

Analysis of Deaf Speakers' Speech Signal for Understanding the Acoustic Characteristics by Territory Specific Utterances

Nirmaladevi Jaganathan¹, Bommannaraja Kanagaraj²

¹Department of Information Technology, Excel Engineering College, Namakkal, India

²Department of Electronics and Communication Engineering, KPR Institute of Engineering and Technology, Coimbatore, India

Email: jvandana25@gmail.com, dr.k.bommannaraja@gmail.com

Received 7 April 2016; accepted 5 May 2016; published 17 June 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

An important concern with the deaf community is inability to hear partially or totally. This may affect the development of language during childhood, which limits their habitual existence. Consequently to facilitate such deaf speakers through certain assistive mechanism, an effort has been taken to understand the acoustic characteristics of deaf speakers by evaluating the territory specific utterances. Speech signals are acquired from 32 normal and 32 deaf speakers by uttering ten Indian native Tamil language words. The speech parameters like pitch, formants, signal-to-noise ratio, energy, intensity, jitter and shimmer are analyzed. From the results, it has been observed that the acoustic characteristics of deaf speakers differ significantly and their quantitative measure dominates the normal speakers for the words considered. The study also reveals that the informative part of speech in a normal and deaf speakers may be identified using the acoustic features. In addition, these attributes may be used for differential corrections of deaf speaker's speech signal and facilitate listeners to understand the conveyed information.

Keywords

Deaf Speaker, Hard of Hearing, Deaf Speech Processing, Assistive Mechanism for Deaf Speaker, Speech Correction, Speech Signal Processing

1. Introduction

Speech is the strongest communication medium for all human beings, while it becomes hardest for deaf speakers

How to cite this paper: Jaganathan, N. and Kanagaraj, B. (2016) Analysis of Deaf Speakers' Speech Signal for Understanding the Acoustic Characteristics by Territory Specific Utterances. *Circuits and Systems*, 7, 1709-1721.

<http://dx.doi.org/10.4236/cs.2016.78148>

or hard of hearing subjects (HoH). This communication medium should sustain, preserve and prolong in its quality to convey their thoughts [1]. Generally the amount of hearing loss ranges from moderate to profound to severe and measured in decibel (dB). The profound to severe hearing loss of ≥ 70 dB occurs because of congenital facts and becomes complex to understand the acoustics as well as verbalization phonetics. The speech output of deaf speakers is less intelligible and hence it is difficult for the listeners to understand the content of articulation. The acoustic characteristics to greater extent can be comprehended with available speech analyzing techniques [2]-[4] to contemplate for the posed problem.

Deaf speakers born with severity or profundity or became deaf in infancy can achieve intelligible speech naturally by training or administering a battery of speech and language tasks [5], but most of such speakers do not produce legible speech due to the fact that speech intelligence is a correlated factor with residual hearing. Even if they attempt to place their articulators accurately, they fail to coordinate the movements of articulation [6]. The negative effect on intelligence of deaf speaker is due to the easy movement of transitional sounds with very long duration, pitch breaks and short pauses that have some positive effects on speech intelligence [7]. Also, the timing of the deaf speaker ranges higher and by correcting the timing errors in deaf speech, the intelligibility can be improved.

In many cases, the utterances are modified and the relative timing error corrections are fed to computer for obtaining the highest intelligence score that results in better speech production [8] [9]. Deaf speakers have a strong relationship between prosodic features and phoneme production that directly influence the speech intelligence. Such intricate relationship can be perceived as pitch variations in deaf speakers [10].

The unintelligent speech information of the deaf speakers is replicated in form of random formant position, abnormality in fundamental frequency (F0) and intensity contours. F0 is an important parameter in assessing the correctness of speech production. Due to low intelligence, the deaf speaker's second formant (F2) range is reduced both by time and frequency [11]. Further the wide variation in F0 with high jitter values is noticed with lack of control of amplitude. Consequently, majority of the hearing impaired speakers need to be educated with the speech skills that normal speakers readily acquire during the initial years of life. To develop such speaking skills, procedures and techniques for the benefit of deaf speakers, efforts have been taken in studying the speech characteristics of the deaf speakers [12]. The progress in the sophisticated processing and analyzing techniques in speech science, electrical engineering and computer science have facilitated to understand the speech production process [13]. These technological growths helped to examine the deaf speakers' speech signal as well as to develop clinical assessment and training procedures [14] [15].

In recent years, attempts are made to clinically separate or identify the differences between the normal and deaf speech signals. The speech parameters like pitch, formants, spectral and cepstral coefficients have been explored to correlate the signals. But each work focuses on specific group of community, differs in origin and demography. Also various assistive technologies have been developed to support the deaf community to communicate using gadgets or modules. These devices provide a way to assess the information through vision and/or vibration method. Few such devices are pitch indicator, nasality indicator, spectrum analyzer [16], talking heads and vocal tract estimator [17]. Multiple computer based training methods [18]-[24] are also developed to improve the deaf speaker's utterances in terms of intelligibility, speech recognition, speech production, speech synthesis, speech enhancement and gender identification.

Further, many researchers, manufacturers and clinicians determine a way for synthetic speech generation using alternative and augmentative communication method, neural networks and face recognition techniques [25]. A touch screen based system has been employed as a communicating device for the deaf speaker as a result the normal speaker can understand the information conveyed by the deaf speaker. A GSM module has been inbuilt in addition, so that the deaf speaker may be able to communicate even over long distances [26]. Advancements in the emotional recognition facilitated to develop talking head interference to assist the deaf people in expressing their emotional status using gestures in various languages [27] [28]. The mentioned techniques confiscate the barrier of a deaf speaker to make know their thoughts to the listeners.

With such background, an attempt is made to instill a technique that enhances the deaf speaker's speech utterances by analyzing the degree and depth of variation in speech signal. This may contribute for developing an assistive mechanism that can generate a speech signal with deaf speaker's tone of voice. The objectives of the present work are 1) to evaluate the speech parameters and understand the deviation (degree) and dislocation (depth), 2) to study the viability of deriving correction measures using quantified deviation and dislocation and 3) to use correction measure for enhancing the deaf speakers speech and produce speech signal with same tone of

voice. Such study may facilitate to modify the deaf speech signal to resemble normal.

2. Materials and Methods

In the state of Tamilnadu, India, a private organization namely, Madras ENT Research Foundation (MERF) has made a survey from 2003-2013 among the children's aged below 12 years and found that 6 out of 100 children's are affected by deafness problem for various congenital reasons. The report also indicates that the deafness in the state is 3 times more than the National average and 6 times greater than the International average. Particularly, the state Tamilnadu is positioned 3rd in south India. Further, the survey proclaims that the deaf population in south India is almost 17.5% higher than that of the other states. Therefore, in this present work, the deaf speakers from various districts of Tamilnadu, speaking Tamil language, are considered for the investigation.

