



Designing visual Speech Training Aids for Hearing Impaired Children

S. Nilashree Wankhede

Assistant professor, Dept. Of Elect. and Telecom., Fr.C.Rodrigues Institute of Technology, Navi Mumbai, Vashi, India¹

ABSTRACT: Lack of auditory feedback for hearing impaired children leads eventually to a speaking disability in them. Hence, hearing impaired children are unable to speak, in spite of having proper speech production mechanism. Even if the hearing impaired child tries to speak by visualizing lip movements, his articulation, accuracy, stress and intonation patterns are affected, since vowels & consonants with tongue movement hidden in the mouth are not distinguishable to him and neither speech intensity or pitch variations are understood. Depending on the severity of the hearing impairment, either auditory, or tactile or visual feedback could be provided to the hearing impaired individuals. Speech-training systems can be designed based on feedback of acoustic parameters such as speech intensity, fundamental frequency, spectral features, or feedback of articulatory parameters such as voicing, nasality, lip and vocal tract movement. Using computer based speech training aids, for the hearing impaired children, the auditory feedback mechanism can be replaced by a visual representation of important acoustic parameters of speech such that as a child speaks he would be able to evaluate and correct his utterance/pronunciation based on expected and actual parameters that are displayed to him. This paper discusses about the speech training aids and the possible speech processing that can be done to obtain visual feedback for the hearing impaired children so that they get an opportunity to learn, speak and communicate properly through these aids. Also, an implementation using LPC analysis of speech is done to obtain vocal tract shape of children speaking vowels from various age groups so that it could be used as visual feedback to design a visual aid.

KEYWORDS: Acoustic, articulatory, feedback, pitch, intensity, speech, tactile, training, vocal tract

I.INTRODUCTION

Speech signal is produced in the form of pressure waves and consists of variations in pressure as a function of time. For each sound produced by human, there is positioning for each of the vocal tract articulators i.e. vocal folds, tongue, lips, teeth, velum and jaw [1]. Normal children can acquire the ability to control these various articulators by the age of four since they receive both visual and auditory feedback. However, hearing impaired children do not have access to the auditory feedback and hence they become dumb, in spite of having proper speech production mechanism. The hearing impaired children have neither auditory loop nor any remembrance of speech by themselves. Lip reading technique to teach hearing impaired people is not a solution since vowels & consonants with tongue movement hidden in the mouth are not distinguishable to them by simply visualizing lip movements. They may be able to understand message but not actually speak.

Nowadays, the speech-training systems can be designed based on visual or tactile feedback of acoustic parameters such as speech intensity, fundamental frequency, spectral features [15]-[23] or based on feedback of articulatory parameters such as voicing, nasality, lip & vocal tract movement [38]-[42]. The visual aid will be such that while a hearing impaired person speaks he would be able to evaluate and correct his utterance or pronunciation based on expected and actual parameters that are displayed to him. For example, they can see the articulation of their vocal tract shapes and compare it with a reference articulation to know where their articulation defects exist and correct them suitably [31]-[37].

Some important requirements for a speech training aid as summarized by Oster [3] are that, clear instructions and manuals must be created and made available for use with different groups of children. The visual feedback of the child's voice and articulation should be shown immediately on a computer screen and without delay. The aid must be acceptable to the speech therapist as well as to the child, which means that the aid must be attractive, interesting, easily



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

comprehensible, easy to handle, and motivating. The visual pattern must be natural, logical and easily understandable. This means that training parameters for example, pitch could be shown vertically as pitch variations occur; intensity through the size of an object that becomes larger as a sound becomes louder and smaller as a sound becomes softer and duration could be shown horizontally. The aid should finally give an objective evaluation of the child's training results.

In this paper section II discusses about use of auditory and tactile aids for hearing impaired children and their drawbacks. Section III discusses about speech training aids providing visual feedback. Section IV presents few implementation results of my preliminary research work carried out towards obtaining vocal tract shape for children from various age groups by performing LPC analysis of their speech. Section V provides conclusion and the future work towards designing visual speech training aid using the obtained results.

