International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878, Volume-8 Issue-3, September 2019

Use of Preprocessor with Frequency Amplitude Modulation Encoding (Fame) for Speech Quality Enhancement in Cochlear Implant to Aid Hearing Impairment

Deepti Gupta, Pratistha Mathur, Peeyush Tewari

Abstract: ***Cochlear Implant is a powerful technical device that facilitates hearing sensation to aid hearing in deaf or partially deaf persons. In this paper a novel method of adopting a preprocessor and using its output for Frequency Amplitude Modulation Encoding (FAME) technique has been studied. This enables filtering of noise from speech before sending it to the electrodes of Cochlear Implant. Preprocessor with FAME gives better results when compared to other techniques which are used for noise removal from the speech for Cochlear Implant users. Also this technique gives a completely new and better speech perception to the Cochlear Implant (CI) users.***

Index Terms: ***Cochlear Implant (CI), Independent Component Analysis (ICA), Speech perception, Hearing in Noise Test (HINT), Frequency Amplitude Modulation encoding (FAME).***

I. INTRODUCTION

Background noises coming from different types of environmental situations cause considerable stress in the life of a deaf person who is using a Cochlear Implant (CI). However, it is known that it is not possible to implement any one type of technique to overcome all kinds of noisy situations. Still focus is on establishing the right technique for all kind of noise condition. This noise removal technique was used with Frequency Amplitude Modulation Encoding (FAME). "Independent Component analysis (ICA) is a method in signal processing for separating multivariate signal into additive sub-components. In this method, the speech is composed of signals which are coming from different locations at each interval of time.

For speech improvement, especially with intelligibility single microphone algorithms are not as much useful as required for CI users.

Manuscript published on 30 September 2019
* Correspondence Author

Deepti Gupta*, Computer Science, Banasthali Vidyapith University, Jaipur, India. Email: 29.deepti@gmail.com

PratisthaMathur, Information Technology, Manipal University Jaipur, Jaipur, India. Email: pratistha.mathur@jaipur.manipal.edu

Peeyush Tewari, Applied Science, Birla Institute of Technology, Noida, India. Email: peeyushtewari@gmail.com

Thus, algorithm with minimum two microphones is required to improve speech with intelligibility in different kinds of noises. Initially in all algorithms like wavelet just did 'denoisation' of the noisy signal to enhance the speech quality which gave some good results. Still improvements are required which give good results in different noisy conditions [1, 2, 3].

II. HISTORICAL BACKGROUND

The objective of noise reduction is to provide better sound to the CI user for carrying out those routine jobs which involve sounds. In Italy the CI related work was started in 1790 by Alessandra Volta and the first electrically stimulated CI was developed by Djourno and Eyries in Paris in 1957. In 1967 Graeme Clark also started working in cochlear implant (CI) area. In 1977 Jim Patrick and Ian Forster made the first circuit for silicon chip. In 1978 Rod Saunders was the first user of CI with multichannel. In 1979 companies like Nucleus and Cochlear took interest in CI work. In 1980, another user was found to use a CI without magnet headset. Finally, in 1982 the first nucleus implant CI22 came up and in 1985 it got FDA approval. CI22 is still in use till date. Currently, CI with FDA approval is being marketed by three most popular companies, namely, Advanced Bionics, Nucleus Freedom and ME-DEL. These companies are manufacturing more than one type of processors and internal implants based on user's needs [4, 5, 6, 7].

Several kinds of test have been used to assess the speech quality in different kinds of noise. Many changes happened in the history and the best test is Hearing in Noise Test (HINT) sentences. HINT sentences are used basically to test the intelligibility present in the speech in noisy situations [8].

III. ICA FUNCTIONALITY:

Independent Component Analysis (ICA) has been used for analyzing multi-variate data. It is used for linear decomposition (transform) of the data and is able to find the underlying components and sources mixed in the data in many cases where the classical methods fail. ICA does not work on beam-forming technique.