32 normal speakers (NS) with no hearing loss and 32 deaf speakers (DS) with hearing loss are selected for the study. The age of the speakers ranges between 8 to 12 years. The deaf speakers are selected from various hearing impairment schools. The amount of hearing loss for the deaf speakers chosen is ≤ 70 dB under the category of moderate hearing loss.

Feature Extraction for the Vocal Characteristics of the Speech

The Praat software is used for extracting the features from the speech signal and it is analysed to understand the speech parameters of the normal and deaf speakers. The subjects are instructed to be seated comfortably and asked to perform deep breath inhalation and exhalation. They utter the chosen 10 Tamil words with the help of a display board and a speech trainer. With the comfortable audio intensity by using the SONY ICD-UX523F voice recorder fixed at suitable height and distance of 5 inches from the mouth in a sound proof room, both the group of speakers utter the words repeatedly for 3 times. The sampling frequency is fixed as 44,100 Hz. The time gap for each utterance for a given word ranges from 1 sec to 3 sec.

The phonetic information of the selected Tamil words with classification details is given in [Table 1](#). The recorded speech samples are processed using signal processing to suppress background noise. To have a better representation of the speech signal, the mean signal is generated for 3 repeated recordings and then speech parameters are estimated, analysed in addition, data collection of the various vocal parameters are measured.

Pitch: The Pitch is an important attribute of sound and it conveys the bulk of the prosodic information in speech. Pitch determines the speech harmonics based on the harmonic peak lowness and highness of a waveform and it is usually related to its fundamental frequency hence, it is called as pitch frequency or fundamental frequency (F0). It is represented in Hertz [Hz].

Normally the speech contains distribution of many peaks and an algorithm is needed to find the possible pitch candidates. Here autocorrelation algorithm is used to determine the Pitch candidates. To analyze the pitch candidates the speech signal is divided into number of segments. The time step is fixed as 0.01 seconds which calculates 100 pitch values per second. The window length is 40 ms or 0.04 seconds because the pitch floor is fixed as 75 Hz and the frequency below this range will be eliminated, also it calculates 4 pitch values in one window

Table 1. Phonetic Information of the Tamil language words with classification.

Test Words	Phonetic Transcription	No of Syllable	Words Classification
Amma	[əməmə]	2	Open Nasal
Appa	[əbəbə]	2	Open Plosives
Bhandhu	[bɦəndɦu]	2	Close Nasal Plosives
Kan	[kənə]	1	Open Nasal Plosives
Kathavu	[kəɖəɦəvu]	3	Open Plosives
Kathu	[kəɖəɦu]	2	Close Plosives
Mani	[məni]	2	Open Nasal
Maram	[məərəmə]	2	Open Nasal
Veedu	[ve:ɛ:du]	2	Open Plosives
Paal	[pəələ]	1	Close Plosives

length. The analysis of each segment results with a number of pitch candidates with an indication of its strength. The best pitch candidate can be found at the maximum peak using the algorithm. In autocorrelation method the Pitch candidates are found in each frame and all the consecutive frames are analyzed. Here the same pitch candidate if occurred that will not be considered second time when the transition is made between the analysis frame. Finally the mean pitch value is obtained using all the pitch candidate of window frames. Thus the appropriate pitch value can be determined using this algorithm which is accurate, non-resistant and robust than other methods.

The following equation is used to calculate the pitch candidates of a windowed signal.

$$a(t) = s(t)w(t). \quad (1)$$

Here windowed signal $a(t)$ is obtained by convolving $w(t)$ which is a window function on the analysis frame $s(t)$.

Formants: The spectral peaks of speech spectrum of a speech signal are called “Formants”. It is mentioned as $[P(f)]$. The formant values are estimated to know the relationship between the words and to find whether they are similar. Usually a maximum of five formants for men are present within every 1000 Hz and the maximum frequency range is 5000 Hz. For women an average of five formants present between every 1100 Hz and the maximum frequency range is 5500 Hz. The pre-emphasis form is fixed up to 50 Hz which boosts the higher formants that will be helpful in finding the formants in the analysis.

The first Formant frequency (F1) will be higher for open vowel and lower for closure ones. The First formant roughly corresponds to the vowel height and the second for vowel location. The remaining formants will get increased or decreased depending on the first Formant (F1) frequency. **Table 1** further shows the classification of ten Tamil words based on nasality, which is used for deaf speech analysis.

Formants can be estimated using the Linear Predictive Coding (LPC) method. Linear Predictive Coding (LPC) analysis is based on the source filter theory of speech production. According to this theory, speech signal is the result of the source signal filtered from the vocal tract system. The LPC method will separate the source and filter from the acoustic signal. The LPC source filter multiplication model relationship is mentioned below:

$$O(f) = H(f) \cdot S(f) \quad (2)$$

$O(f)$ —the spectrum of sound, $H(f)$ —the spectrum of filter $S(f)$ —the spectrum of the source.

The formants are analysed to find the various acoustic transitions of speech on the vocal passage where F1 is related to pharynx, F2 for oral cavity, F3 for nasal cavity, F4 for sinuses and above F5 normally the formants range will be very poor to identify. In the present work maximum of five formants values are estimated.

Jitter and Shimmer: The Jitter and Shimmer are the disturbance indexes of the fundamental frequency. The cycle variations of fundamental frequency and amplitude are called “Jitter and Shimmer” respectively. The Jitter is finding the average absolute difference between consecutive periods by average period and the threshold of Jitter for pathological voices is 1.040%.

$$jitter = \frac{\sum_{i=2}^{N-1} |2 \cdot T_i - T_{i-1} - T_{i+1}|}{\sum_{i=2}^{N-1} T_i} \quad (3)$$

where, $T_i = i^{\text{th}}$ interval and N = number of intervals.

The Shimmer is found by the average absolute difference between the amplitudes of consecutive periods by average amplitude. It is also called “Shimmer (Local)”. The threshold of shimmer for pathological voices is 3.180%.

$$Shimmer = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} [A^{(i)} - A^{(i+1)}]}{\frac{1}{N} \sum_{i=1}^N (i)} \quad (4)$$

where, $A(i)$ —which is the extracted peak-to-peak amplitude data.

Signal-to-Noise Ratio: Signal-to-noise ratio (SNR) is a standard measure of the amount of background noise present in a speech signal. It is defined as the ratio of signal intensity to noise intensity and is expressed in decibels.

$$SNR(\text{dB}) = 20 \log_{10} (S/N) \quad (5)$$

where S is the root-mean-square of the speech signal without any noise present and N is the root-mean-square level of the noise without including speech.