II. RELATED WORK: AUDITORY AND TACTILE DEVICES FOR HEARING IMPAIRED

Hearing loss exists when there is diminished sensitivity to the sounds normally heard. Deafness would be defined as the degree of poor hearing such that a person is unable to understand speech through hearing even in the presence of amplification. The severity of a hearing loss is categorized according to the additional intensity above a nominal threshold that a sound must be before being detected by an individual. Hearing impairment is measured in decibels of hearing loss or dBHL [4]. It may be ranked as mild, moderate, moderately severe, severe, profound or totally deaf as tabulated in Table I. Auditory aids may not be useful for severe and profound deafness.

TABLE I. Categorization of hearing loss

Severity of hearing loss	Hearing impairment Range in dBHL [4]
Mild	20 to 40 dBHL
Moderate	41 to 40 dBHL
Moderately severe	55 to 40 dBHL
severe	71 to 40 dBHL
Profound deafness	91 dBHL
Total deafness	Above 91 dBHL

A hearing aid selection is based on the comparison in four dimensions i.e. sensitivity to sound, tolerance limit, efficiency in background noise and efficiency in distinguishing small sound differences. Traditional speech training with profoundly hearing-impaired children is based on methods that help the children to learn speech by looking at the therapist's face and lips, through residual hearing or by tactile feedback where feeling of the therapist's face, throat and expiration air, etc. is obtained to establish control of their speech. A very wide range of devices have been developed for people with hearing loss [9]-[11]. Levitt [1][5][6] categorized these devices not only by the modality of stimulation [i.e., auditory, visual, tactile, or direct electrical stimulation of the auditory nerve (auditory-neural)] but also in terms of the degree of speech processing that is used.

The assistive listening devices (ALDs) using auditory channel, include high-gain telephones and listening systems for rooms and auditoria in which the signal is transmitted by electromagnetic means to a body-worn receiver [5]. Low-power FM radio or infrared transmissions are typically used for this purpose. The primary function of most assistive listening devices is to avoid the environmental noise and reverberation that are typically picked up and amplified by a conventional hearing aid. FM transmission is typically used in classrooms and infrared transmission in theatres, auditoria, houses of worship, and other public places.

Uptil now the major thrust in hearing-aid development has been toward instruments of smaller and smaller size because most hearing-aid users do not wish to be seen wearing these devices. Miniature hearing aids that fit entirely in the ear canal and are barely visible are extremely popular. Even smaller hearing aids have recently been developed that occupy only the innermost section of the ear canal and are not visible unless one peers directly into the ear canal [5].

A common characteristic of sensorineural hearing impairment is that the degree of impairment increases with increasing frequency. It has been suggested that speech intelligibility could be improved for this form of hearing loss

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

by transposing the high-frequency components of speech to the low-frequency region, where the hearing impairment is not as severe. A series of investigations on frequency-lowering schemes and its use for hearing impairments were carried out by Posen [7] and Johansson [8] respectively. Commercially available speech training devices based on auditory feedback are Chatter Phone (Model PVC-C) [9] and Speak Easy (Model 4392) [10].

Sometimes the deafness in children may be due to damage to the hair cells in the cochlea which convert mechanical vibration to neural firings. But if neurons connected to these hair cells are still functional, they can be triggered by the electromagnetic field generated by a cochlear implant. It artificially stimulates the cochlear nerve by providing an electric impulse substitution for the firing of hair cells. If cochlear implant is done at a very young age, some profoundly impaired children can acquire effective hearing and speech, particularly if supported by appropriate rehabilitation. However, cochlear implants are not only expensive, but require sophisticated programming in conjunction with training for effectiveness. Cochlear implant recipients may be at higher risk for meningitis which becomes a disadvantage [11] [12].

Speech can also be presented through vibrators at the fingertips and other parts of the body to indicate various elements of speech such as voicing or nasalization [1]. Such an aid is called a tactile aid which could be used as a supplement to speech-reading and for speech training. An important advantage of tactile sensory aids is that these devices can be worn conveniently without interfering with other important sensory inputs. There are examples of deaf individuals who have worn spectrum-based tactile aids for many years and who are able to communicate effectively in face-to-face situations using their tactile aid [6] [38].