It worked on mixed signal which are coming from different directions. Thus ICA gives best results when the signal is minimum two dimensional and output qualities get reduced when signal is one dimensional only.

It uses their statistical properties to find the original components. It uses the non-Gaussian structure of the data. It is an unsupervised method in the sense that it takes the input data in the form of a single data matrix. It is data-driven method because by it the measurement of system can be possible without designing different experimental conditions.

It can be used to investigate the structure of the data when suitable hypotheses are not available, or they are considered too constrained or simplistic. Human speech and electrical signals from brain areas are all not normally distributed. The "cocktail party effect" is the best example for it. Let us imagine two people standing in a room and speaking simultaneously. If two microphones are placed in two different places in the room, they will each record a particular linear combination of the two voices. If non-Gaussianity is assumed, then ICA is used and it is able to identify different speakers sound. Thus it is beneficial when need to identify different speakers sound in noisy environment. But sound should be minimum two dimensional [9, 10, 11]. Figure (1) gives details of ICA in a block diagram.

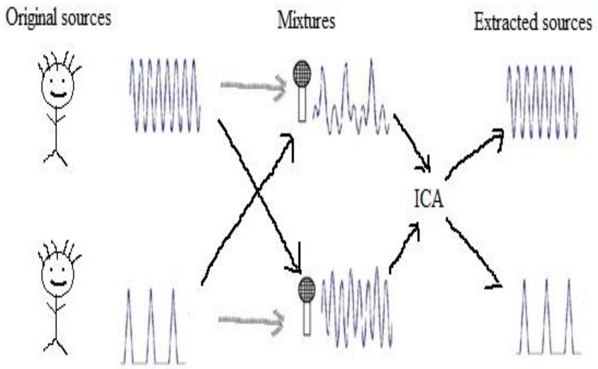

Fig.1 ICA Block diagram

IV. FAME FUNCTIONALITY:

In Frequency Amplitude Modulation Encoding (FAME), amplitude and frequency modulation have been calculated while in Continuous Interleaved Sampling (CIS) only amplitude modulation calculated. Therefore, with FAME technique it is possible to identify the speaker by using the frequency formants [12, 13].

The Frequency amplitude modulation encoding (FAME) separately extracts the slowly varying amplitude AM and frequency FM modulations within each frequency band (Figure2. Block diagram). The FM codes the temporal fine structure of the speech wave form, whereas the AM separately codes the temporal envelope. The instantaneous amplitude of the FM carrier frequency is determined from the temporal envelope in the corresponding band. The signal is first divided into *n* narrow bands. The narrowband signals are then transmitted to separate AM and FM extraction pathways.

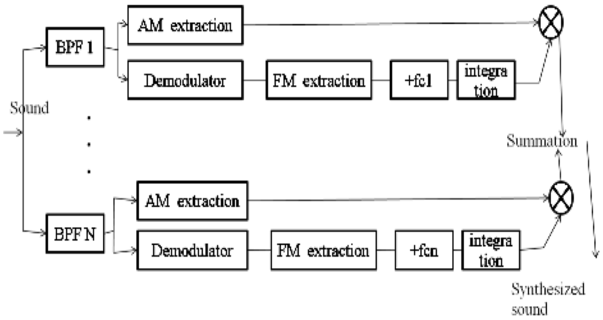

Fig.2 FAME Block Diagram

The AM pathway involves full-wave rectification followed by low-pass filtering at 500 Hz to obtain the slowly varying envelope. The FM pathway involves removal of each narrow band's center frequency followed by low-pass filtering of the FM bandwidth with a cut-off of 500 Hz [14, 15, 16]. Figure (2) gives the details of FAME in a Block diagram.

For speech perception temporal and spectral cues have been used always. In the same way for the extraction of envelope from speech there are three methods half-wave rectification, full-wave rectification and Hilbert transformation. All three methods have been tried in FAME. Out of three methods full-wave rectification and Hilbert transformation give best result. In respect of good envelope detection many variation did in the experiment like changes the cut off frequency and varying the spectral bands systematically [17].