3. Results and Discussion

The quantitative result for the speech parameters of NS and DS groups to understand the significant variations in their speech characteristics is discussed. **Table 2** shows the values of pitch, jitter and shimmer of NS and DS groups respectively. The dissimilarity within the group is noticed which indicate the uniqueness of voice parameters of a speaker. From the overall pitch study, it is inferred that the significant pitch variations are noticed for two syllable words.

From the results as shown in **Table 2**, it is noticed that the minimum pitch value (#NS08) is 263.97Hz for the word “Maram” (Two syllable word) and maximum pitch value (#NS30) is 298.03 Hz for the word “Appa” (Two Syllable word). In the case of DS group the minimum pitch value (#DS04) is 413.95 Hz for the word “Kan” (One syllable word) and maximum pitch value (#DS12) is 433.04 Hz for the word “Veedu” (Two syllable word). The NS group pitch value ranges between 263.97 ± 34.06 Hz and for deaf speaker it is between 413.95 ± 19.05 Hz for deaf speaker, which is higher in frequency. The deviation between the NS and DS groups for the mean pitch values is 149.98 Hz, which states that the deaf speakers pitch value starts at a higher value compared to a normal speaker.

A plot is presented for the word “Appa” of both the groups taken F0 in x-axis and F1 in y-axis as shown in **Figure 1**. The mean value is calculated for F0 and F1 for both the groups and an average distance is measured between the two clusters as shown in the figure using the distance formula for two independent groups. The average distance for the word “Appa” is 407.32 Hz. The figure shows that the values of pitch and formant values of normal speakers form a separate cluster and the same form a separate cluster for deaf speakers.

A formant chart for the NS and DS group is depicted **Figure 2** and **Figure 3** for the chosen Tamil words.

Table 2. Descriptive analysis of pitch, jitter and shimmer for normal and deaf speakers.

Test Words	Subjects	Pitch (Hz)	Jitter (%)	Shimmer (%)
Amma	NS	271.58	1.12	3.27
	DS	416.68	2.15	6.89
Appa	NS	298.03	1.00	3.05
	DS	419.81	2.23	6.08
Bhandhu	NS	267.17	1.33	3.00
	DS	418.68	2.60	6.91
Kan	NS	266.89	1.34	3.01
	DS	413.95	2.85	6.95
Kathavu	NS	273.78	1.32	3.24
	DS	416.30	2.58	5.96
Kathu	NS	277.69	1.22	3.13
	DS	419.44	2.67	6.80
Mani	NS	268.12	1.35	3.08
	DS	415.65	2.70	7.25
Maram	NS	263.97	1.25	3.31
	DS	420.92	2.54	5.84
Veedu	NS	271.08	1.18	3.27
	DS	433.04	2.53	6.89
Paal	NS	269.63	1.37	3.05
	DS	427.11	3.00	6.08

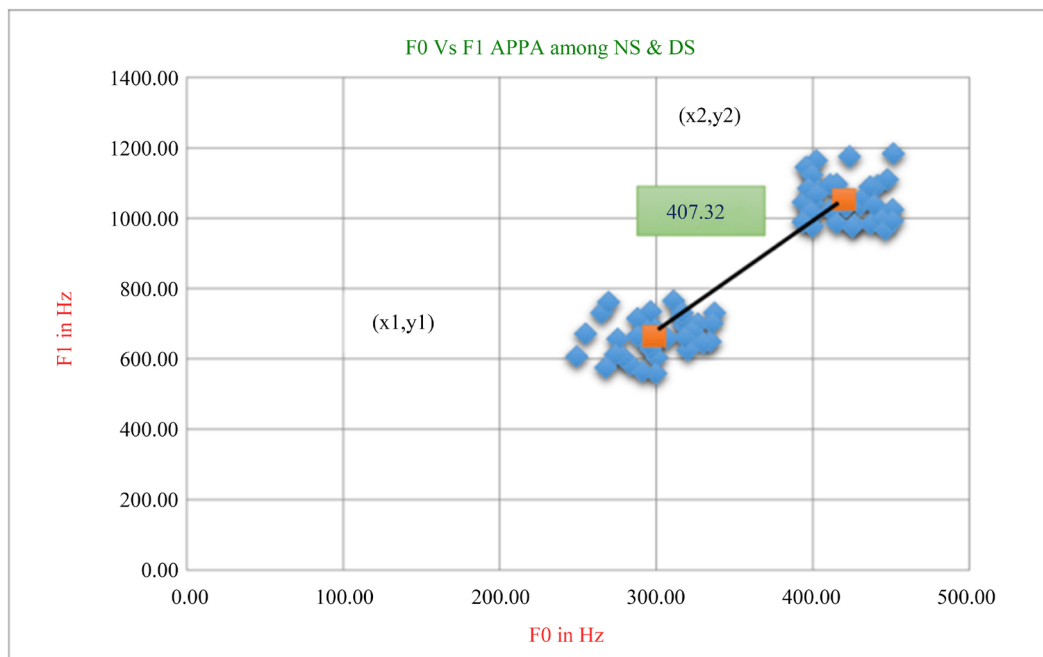


Figure 1. Cluster formation for the Tamil word “Appa” using F0 and F1.