Figure 1[6] shows a schematic diagram of an example of a tactile aid where speech signal processing is involved. This aid is worn on the lower forearm that provides information on voice fundamental frequency, F_0 . Speech signals from the input transducer (an acoustic microphone or a surface-mounted accelerometer) are delivered to a pitch extractor that generates a square wave whose frequency equals one-half that of the fundamental voice frequency. The electronic components are housed in a body-worn unit, the output of which is connected by wire to a transducer array worn on the wrist as shown in Figure 1. The locus of tactile stimulation is proportional to the value of F_0 . When F_0 is low, the region near the wrist is stimulated; as F_0 is increased, the locus of stimulation moves away from the wrist.

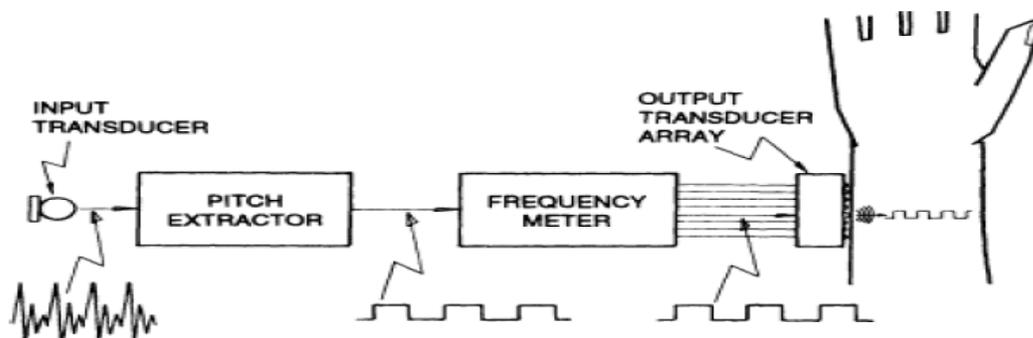


Fig 1. Wearable tactile display for voice fundamental frequency [6].

The tactile F_0 display has also been found to be of value as a speech-training aid. A very useful feature of this aid is that it can be worn outside the classroom, thereby providing students with continuous feedback of their speech production in everyday communication [6]. The central problem facing the development of practical speech-feature tactile wearable aids is that the signal picked up by the microphone on the wearable aid is contaminated by environmental noise and reverberation.

III. SYSTEMS BASED ON VISUAL FEEDBACK FOR SPEECH TRAINING

Despite the fact that many of auditory and tactile aids have reported to improve the speech of some children, the use of them was limited. This was probably due to the fact that the feedback provided by this type of speech training aids was difficult to understand, unnatural, delayed, unattractive, and had no motivational impact on the children. The earliest visual sensory aids were concerned primarily with making sounds visible to a deaf person. The limitations of these

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

uniform growth of vocal tract section length is observed as age increases [45]-[46]. There is an increase in vocal tract length predominantly due to growth in the pharyngeal region. Longer pharynx is observed in adult men compared to women and children. Also, there is descent of the larynx, the hyoid bone and the tongue and lengthening of the vocal tract with a decrease in the oro-laryngo-pharyngeal angle [46]. Therefore, if we observe the entire vocal tract length, we can say that as compared to adults the more variations in vocal tract areas could be seen closer to glottis. Overall vocal tract shape in children will still remain the same as that of adults except that there will be shift of all vocal tract area values towards the glottis end indicating a very small length of pharyngeal region where hardly any changes are obtained in area values.

In my research work, implementation of LPC analysis on speech produced by children is done to obtain VTS for children. Wakita's method based on LPC analysis of speech has been used, where spoken vowels /a/, /i/ and /u/ were acquired from children in various age groups at decided sampling frequencies. Based on the observations using MRI images [44][46], for different age groups of children, there is a need to appropriately select vocal tract length, LPC order and speech sampling rate while estimating their realistic vocal tract shapes using LPC analysis. For different age groups of children various set of optimum parameter values were found. Pre-emphasis was done before applying a Hamming window of 20 ms to the speech signal. Autocorrelation coefficients and reflection coefficients were obtained using LPC analysis. Using the reflection coefficients, finally vocal tract shape in terms of vocal tract area function were obtained for different prediction orders. Figures 8, 9 and 10 shows VTS obtained for the age group of 2 to 6 years for vowels /a/, /i/ and /u/, respectively. For every vowel, five VTS are displayed by varying prediction order from 8 to 12 in steps of one.