V. DESIGN OF EXPERIMENTS:

The paper mainly focuses on the experiments conducted by taking speech under noisy environments and its reduction using pre-processing activity.

In this study, the main focus is on two important parameters like quality and intelligibility. It is noticed that sometimes when the quality is good the intelligibility gets reduced. Therefore, it is advisable to have some noise level to avoid distortion of intelligibility. This point has been taken into consideration while carrying out the experiments. The experiments have been conducted for 8 noisy environments, such as, bus, train, babble, airport, car, traffic, restaurant, and submarine. Under each of these noise environments, Hearing in Noise Test (HINT) has been used.

All sentence recording has been done by a male speaker in English language. The sentences were recorded digitally on a data acquisition system at 44.1 kHz sampling frequency and using a 16-bit A/D converter in a sound treated room with Avid Pro Tools, Pro Tools11 model. Each CI subject need two hours for an experiment. In case of normal user one hour is sufficient.

A. Training Protocol:

For CI subject training is required to make them familiar with apparatus and protocol of the experiments. New words need to be exposed for CI subjects. All CI subjects are not old enough to understand all words present in the sentences, so it is required. Familiarization of all words by the all subjects is necessary for the correct output of the experiment.

B. Creation of Test Material:

A number of sentences have been taken for recording in which length is uniform and naturalness also present of all sentences. The digitally recorded sentences level is adjusted for the listeners. Thus it seems natural and intelligibility also remains when noise added with speech. Total 9 phonetically balance with 10 sentences each list has been made for adaptive measurement of Hearing in noise test (HINT).

C. *Development of sentence materials:*

All the sentences have equal Mean square amplitude and intelligibility when using as a noise test material. In these sentences noise may not be equal. However, phonetic balance, word familiarity and variation in notation need to be take care so that intelligibility remains present in the sentences along with noise and user is able to identify all words in noise at a fixed signal to noise ratio.

After recording of HINT sentences 8 kinds of noises like bus, train, babble, airport, car, traffic, restaurant, and submarine are added in the sentences with different decibels like -6dB, -5dB, -3dB, 0dB, 3dB, 5dB, 6dB, 10dB SPL. To do all these changes Audacity software has been used.

Experiment was conducted using normal hearing subjects to find out the best output from pre-processor. ICA technique was used for all kinds of noisy environment. The pre-processor output was used as input with CIS and FAME coding technique. The study was done to check the difference between CIS and FAME with pre-processor for different kinds of noises using HINT sentences. Thus, in this experiment evaluation was carried out between CIS and FAME to understand the difference and to use the best technique to the CI users.

Total 30 normal hearing 18 male and 12 female were participated. All subjects screening has been done with the questionnaire test and all are having normal hearing. All subjects age ranged from 20 to 35 year and who were exposed to English language at least for 5 years was selected. All are Graduate and post graduate students. Consent forms have been signed by all subjects.

Cochlear Implant subjects also participated in the experiment. Total 7 participants were present. 2 participants are post-lingual and 5 are pre-lingual. Participants age range from 8 to 34 year with minimum of 2 year cochlear implant experience and unilateral sensorineural hearing loss. All CI participants if adults otherwise their parents have signed consent form.

All subjects were tested at fixed signal to noise ratio with different noise types. PC with digital dual core processor has been used. The output is two channels. SONY headphones of MDR-ZX110A model with 1000 mw amplifier and 98dB/mw presentation level were used as transducers. In case of CI user's loudspeaker of Sony Mega Bass XS-FB162E 6.5-inch Speakers model, 12volts amplifier, 260 watt presentation level is used.