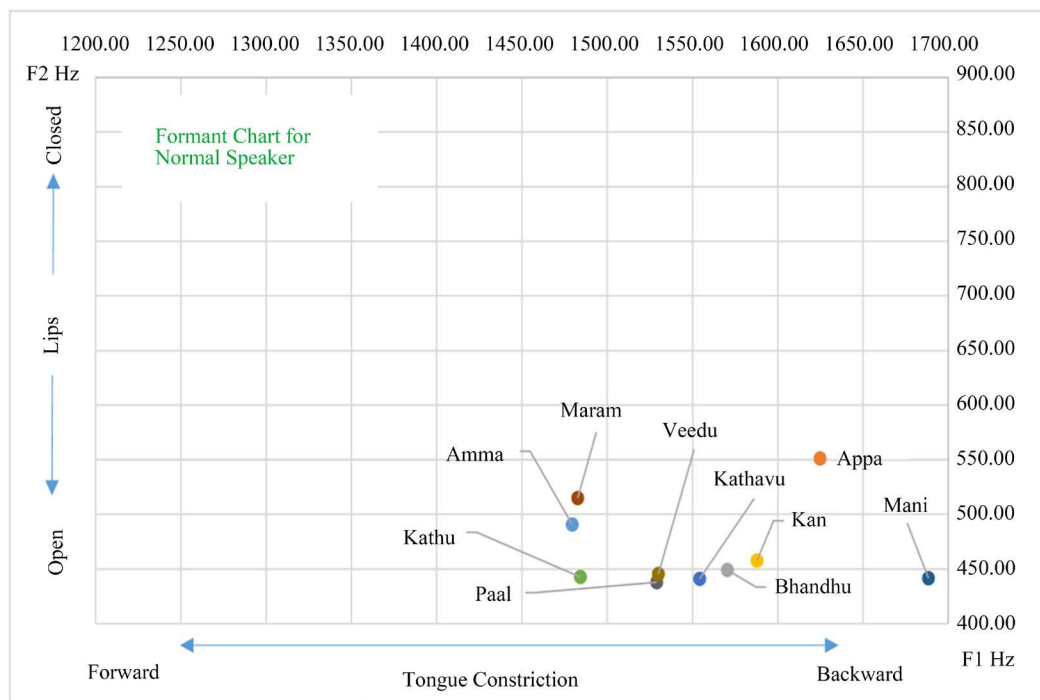


Figure 2. Formant chart for normal speakers.

The formant chart is obtained by placing the F1 and F2 formants in the graph in order to observe the lips movement and tongue constriction positions of the speakers. Figure 2 and Figure 3 illustrate that all the words are present in the Lips open tongue backward constriction region for NS group whereas none of the words of DS group is seen in the Lips open tongue backward constriction region but alternatively all the words of normal speakers are present in that region only. Based on the formant chart the word positions for the normal and deaf speakers are obtained and shown in Figure 4.

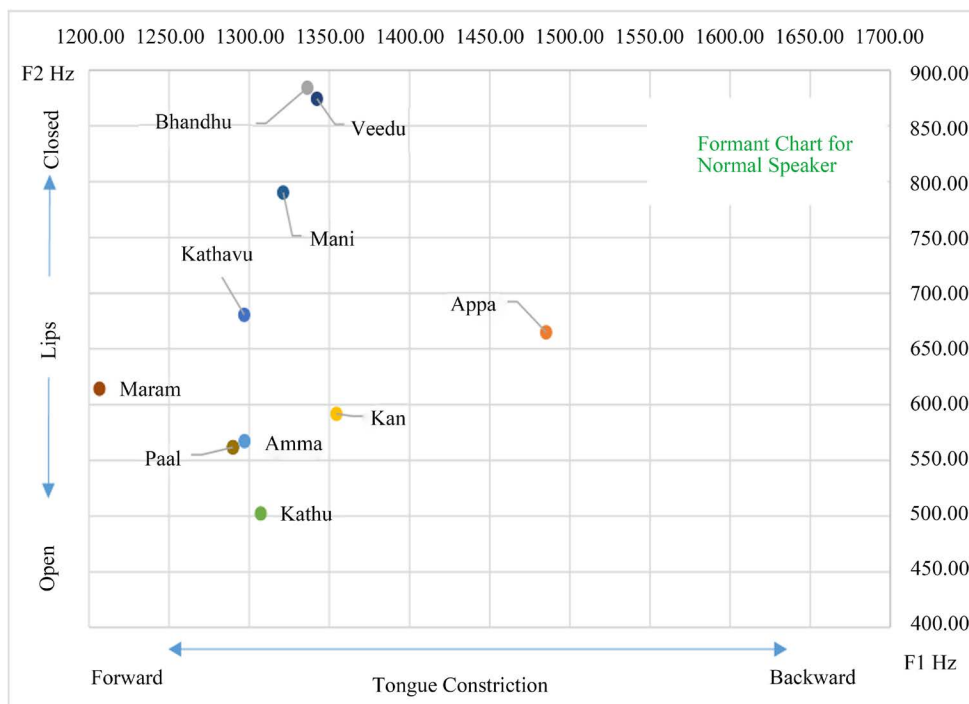


Figure 3. Formant chart for deaf speakers.

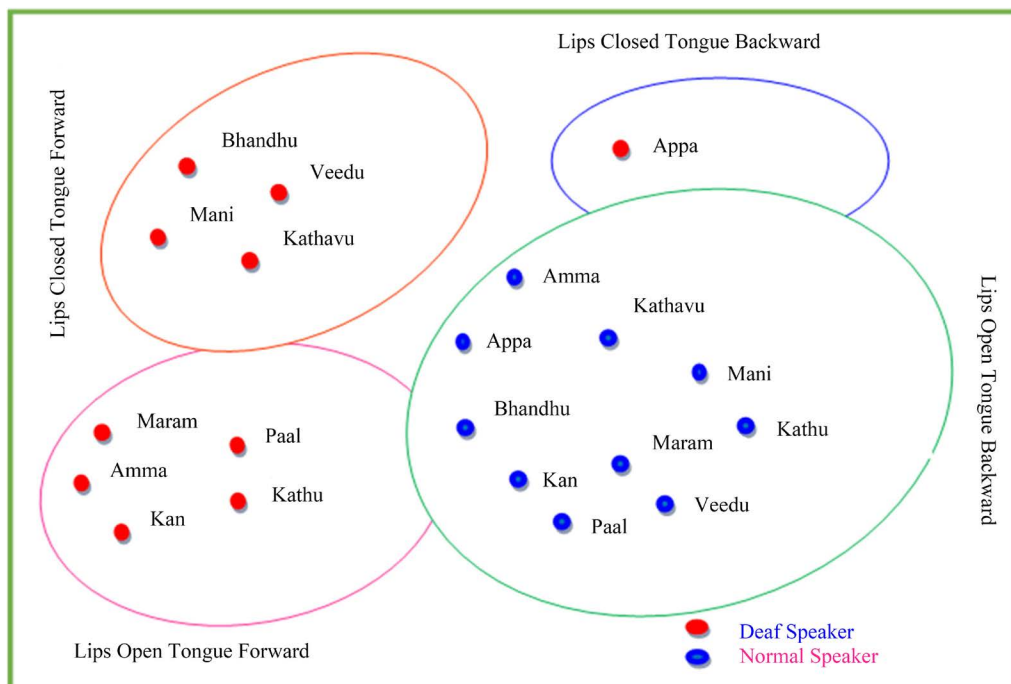


Figure 4. Word positions of normal and deaf speakers.

The first five formant values (F1 to F5) are shown in Table 3 for NS and DS group. From the table, the data shows that the formant values appear deviant between the groups.

When comparing the speaker results the following is understood namely:

Pitch range is higher for deaf speakers and lower for normal speakers,

F1 range is higher for deaf speakers and lower for normal speakers,

Table 3. Descriptive analysis of normal and deaf speaker formant analysis.