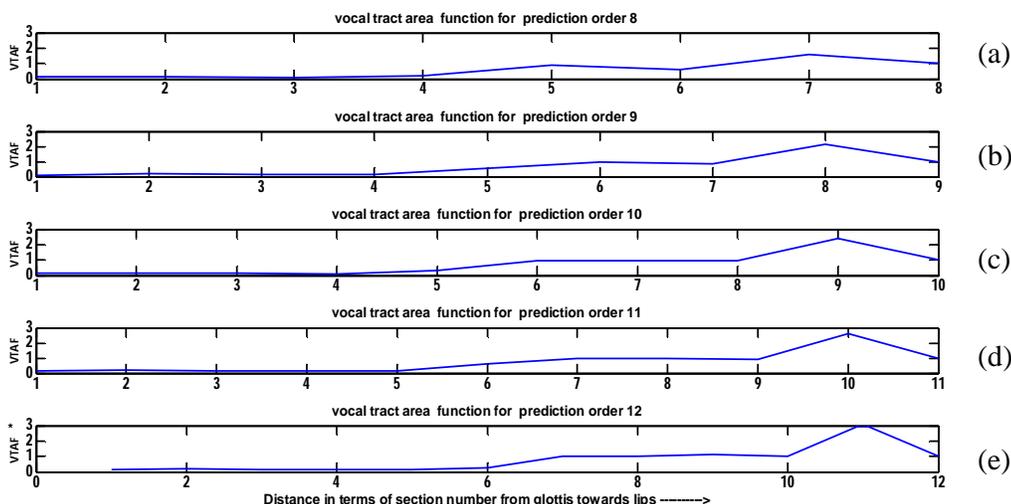


Fig. 3 VTAF results obtained for a 6 year old girl speaking vowel /a/ with LPC order (a) M=8 (b) M=9 (c) M=10 (d) M=11 (e) M=12, (Fs= 14 kHz, average VTL, l=10cm)

Likewise we had divided the children into 5 age groups, i.e. from 2 to 6, 7 to 9, 10 to 12, 13 to 16 and 17 to 21 years and found VTS in each case for different prediction order. The estimated shapes based on LPC analysis are presented in terms of cross-sectional area along y-axis vs. section number from glottis to lips along x-axis. Realistic vocal tract shapes for a particular age group of child is obtained only at specific prediction order as tabulated below in Table II.

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

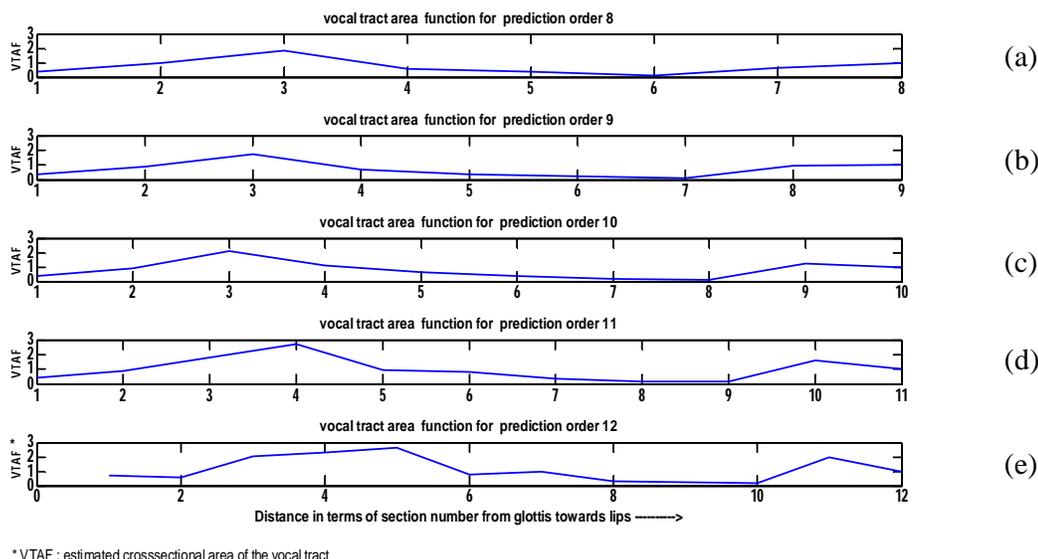


Fig. 4 VTAF results obtained for a 6 year old girl speaking vowel /i/ with LPC order (a) M=8 (b) M=9 (c) M=10 (d) M=11 (e) M=12, ($F_s = 14$ kHz, average VTL, $l = 10$ cm)

