3,}$ and compute $\boldsymbol{V_1, V_2,}$ and $\boldsymbol{V_3}$
Step 7: Perform peak and valley enhancement for $\boldsymbol{S(f)}$

for i=1 to 4 **do**
  for $\boldsymbol{f_i = F_i - \delta f}$ to $\boldsymbol{F_i + \delta f}$ **do**
    $\boldsymbol{S^{'}(f_i) = p_i * S(f_i)}$
  end for
end for
for i=1 to 3 **do**
  for $\boldsymbol{f_i = D_i - \delta f}$ to $\boldsymbol{D_i + \delta f}$ **do**
    $\boldsymbol{S^{'}(f_i) = v_i * S(f_i)}$
  end for
end for

Step 8: Recompute the LP coefficients for $\boldsymbol{S^{'}(f)}$
Step 9: Synthesize the speech signal (stage-1 enhanced signal) using LPCs of step 3 and residual of step 2
Step 10: Synthesize the speech signal (stage-2 enhanced signal) using LPCs of step 8 and residual of step 2

spectrum computed for the LP order of 16. The removal of nasal formant stage consists of identification and removal of poles corresponding to nasal formants, P_0 in /a/ and P_1 in /i/. The speech enhancement is carried out on the samples consists of vowels /a/ and /i/, without using any phone level transcriptions. Simply removal of formants near 250 and 1000 Hz may result in undesired modifications in the resonant structure of vowels. For a given nasalized /a/ formant structure resulted from LP analysis consists of P_0, F_1, F_2, F_3 , and F_4 and that for /i/ is F_1, P_1, F_2, F_3, and F_4. Removal of formant near 250 Hz in /i/ and 1000 Hz in /a/ may result in the elimination of F_1 in /i/ and F_2 in /a/, which is not desired. Hence, it is necessary to decide, whether the given pole corresponding to nasal or vowel. In the Algorithm 1, step-3 represents the procedure to select the pole corresponding to nasal formant. The LP spectrum ($S(f)$) is estimated from LPCs resulted after the elimination of pole corresponding to nasal formant. Fig. 4(a) shows LP spectra before and after removal of nasal format. Further, spectral enhancement is carried out using steps from 5 to 7. The peak picking procedure is used to locate the formants F_1, F_2, F_3, and F_4 with amplitudes A_1, A_2, A_3, and A_4. Similarly, D_1, D_2, and D_3 represent the valleys between the formants $F_1 - F_2$, $F_2 - F_3$, and $F_3 - F_4$ with amplitudes V_1, V_2, and V_3, are located. The formant peak scaling factor p_i for i=1, 2, 3, and 4 is computed as,

$$p_i = A_{N_i}/A_{CLP_i} \quad (2)$$

where, A_{N_i} and A_{CLP_i} are the average amplitudes of i^{th} formant computed from LP spectra of normal and CLP speakers respectively. Similarly valley suppression factor v_i for i=1,2, and 3 is computed as,

$$v_i = V_{N_i}/V_{CLP_i} \quad (3)$$

where, V_{N_i} and V_{CLP_i} are the average amplitudes of i^{th} valley computed from LP spectra of utterances of normal and CLP speakers, respectively. Utterances of normal and CLP subjects present in the entire database are used to compute scaling factor. A very small frequency components of δ Hz around detected peaks and valleys are multiplied by computed scaling factors to get enhanced spectrum $S^{'}(f)$. Fig. 4(b) shows the spectrum after peak-to-valley enhancement. The LPCs are re-computed for $S^{'}(f)$. Fig. 4(c) shows the spectra before enhancement, stage-1, and stage-2 enhancement. Synthesized speech only by the removal nasal formants is referred as stage-1 enhanced speech. Signal after the removal of nasal formants and formant enhancement is referred as stage-2 enhanced speech. In this work, the formant enhancement refers to the spectral peak enhancement and valley suppression. The LP based speech enhancement algorithm is applied for every frame of 25 ms with a shift of 5 ms to get stage-1 and stage-2 enhanced speech.

4. Evaluations

Improvement in CLP speech quality after enhancement is evaluated using both subjective and objective methods. First, the quality of stage-1 and stage-2 enhanced signals are compared with original CLP samples using comparative mean opinion score (CMOS) test. Further, the speech samples, before the application of prosthesis or surgery are enhanced and compared with that of speech samples after prosthesis or surgery using preference and objective tests.