Test Words	Subjects	F1 (Hz)	F2 (Hz)	F3 (Hz)	F4 (Hz)	F5 (Hz)
Amma	NS	490.70	1479.40	2427.89	3391.45	3614.40
	DS	567.33	1297.06	2356.31	3664.95	4002.53
Appa	NS	551.20	1624.75	2333.80	3424.53	3494.56
	DS	665.01	1485.25	2205.74	3778.42	4102.27
Bhandhu	NS	449.01	1570.50	2500.76	3689.85	3530.04
	DS	884.29	1336.20	2400.10	3891.98	3940.92
Kan	NS	457.85	1587.78	2327.31	3383.48	3693.54
	DS	591.88	1354.35	2319.90	3631.61	4010.70
Kathavu	NS	441.05	1554.16	2441.40	3325.86	3520.50
	DS	680.58	1296.87	2469.16	3710.54	3893.62
Kathu	NS	442.63	1484.21	2428.10	3435.88	3791.28
	DS	502.44	1307.20	2285.51	3960.12	4260.13
Mani	NS	441.50	1688.29	2379.13	3400.01	3499.98
	DS	790.42	1321.18	2398.96	3630.57	4131.07
Maram	NS	514.93	1482.67	2446.16	3611.28	3422.91
	DS	614.42	1206.44	2384.25	3919.66	4223.65
Veedu	NS	437.85	1529.08	2365.50	3396.18	3595.23
	DS	561.88	1289.79	2285.38	3802.51	3989.20
Paal	NS	445.52	1530.04	2311.22	3567.56	3997.04
	DS	874.45	1342.28	2503.67	3780.93	4121.28

F2 range is higher for normal speakers and the same was reduced in deaf speaker,

F3 range is higher for normal speakers and lower for normal speakers,

F4 range is higher for normal speakers and lower for normal speakers and finally

F5 range is higher for deaf speakers and lower for normal speakers.

The spectrogram mentioned with various levels of formants is mentioned in **Figure 5**. The first line indicates the first formant and it is correlated with the height of the vocal tract. The deaf speaker vocal tract height is higher than the normal speaker which is justified by the F1 values. The first formant values for the normal speaker subject (child2_sample1) for the words “Kan, Kathu and Kathavu” are 457.65 Hz, 442.63 Hz and 441.05 Hz respectively. The deaf speaker first formant values for the subject (child1_giri_sample1) are 591.88 Hz, 502.44 Hz and 680.58 Hz. Hence it is proved that the vowel height is higher for deaf speaker.

The energy level can also be seen in the spectrogram which is dark in colour. In addition the energy level of deaf speaker is heavy compared to normal speaker and the energy is seen at particular location where as in normal speaker the same is spread on various levels. The reason for the accumulation of energy is due to the minimized range of articulators like lips, jaw and tongue positions in deaf speakers. The energy level of deaf speaker is missing at the ranges 2809.90 Hz, 4205.12 Hz and 4225.54 Hz for the words “Kan, Kathu and Kathavu” respectively. Normally the deaf speakers have to put more efforts in order to pronounce a word which in turn reflects in their low intelligibility and have major differences in their speech characteristics. It is well noticed that the energy levels for formants are spread on some specific positions.

The spectrogram analysis indicates that the formant levels are good in normal speaker and it is deteriorated in deaf speakers. The second formant range is reduced in frequency and time in deaf speaker compared to the normal. The first and second formants value of normal speaker for the one syllable word “Kan” is child2_sample1, 457.85 Hz (F1) and child2_sample1, 1587.78 (F2). For deaf speaker the formant values are child1_giri_sample1, 591.88 Hz (F1) and child1_giri_sample1, 1354.35 Hz (F2) which is reduced by 367.46 Hz by frequency. The same is reduced by 236.82 Hz for the two syllable word “Kathu” and by 496.61 Hz for the three syllable word “Kathavu”.

Table 2 shows the distribution indexes of frequency and amplitude namely the Jitter and Shimmer values of

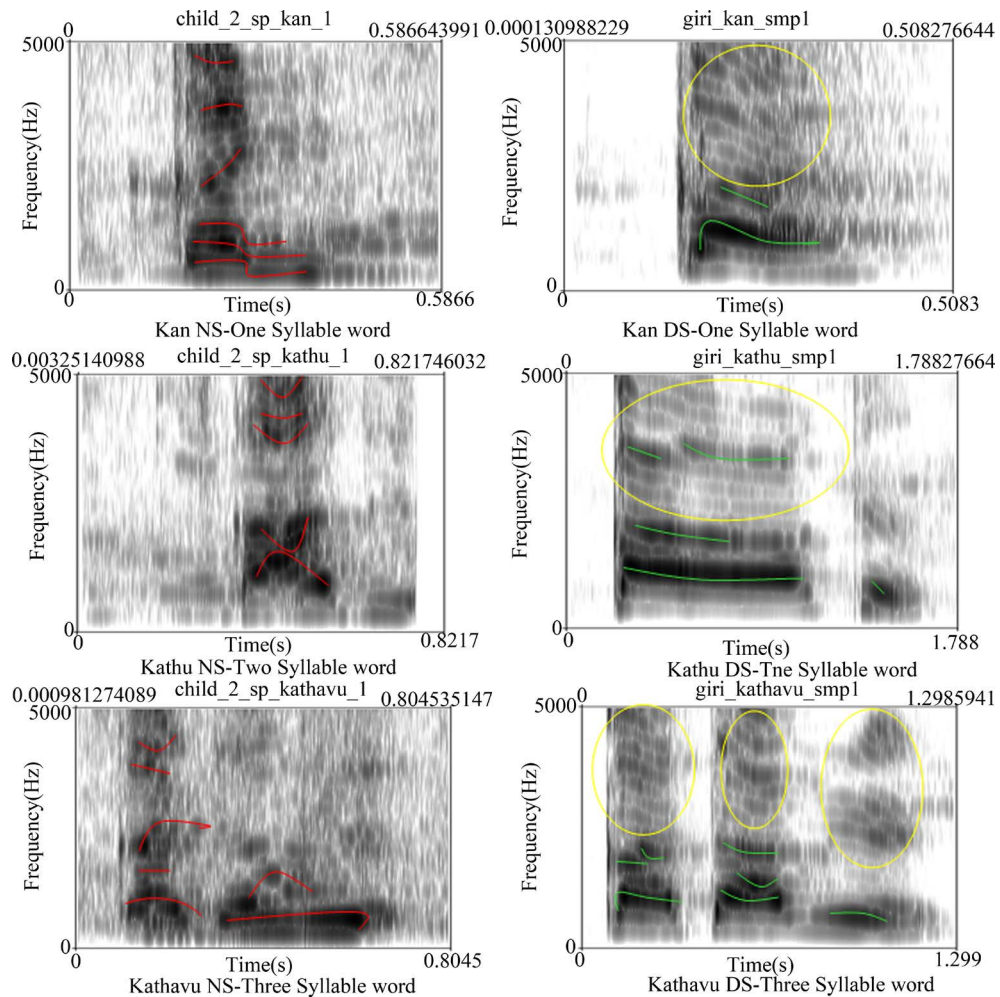


Figure 5. Spectrogram, formant and energy levels.