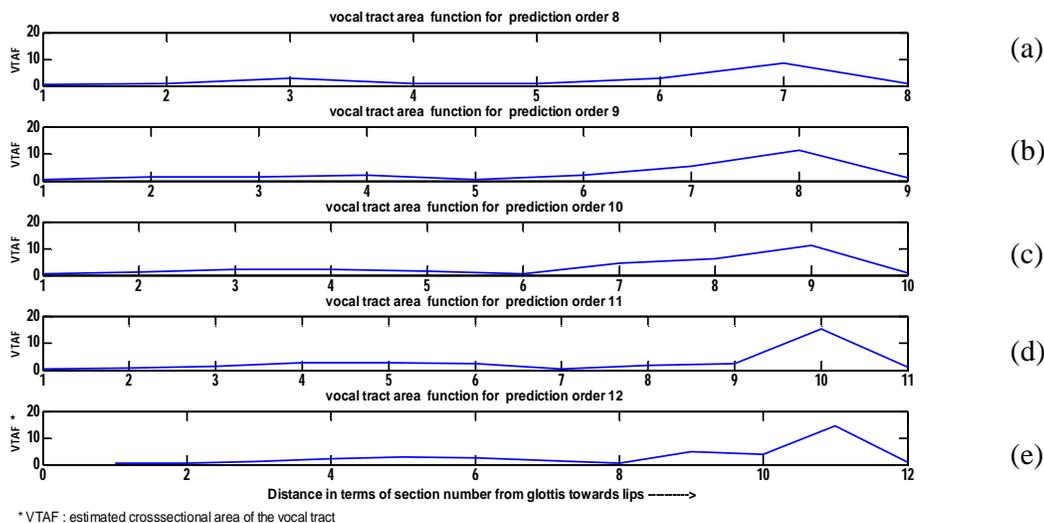


Fig. 5 VTAF results obtained for a 6 year old girl speaking vowel /u/ with LPC order (a) M=8 (b) M=9 (c) M=10 (d) M=11 (e) M=12, ($F_s = 14$ kHz, average VTL, $l = 10$ cm)

The VTS obtained for different age groups using our implementation could be used as a reference shape in speech training aids where for that particular age group, the hearing impaired child can compare his own VTS with reference shape for a spoken vowel.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

TABLE IILPC order for realistic VTS In that age group

Age group in years	LPC order
	M
2 to 5	8
6 to 9	8 or 9
10 to 12	9 or 10
13 to 16	10 or 11
17 to 21	11 or 12
Above21(adult)	12

V. CONCLUSION

The paper discussed various ways to train the hearing impaired children so as to improve their speaking abilities. It leads to the conclusion that for a child to develop intelligible speech, a visual speech training aid is better than tactile aid. Using visual aids, a hearing impaired child can better understand speech and try to learn speaking. The motive behind implementation of LPC analysis of speech to obtain VTS for a child is that using a computer based aid, the child can see the articulation of his own vocal tract shape and compare it with a reference articulation provided by the speech training aid and try to minimize the mismatch in his VTS. But, to design such a speech training aid, vocal tract areas need to be estimated with consistency and appropriate dynamic response. For hearing impaired children, appropriate displays, cartoons, or games based on dynamically varying vocal tract shape need to be devised. Hence, obtaining realistic shapes for children in various age groups is a must so that a designed speech training aid could be used by children from any age group effectively. Help and assessment results need to be given regularly by a therapist or a teacher to the hearing impaired child until he masters to use such a speech training aid.