4.1. CMOS test

The CLP samples in the database (both repaired and unrepaired) are passed through the enhancement algorithm. Two different enhanced speech signals are obtained from stage-1 and stage-2. The listeners were given with CLP speech files of before enhancement, stage-1, and stage-2 enhancement. The listeners were instructed to compare the degree of nasality in vowels before and after the enhancement, and rate the enhanced signals using 5-point grading scale. The 5 point scale given by 1-unsatisfactory (before and after enhancement are same), 2-Poor (slightly reduced in nasality), 3-fair (reduction in nasality, but not near to normal), 4-good (reduction in nasality and near to normal), and 5-excellent (complete reduction in nasality and very near to normal). CMOS scores obtained for the stage-1 and stage-2 enhanced signals are as presented in Table 2. The scores for stage-1 enhanced speech indicate the reduction in nasality, but not near to normal speech. The scores of stage-2 enhanced signals show both reduction in nasality and nearness to normal speech. Thus, scores of CMOS test indicate that stage-2 enhanced signals are having better quality than stage-1 enhanced speech.

Table 2: MEAN OPINION SCORES OBTAINED FOR ENHANCED SPEECH SIGNALS AFTER STAGE-1 AND STAGE-2

CMOS stage-1	CMOS stage-2
3.4	4.5

4.2. Comparison with clinical methods

In order to compare the improvement by the proposed work with that of clinical methods, in this work, two clinical methods : palatal surgery and palatal prosthesis are considered. Speech before the application of clinical methods is enhanced by proposed method and preference test is used compare it with the speech after surgery / prosthesis. In preference test (perceptual test), the listener is given with speech files, before and after the prosthesis / surgery, and stage-2 enhanced files. Listener were instructed to listen to a pair of files and indicate the file, which is better in-terms of reduced nasality and near to normal sounds. Preference test scores between before-after clinical procedure (test-1), before clinical procedure-enhanced signal (test-2), and after clinical procedure-enhanced signal (test-3) are considered. Here, enhanced signal refers to speech after stage-2 enhancement. The results of preference test are shown in Table 3. Test-1 results indicate that the therapy is significant and speech samples, after therapy are more preferred. Test-2 results show that enhanced signals got more preference than that of speech before surgery / prosthesis. From the results of the test-3, enhanced speech got high perceptual preference than speech after surgery / prosthesis, which indicates that further treatment is required to improve the speech quality. Hence, the proposed method can be used to set a benchmark to compare the enhanced speech obtained after clinical applications.

The objective evaluation in speech synthesis and enhancement under noisy conditions are conducted using conventional

Table 3: RESULTS OF PREFERENCE TEST.

Test. No	Preference (1 and 2)	1	2
1	Before and after surgery/prosthesis	40%	60%
2	Before surgery / prosthesis and after enhancement	16%	84%
3	After surgery / prosthesis and after enhancement	33%	67%

methods, such as spectral distance measurement between original (before degradation) and enhanced signals. In current work, it is impractical to get the speech before nasalization from CLP subjects. Therefore, we define an objective criterion in-terms of percentage of presence of nasal resonance (PNR). The PNR is computed from LP spectrum, where the LPCs are computed for every frame of size 25 ms with a rate of 5ms. For each frame, nasal peak is detected using the method given in step-3 of Algorithm 1. The process is repeated for entire utterance and PNR is computed as the ratio between number of frames detected with nasal resonance and the total number of frames. In addition to PNR, peak-to-valley ratios, $A_1 - V_1$, $A_2 - V_1$, and $A_2 - V_2$ are also used for objective evaluation. The results of objective test for speech before and after the clinical applications,

and after the enhancement are mentioned in Table. 4. Compared to speech before surgery / therapy, the nasal resonance is fairly reduced and peak-to-valley ratio is significantly increased in speech after therapy. After therapy, the peak-to-valley ratio reaches that of enhanced signal, but the nasal formant not significantly suppressed. Hence, between enhanced speech and speech after therapy, the enhanced signals got high preference score (test-3 in Table. 3). From the results of subjective and objective evaluations, the enhanced signals can be used to compare the improvement after the prosthesis / surgery.

Table 4: RESULTS OF OBJECTIVE EVALUATION

Parameters	Before Surgery/ Prosthesis	After Surgery/ Prosthesis	Enhanced Speech
% Nasal resonance	99.25	22.39	6.25
$A_1 - V_1$	5.62	3.8	6.25
$A_1 - V_2$	9.05	10.58	11.89
$A_2 - V_2$	8.80	15.80	12.35

5. Conclusion and Future Work

Motivated by the characteristics of speech after palatal prosthesis application, a two stage CLP speech enhancement algorithm is proposed by addressing the hyper-nasality issue in CLP. The algorithm involves the removal of extra nasal formant and formant enhancement. Perceptual evaluation results show that enhanced speech samples are having better quality in-terms of reduction in nasality. Enhanced speech signals resulted from proposed method are compared with clinical methods, i.e., surgery and prosthesis. The results of preference test and objective test show that stage-2 enhanced signals are nearer or better than that of speech after surgery / prosthesis. As the perceptual evaluation results of speech after therapy / surgery are highly biased to SLPs, enhanced speech by proposed method can be used as a perceptual standard to compare the improvement from clinical methods. Further, the work can be extended to enhance the consonant sounds like stops and fricatives, which are severely affected in CLP.