NS and DS groups. The Jitter values for the words of DS group are found greater than the corresponding NS group samples and the variation is found in the two syllable words. The box plots for the jitter and shimmer comparison of NS and DS groups is shown in [Figure 6](#) and [Figure 7](#). From the box plot is understood that both jitter and shimmer values have greater variation in the DS group. The positive quartile is negligible for jitter in DS group, whereas the negative quartile is higher in both jitter and shimmer which determines that the deaf speakers have more dissimilar values and ranges for the distribution indexes of amplitude and frequency.

It is understood that the cycle-to-cycle variation in the pitch (Jitter) and amplitude (shimmer) is maximum for deaf speakers. The Jitter and shimmer values for the Tamil words chosen are $<1.98\%$ and $<3.83\%$ respectively for the normal speakers. In deaf speakers, the same is greater. The results show contradictory values among the two groups and the values are never similar for both Jitter and Shimmer disturbance indexes.

Intensity problems are mainly due to loudness and stress during the voice production of deaf speakers. Specifically the intensity for one syllable word “Kan” ranges in low intensity than the two and three syllable words and it has no fluctuations. The word “Kan” is totally a consonant whereas “Kathu” and “Kathavu” is mixed with consonant and vowel at the end. The intensity fluctuations in deaf speakers occurred when there is a transition from consonant-to-vowel or from vowel-to-consonant which is indicated in [Figure 8](#).

The intensity contour of the deaf speakers as shown in Figure is highly fluctuating between each consonant to vowel or vowel to consonant transitions. Due to this, the energy levels of the deaf speaker may have more variations. The distribution indexes namely the jitter and shimmer of deaf speaker are with greater values which reflect in their amplitude and frequency parameters. The deaf speakers have unusual voice quality and they often have inadequate pitch frequency, which further affects all the other speech parameters.

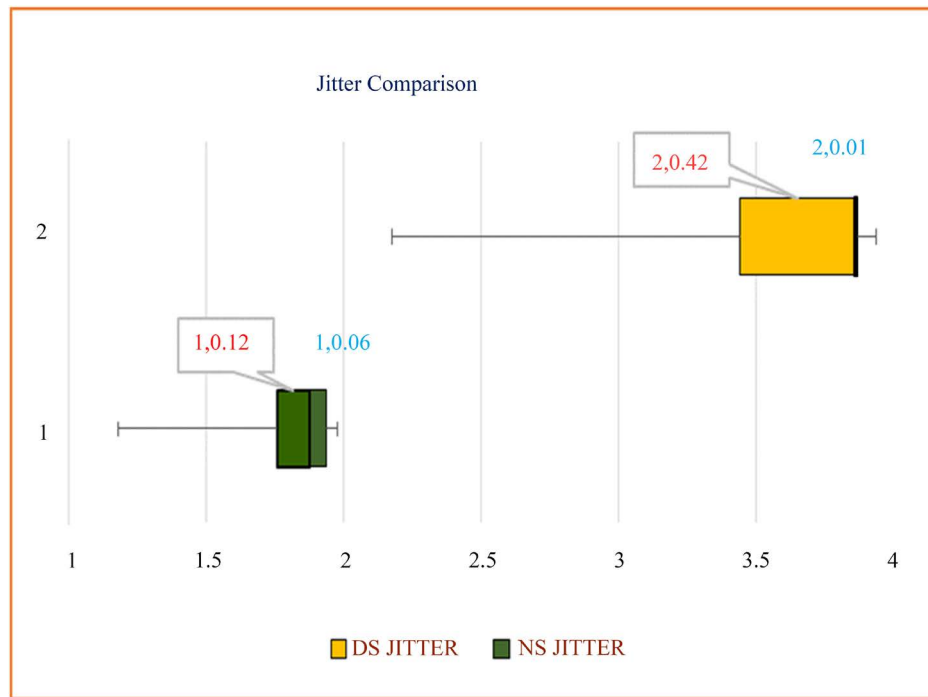


Figure 6. Box plot for jitter.

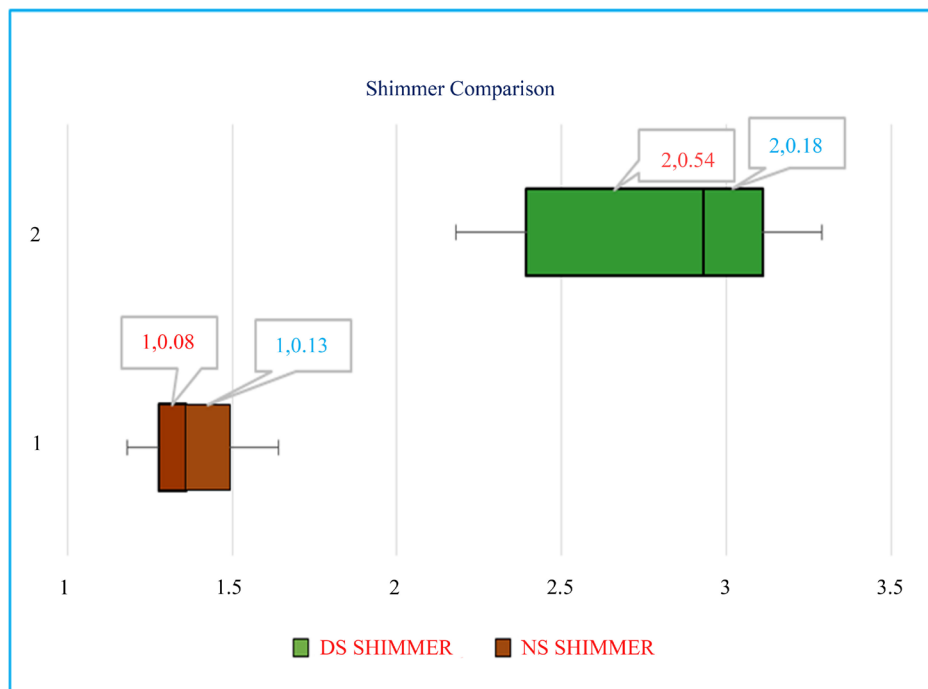


Figure 7. Box plot for shimmer.