REFERENCES

- [1] L. R. Rabiner and R. W. Schafer, Digital Processing of Speech Signals. Englewood Cliffs, NJ: Prentice Hall, 1978.
- [2] B. R. David and G.W. Jay, (Eds.), Voice Communication Between Humans And Machines. Washington : National Academy Press, 1994.
- [3] A.M. Oster, "Clinical applications of computer-based speech training For children with hearing-impairment", in Proc. of 4th International Conference on Spoken Language Processing, Philadelphia, USA, pp.157-160,1996
- [4] Available :<http://en.wikipedia.org/wiki/Ear>. Last accessed on 30th July 2012.
- [5] H. Levitt, J. M. Pickett, and R. A. Houde, (Eds.), Sensory Aids For the Hearing Impaired. New York: IEEE Press, 1980.
- [6] S. Rosen, J. R. Walliker, A. Fourcin, and V. Ball , "A micro-processor-based acoustic hearing aid for the profoundly impaired listener", J. Rehabil. Res. Dev., vol4, pp23, 1987.
- [7] M. Posen, C. M. Reed, and L. D. Braida , "The intelligibility of frequency-lowered speech produced by a channel vocoder", J. Rehabil. Res. Dev., 30(1), pp.26-38, 1993.
- [8] B. Johansson, "The use of the transposer for the management of the deaf child.", Int. Audiol., 5: pp 362-373,1966.
- [9] Available : <http://www.aimeesolutions.com>. Last accessed on 4th Aug 2012.
- [10] Available : <http://www.enablingdevices.com>. Last accessed on 4th Aug 2012.
- [11] Available : <http://en.wikipedia.org/wiki/Deaf> . Last accessed on 30th July 2012.
- [12] W. House and J. Urban, "Long-term results of electrical implantation and electronic stimulation of the cochlea in man", J. Ann. Otol. Rhinol. Laryngol., pp. 82:504510, 1973.
- [13] R.E. Stark, "Teaching /ba/ and /pa/ to deaf children using real-time spectral displays", J. Lang. Speech, 15: pp.14-29, 1972.
- [14] R. S Nickerson and K.N. Stevens. "Teaching speech to the deaf: can a computer help?", IEEE Trans. Audio Electroacoust., vol. 21,no. 5, pp.445-455,1973.
- [15] A. E. Coyne, " The Coyne voice pitch indicator", Volta Review, vol 40, pp.437-439, 1938.
- [16] O.O. Gruenz, "Extraction and portrayal of pitch of speech sounds." J. Acous. Soc. Am., vol 21, pp.487, 1938.
- [17] L.O. Schott, "Extraction of pitch of speech sounds", J. Acous. Soc. Am., vol 24, pp.211, 1949.
- [18] Risberg, "Speech processing aids for the deaf", J. of speech and hearing research, Vol.13, pp.22-24,1968
- [19] Fletcher, "Visual speech apparatus feedback ", J. of Speech and Hearing Disorders, Vol.48, pp.178-185, 1983.
- [20] L.E. Bernstein and Mahshie, "Speech training aids for hearing-impaired individuals : I . Overview and aims", J. Rehabil. Res. Dev. , Vol.25, no.4, pp.59-62, 1988.
- [21] S. Zahorian and S. Venkat, "Vowel articulation training aid For the deaf". In Proc. of Int. Conf. on Acoustics, Speech, and Signal Processing, pp.1121-1124,1990.
- [22] E.B. Andrew and S. A. Zahorian, "Transformations of speech spectra to a two dimensional continuous valued phonetic feature space for vowel training", in Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing, vol.2, pp. 241 – 244, 1992.
- [23] B. Xun ,D. Yang, F. Zhang, X. Wang, "Study of computer aided speech training method For deaf children based on Learning vector Quantization", Chinese Control and Decision Conference, China, pp. 4266-4269, 2008.