6. Acknowledgement

The authors would like to thank Dr. M. Pushpavathi, AIISH Mysore, for providing CLP samples and sharing the knowledge about CLP speech.

7. References

[1] Nman DS, Thomas P, Hodgkinson PD, Reid CA,"Oro-nasal fistula development and velopharyngeal insufficiency following primary cleft palate surgery - an audit of 148 children born between 1985 and 1997," *Br. J. of Plast. Surg.*, vol. 58, pp. 1051–1058, 2005.

[2] Koog T, "The pharyngeal flap operation in cleft palate," *Br. J. of Plast. Surg.*, vol. 18, pp. 265–283,1965.

[3] Murthy J, Sendhilnathan S, Hussain S A, et. al.,"Speech Outcome Following Late Primary Palate Repair," *Cleft PalateCraniofacial Journal,* vol.47, no.2, pp. 156-161, 2010.

[4] Joao Henrique Nogueira Pinto, Giseleda Silva Dalben, Maria Ines Pegoraro-Krook, "Speech Intelligibility of Patients With Cleft Lip and Palate After Placement of Speech Prosthesis," *The Cleft Palate-Craniofacial Journal*, vol. 44, no. 6, pp. 635–641, 2007.

[5] Debbie Sell,"Issues in perceptual speech analysis in cleft palate and related disorders: a review," *Int. J. Lang. Comm. Dis.*, vol. 40, no. 2, pp. 103–112, 2005.

[6] Gunilla Henningsson, David P. Kuehn, Debbie Sell, Triona Sweeney, Judith E. Trost-Cardamone, and Tara L. Whitehill, "Universal Parameters for Reporting Speech Outcomes in Individuals With Cleft Palate," *Cleft Palate-Craniofacial Journal*, vol. 45, no. 1, 2008.

[7] K. N. Stevens, Acoustics Phonetics, Cambridge, MA, USA: MIT Press, 1999.

[8] Gang Feng and Eric Castelli, "Some acoustic features of nasal and nasalized vowels: A target for vowel nasalization," *J. Acoust. Soc. Amer.*, vol. 99, no. 6, pp. 3694–3706, 1996.

[9] Yukie Kozaki-Yamaguchi, Noriko Suzuki, Yukihiro Fujita, Hidemi Yoshimasu, Masato Akagi and Teruo Amagasa, "Perception of Hypernasality and its Physical Correlates," *J. of Oral Science International*, vol. 2, no.5, pp. 21–35, 2005.

[10] Ryuta Kataoka, Donald W. Warren, David J. Zajac, Robert Mayo, and Richard W. Lutz, "The relationship between spectral characteristics and perceived hypernasality in children," *J. Acoust. Soc. Amer.*, vol.109, no. 5, pp. 2181–2189, 2001.

[11] Maeda, S. "Acoustic Correlates of Vowel Nasalization: A Simulation Study," *J. Acoust. Soc. Am. Suppl.*, vol. 1, no. 72, 1982.

[12] Betty Jane Philips and R.D.Kent, "Acoustic-Phonetic Descriptions of Speech Production in Speakers with Cleft Palate and Other Velopharyngeal Disorders," in Lass. Nj., Speech and Language: Advances in Basic Research and Practice, Academic press, vol 11, pp.132–160, 1984.

[13] Ryan Shosted, Christopher Carignan and Panying Rong,"Managing the distinctiveness of phonemic nasal vowels: Articulatory evidence from Hindi," *J. Acoust. Soc. Amer.*, vol. 131, no. 1, pp. 445–465, 2012.

[14] Tuomo Raitio, Antti Suni, Hannu Pulakka, Martti Vainio and Paavo Alku, "Comparison of Formant Enhancement Methods for HMM-Based Speech Synthesis," in *Proc. Interspeech*, Sept., 2010.

[15] P. Vijayalakshmi and M. R. Reddy, Analysis of hypernasality by synthesis, in *Proc. Int. Conf. Spoken Language Processing*, Oct. 2004.

[16] P. R. Vijayalakshmi, M. Ramasubba, and O. Douglas, "Acoustic analysis and detection of hypernasality using a group delay function," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 4, pp. 621–629, 2007.

[17] All Indian Institute for Speech and Hearing, Mysore, India: website: $http://www.aiishmysore.in/en/about_aiish.html$

[18] Makhoul, John, "Linear prediction: A tutorial review," *Proc. of the IEEE*, pp. 561–580, 1975.