The SNR value of the groups is depicted in Table 4. It is understood that the deaf speakers have more noise in the background compared to normal speakers. The maximum amount of noise is seen in three syllable word "Kathavu". The reason for the background noise to be higher in deaf speaker is due to the fact that they often misplace their articulators during speech production and in addition they have stress with breathing problems during the word utterances.

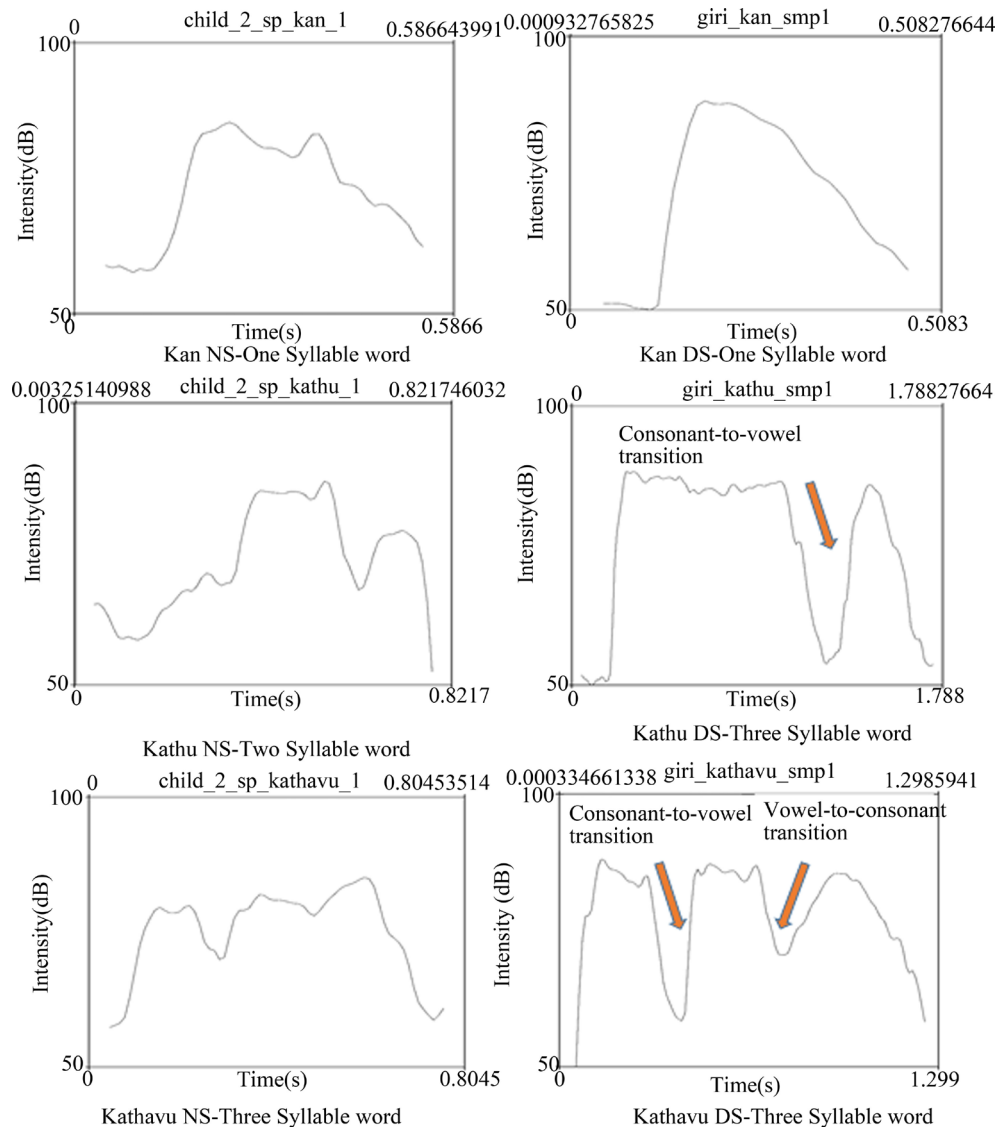


Figure 8. Intensity contours of normal and deaf speakers.

Table 4. SNR Values between normal and deaf group.

Test Words	SNR Values (dB)		No of Syllables
	NS	DS	
Amma	16.77	21.76	2
Appa	13.91	19.88	2
Bhandhu	15.50	22.35	2
Kan	15.06	19.35	1
Kathavu	17.73	23.81	3
Kathu	16.18	22.60	2
Mani	16.53	21.65	2
Maram	14.97	21.45	2
Veedu	18.24	22.84	2
Paal	13.82	16.44	1

4. Conclusion

The comparative study on normal and deaf speech signal shows that the quantitative measures, pitch, F1, F3, F4, F5 are higher in deaf speakers than normal speakers, whereas F2 is lower in deaf speakers than normal speakers. The jitter, shimmer and frequency fluctuations are higher in deaf speakers. Also a significant difference in speech parameters between the two groups except F2 is noticed. Further it is understood, the magnitude of deviation in time domain gives an insight on “depth” and in frequency domain on “deviation”. Based on this information, a correction factor may be estimated, so as to correct the quantitative measures of deaf speech signal. This may result in production of signal that acceptably correlates with normal speech production. Such method may substitute onerous training modules by producing good quality speech in deaf speaker’s tone of voice.