International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2014

- [24] Available : <http://www.kayelemetrics.com> . Last accessed on 5th Aug 2012.
- [25] Available : <http://www.drspeech.com/SpeechTrain2.html>. Last accessed on 5th Aug 2012
- [26] C. S. Watson, M. Elbert and G. DeVane, "The Indiana Speech Training Aid (ISTRA)", J. Acoust. Soc. Am., Vol. 81, Issue S1, pp.95, 1987.
- [27] A.M. Oster, "Auditory and visual feedback in spoken L2 Teaching", Reports from the Dept of Phonetics, Umeå University, PHONUM 4,1997.
- [28] Available : www.videovoice.com . Last accessed on 5th Aug 2012.
- [29] J. B. Ferguson, L. E. Bernstein and M. H. Goldstein, "Speech training aids For hearing-impaired individuals:II. Configuration of the Johns Hopkins aids", J. Rehabil. Res. Dev., Vol . 25 , pp. 63-68,1988.
- [30] J. J. Mahshie, V.M. Diane, W.S. Betty and L.E. Bernstein, "Speech training aids For hearing-impaired individuals: III.Preliminary observations in the clinic and childrens' homes", J. Rehabil. Res. Dev., Vol . 25, No . 4, pp.69-82, 1988.
- [31] H. Wakita, "Direct estimation of the vocal tract shape by inverse filtering of acoustic speech waveForms," IEEE Trans. Audio Electrocaust., vol. 21, no. 5, pp. 417-427, 1973.
- [32] A.Seshadri, K.Makwana, T.Joseph, M. Varde, "Animation of vocal tract shape for speech training of hearing impaired", Project Report(Guidance:Dr. Milind Shah), Department Of Electronics And Telecommunication Engineering FCRIT, Univ. Of Mumbai , 2009.
- [33] S.H Park , D. J. Kim , J. H . Lee and T. S. Yoon , "Integrated speech training system For hearing impaired," IEEE Trans, Rehab, Engg., vol. 2, no.4,pp, 189-196, 1994.
- [34] A.E. Mahdi," Visualisation of the vocal-tract shape for a computer-based speech training system for the hearing-impaired", The Open Electrical and Electronic Engineering Journal, vol. 2, pp. 27-32, 2008.
- [35] M.S.Shah and P.C.Pandey, "Estimation of place of articulation during stop closures of VCV utterances", IEEE Trans. Audio, Speech, Language Processing, vol.17, no.2, pp.277-286, 2009.
- [36] M. S. Shah and P.C. Pandey, "Estimation of vocal tract shape for VCV syllables for a speech training aid", in Proc. 27th annual conference of the IEEE Engineering in Medicine and Biology Society (Shanghai, China), pp.6642-6645,2005.
- [37] P. C. Pandey and N. Nagesh, "Estimation of lip opening for scaling of vocal tract area function for speech training aids", National Conf. on communications, pp3-5, 2012.
- [38] J. D. Miller, A. M. Engebretsen and C. L. DeFilippo, " Preliminary research with a three-channel vibrotactile speech-reception aid For the deaf", in Proc. of the Seminar on Speech Communication, Vol. 4,1974, pp 230-234.
- [39] H. Javkin, N. A. Barroso, A. Das, D. Zerkle, H. Levitt and K.Youdelman, "A motivation sustaining articulatory/acoustic speech training system for profoundly deaf children," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., pp. 145-148, 1993.
- [40] L. E. Bernstein, J.B. Ferguson and M . H. Goldstein , "Speech training devices for profoundly deaf children," in Proc. IEEE Int. Conf. Acoust. Speech, Signal Process., pp. 633-636, 1986.
- [41] A. Boothroyd, L. Hanin, E. Yeung, Q. Chen," Video-game for speech perception testing and training of young hearing-impaired children", in Proc. of Int. conf. Computing Applications to Assist Persons with Disabilities, NY, pp.25-28, 1992.
- [42] J. M. Pardo, "Vocal tract shape analysis for children," in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process, pp.763-766,1982.
- [43] Vorperian,H.K., Wang,S., Chung,M., Michael,E., Reid R., Kent, Andrew,Z., Lindell,R., "Anatomic development of the oral and pharyngeal portions of the vocal tract: an imaging study", J. Acoust. Soc. Am., vol 125, no.3, pp. 1666-1678, 2009.
- [44] Vorperian,H.K., Kent,R., Gentry,L., Yandell,B., "Magnetic resonance imaging procedures to study the concurrent anatomic development of vocal tract structures: preliminary results", International Journal of Pediatric Otorhinolaryngology, vol 49, pp.197-206, 1999.
- [45] Vorperian,H.K., Kent,R., Gentry,L., Yandell,B., "Development of vocal tract length during early childhood: A magnetic resonance imaging study", J. Acoust. Soc. Am., vol 117, no.1, pp. 338-350, 2005.
- [46] Fitch,W., and Giedd,J., "Morphology and development of the human vocal tract: A study using magnetic resonance imaging", J. Acoust. Soc. Am., vol 106, no.3, pp. 1511-1522, 1999.