There are 9 sets of HINT sentences and each set contain 10 sentences. So total 90 sentences. Sentences are managed with 8 kinds of noise with different SNR level. In testing random list of sentences has been played. All other things have been taken like the minimum threshold and maximum threshold level of the CI subject so that CI subject will not feel discomfort during experiment. If subject feel comfortable then give correct output. First participants need to hear the sounds. After that whatever listeners were listening, have to write down on a paper if subjects are adults otherwise whatever subjects are listening they have to speak and their parents have to write. The sentences were played at different dB SPL levels like -6dB, -5dB, -3dB, 0dB, 3dB, 5dB, 6dB, 10dB and with different kinds of noises like bus, train, babble, airport, car, traffic, restaurant, and submarine. Presentation levels of sentences were increased by 2dB if subject is not able to understand sentence. Mostly sentences were played between 60 to 70dB SPL for normal users and 70 to 80dB SPL for CI users. Sentences were played for normal hearing participants by both coding techniques Continuous Interleaved Sampling and Frequency Amplitude Modulation Encoding to do the comparison. The percentage calculation has been done on the basis of number of words clearly understand by listener in a sentence. Proper understanding of words in a sentence is a primary measurement which is considered. In that way if a sentence has 7 words then out 7 if listener understand only 5 words then percentage will be 71%. From the results can get which one is better for CI subjects regarding good perception of speech in different kinds of noises at different decibels. Figure (3) gives the bock diagram of pre-processor adapted to FAME technique. In the case of CI users only ICA technique output is played without FAME and CIS.

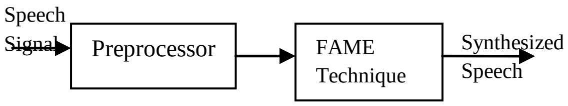

Fig.3 Pre-processor Functionality

By the comparison of the response of the subject in the form of written sentences at different decibels with different types of noises can judge accuracy of sentence understanding. For this mean in table1 has been calculated. For more clarity linear regression also calculated for CIS and FAME individually on the basis of table1 values and can see in figure 6 and 7. Table 2 shows the details of CI users and all users' electrodes are fully inserted. The English word recognition of users also shown in respect of pre-lingual (who are implanted at early age) and post-lingual. Post-lingual users, who are late implanted. For this user's need to listens the words carefully. After that whatever listeners were listening, have to write down on a paper if subjects are adults otherwise whatever subjects are listening they have to speak and their parents have to write.

Table- I: Performance of CIS and FAME in percentage with ICA noise removal technique

Noise Type	-6 dB	-5 dB	-3 dB	0 dB	3 dB	5 dB	6 dB	10 dB
Airport CIS	73.93	67.43	81.83	87.00	87.00	82.50	93.20	94.80
Airport FAME	74.50	61.43	83.50	86.33	88.83	83.17	94.33	94.23
Bus CIS	43.83	70.67	85.14	88.72	93.39	91.58	96.33	97.50

Bus FAME	52.17	72.73	88.20	90.77	92.07	90.70	97.80	98.50
Car CIS	79.77	73.00	75.45	82.50	92.00	88.93	61.72	96.67
Car FAME	79.10	72.33	77.24	88.93	93.80	84.17	65.10	96.67
Babble CIS	63.57	59.17	79.00	71.67	96.00	80.83	75.83	96.33
Babble FAME	65.40	60.00	80.67	69.10	93.33	82.50	79.83	96.00
Train CIS	1.48	0.00	0.00	17.78	10.00	18.70	14.44	52.96
Train FAME	1.48	0.00	0.00	18.52	10.74	21.48	17.26	53.14
Traffic CIS	1.69	14.07	20.00	36.66	55.00	61.11	52.41	62.86
Traffic FAME	2.32	14.57	27.86	36.21	63.28	63.21	56.55	64.64
Restaurant CIS	0.69	0.69	2.67	2.07	40.59	37.59	42.21	62.90
Restaurant FAME	1.72	0.00	6.21	2.07	46.31	38.28	49.62	64.10
Submarine CIS	0.00	0.00	14.07	26.29	17.78	43.45	8.57	91.59
Submarine FAME	0.00	0.71	20.00	31.07	25.19	42.07	11.79	91.59