References

- [1] Gold, T. (1980) Speech Production in Hearing-Impaired Children. *Journal of Communication Disorders*, **13**, 397-418. [http://dx.doi.org/10.1016/0021-9924\(80\)90042-8](http://dx.doi.org/10.1016/0021-9924(80)90042-8)
- [2] Archbold, S., Ng, Z.Y., Harrigan, S., Gregory, S., Wakefield, T., Holland, L. and Mulla, I. (2015) Experiences of Young People with Mild to Moderate Hearing Loss: Views of Parents and Teachers. The Ear Foundation Report to NDCS: Mild-Moderate Hearing Loss in Children, 1-12, May 2015.
- [3] Errede, S. (2016) The Human Ear-Hearing, Sound Intensity and Loudness Levels (2002-2015). Department of Physics, University of Illinois at Urbana-Champaign, Illinois, 1-34.
- [4] World Health Organization (2015) Hearing Loss Due to Recreational Exposure to Loud Sounds. A Review.
- [5] Montag, J.L., AuBuchon, A.M., Pisoni, D.B. and Kronenberger, W.G. (2014) Speech Intelligibility in Deaf Children after Long-Term Cochlear Implant Use. *Journal of Speech, Language, and Hearing Research*, **57**, 2332-2343. http://dx.doi.org/10.1044/2014_JSLHR-H-14-0190
- [6] Eriks-Brophy, L. and Tucker, S.-K. (2013) Articulatory Error Patterns and Phonological Process Use of Preschool Children with and without Hearing Loss. *The Volta Review*, **113**, 87-125.
- [7] Parkhurst, B.G. and Levitt, H. (1978) The Effect of Selected Prosodic Errors on the Intelligibility of Deaf Speech. *Journal of Communication Disorders*, **11**, 249-256.
- [8] Osberger, M.J. and Levitt, H. (1979) The Effect of Timing Errors on the Intelligibility of Deaf Children’s Speech. *The Journal of the Acoustical Society of America*, **66**, 1316-1324.
- [9] Boothroyd, A., Eran, O. and Hanin, L. (1996) Speech Perception and Production in Children with Hearing Impairment. *Auditory Capabilities of Children with Hearing Impairment*, 55-74.
- [10] Osberger, M.J. and McGarr, N.S. (1982) Speech Production Characteristics of the Hearing Impaired. *Speech and Language*, **8**, 221-283. <http://dx.doi.org/10.1016/b978-0-12-608608-9.50013-9>
- [11] Monsen, R.B. (1983) Voice Quality and Speech Intelligibility among Deaf Children. *American Annals of the Deaf*, **128**, 12-19. <http://dx.doi.org/10.1353/aad.2112.0015>
- [12] Levitt, H. (1972) Acoustic Analysis of Deaf Speech Using Digital Processing Techniques. *IEEE Transactions on Audio and Electroacoustics*, **20**, 35-41. <http://dx.doi.org/10.1109/TAU.1972.1162351>
- [13] Maassen, B. and Povel, D.J. (1985) The Effect of Segmental and Supra segmental Corrections on the Intelligibility of Deaf Speech. *Journal of Acoustic Society of America*, **78**, 877-886. <http://dx.doi.org/10.1121/1.392918>
- [14] Jeyalakshmi, C., Krishnamurthi, V. and Revathi, A. (2014) Development of Speech Recognition System for Hearing Impaired in Native language. *Journal of Engineering Research*, **2**, 81-99. <http://dx.doi.org/10.7603/s40632-014-0006-z>
- [15] Lien, Y.-A.S., Michener, C.M., Eadie, T.L. and Stepp, C.E. (2015) Individual Monitoring of Vocal Effort with Relative Fundamental Frequency: Relationships with Aerodynamics and Listener Perception. *Journal of Speech, Language, and Hearing Research*, **58**, 566-575. http://dx.doi.org/10.1044/2015_JSLHR-S-14-0194
- [16] Nickerson, R.S. and Stevens, K.N. (1975) Teaching Speech to the Deaf: Can a Computer Help? *IEEE Transactions on Audio and Electroacoustics*, **21**, 445-455. <http://dx.doi.org/10.1109/TAU.1973.1162508>
- [17] Wankhede, N.S. and Shah, M.S. (2013) Investigation on Optimum Parameters for LPC Based Vocal Tract Shape Estimation. *Proceedings of IEEE International Conference Emerging Trends in Communication, Control, Signal Processing & Computing Applications (C2SPCA)*, Bangalore, 10-11 October 2013, 1-6.
- [18] Bernstein, L.E., Ferguson III, J.B. and Goldstein Jr., M.H. (1986) Speech Training Devices for Profoundly Deaf Children. *Proceedings of Acoustics, Speech, and Signal Processing, IEEE International Conference on ICASSP’86*, **11**, 633-636. <http://dx.doi.org/10.1109/ICASSP.1986.1169018>
- [19] Ertmer, D.J. and Maki, J.E. (2000) A Comparison of Speech Training Methods with Deaf Adolescents. Spectrographic

- versus Noninstrumental Instruction. *Journal of Speech, Language, and Hearing Research*, **43**, 1509-1523. <http://dx.doi.org/10.1044/jslhr.4306.1509>
- [20] Stacey, P.C. and Summerfield, A.Q. (2007) Effectiveness of Computer-Based Auditory Training in Improving the Perception of Noise-Vocoded Speech. *The Journal of the Acoustical Society of America*, **121**, 2923-2935. <http://dx.doi.org/10.1121/1.2713668>
- [21] Al-Rousan, M., Assaleh, K. and Tala'a, A. (2009) Video-Based Signer-Independent Arabic Sign Language Recognition Using Hidden Markov Models. *Applied Soft Computing*, **9**, 990-999. <http://dx.doi.org/10.1016/j.asoc.2009.01.002>
- [22] Pichora-Fuller, M.K. and Levitt, H. (2012) Speech Comprehension Training and Auditory and Cognitive Processing in Older Adults. *American Journal of Audiology*, **21**, 351-357. [http://dx.doi.org/10.1044/1059-0889\(2012\)12-0025](http://dx.doi.org/10.1044/1059-0889(2012)12-0025)
- [23] Ferguson, M.A., Henshaw, H., Clark, D.P. and Moore, D.R. (2014) Benefits of Phoneme Discrimination Training in a Randomized Controlled Trial of 50- to 74-Year-Olds with Mild Hearing Loss. *Ear and Hearing*, **35**, e110-e121. <http://dx.doi.org/10.1097/aud.0000000000000020>
- [24] Shafiro, V., Sheft, S., Kuvadia, S. and Gygi, B. (2015) Environmental Sound Training in Cochlear Implant Users. *Journal of Speech, Language, and Hearing Research*, **58**, 509-519. http://dx.doi.org/10.1044/2015_JSLHR-H-14-0312
- [25] Rakesh, K., Dutta, S. and Shama, K. (2011) Gender Recognition Using Speech Processing Techniques in LABVIEW. *International Journal of Advances in Engineering & Technology*, **1**, 51-63.
- [26] Ohman, T. (1998) An Audio-Visual Speech Database and Automatic Measurements of Visual Speech. Quarterly Progress and Status Report, Department of Speech, Music and Hearing, Royal Institute of Technology, 1-2.
- [27] Rajamohan, A., Hemavathy, R. and Dhanalakshmi, M. (2013) Deaf-Mute Communication Interpreter. *International Journal of Scientific Engineering and Technology*, **2**, 336-341.
- [28] Shrote, S.B., Deshpande, M., Deshmukh, P. and Mathapati, S. (2014) Assistive Translator for Deaf & Dumb People. *International Journal of Electronics Communication and Computer Engineering*, **5**, 87-89.