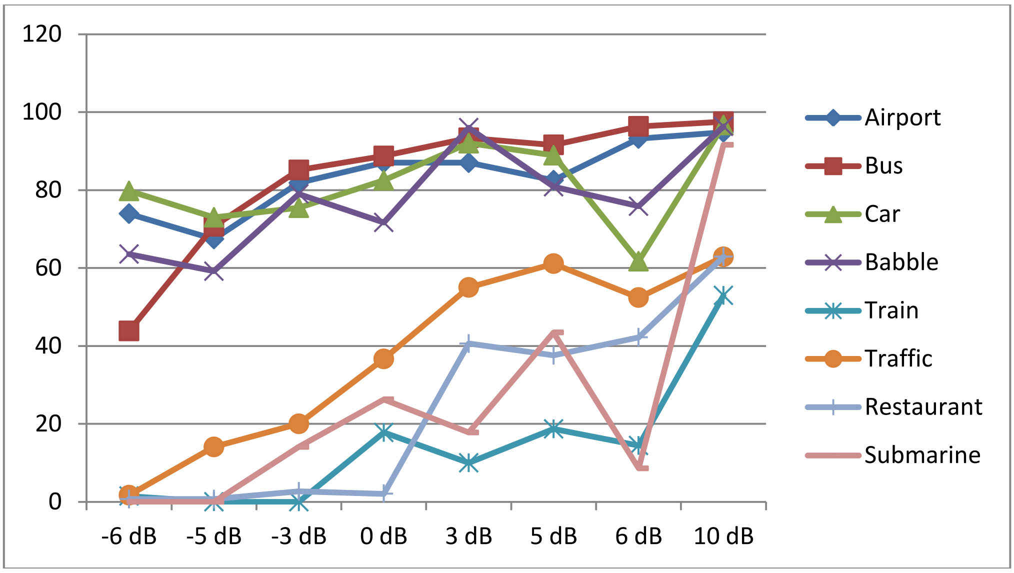

Fig.4 Graph based on the mean values of CIS

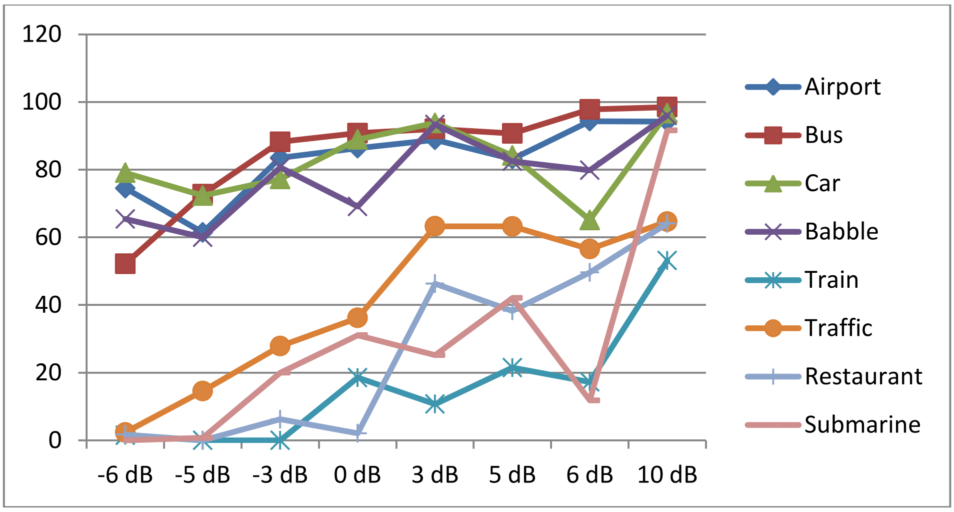

Fig.5 Graph based on the mean values of FAME

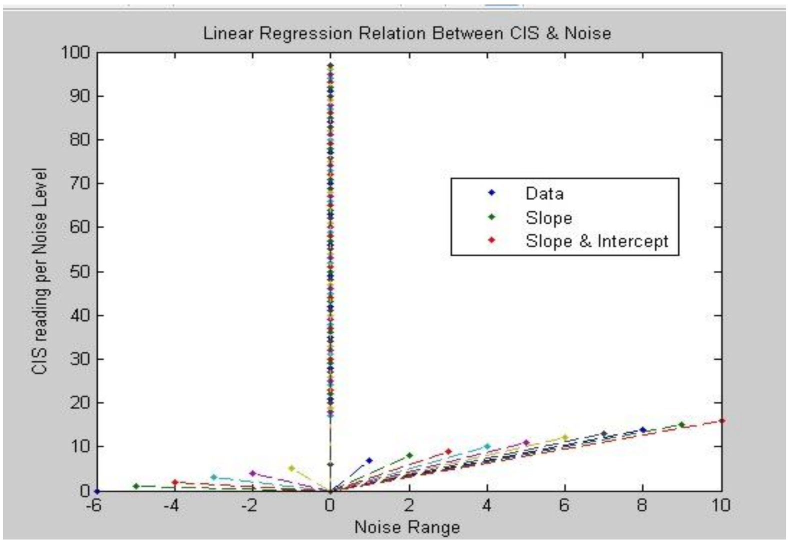

Fig.6 Linear Regression for CIS values

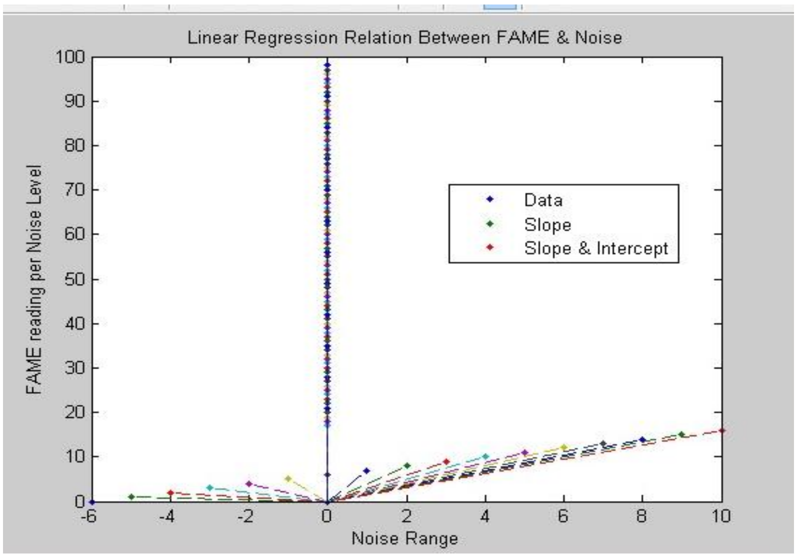

Fig.7 Linear Regression for FAME values

Table II: Cochlear Implant participant's information

S. No.	Company Name	Etiology	Year of hearing loss start	Year of profound hearing loss	Age at implant	Ear	Type	Model type	Gender
MD1	MEDEL	Fever	14	14	15	Right	Unilateral	Opus1	Female
MD2	MEDEL	Fever	20	20	20.2	Right	Unilateral	Opus1	Male
CH3	Cochlear	Unknown	1	11.5	12	Right	Unilateral	CP-802	Female
CH4	Cochlear	Mascular dystrophy	1.5	2.2	2.5	Left	Unilateral	CP-800	Female
CH5	Cochlear	Hereditary	0.5	0.5	2.5	Right	Unilateral	CP-910	Female
CH6	Cochlear	Unkown	1.5	4.5	8	Right	Unilateral	Nucleus freedom 24	Female
CH7	Cochlear	Fever	10	14.5	15	Left	Unilateral	CP-810	Male

VI.RESULTS AND DISCUSSION:

It was proved (table 1) that FAME is better than CIS with respect to clarity of speech with the pre-processor ICA functioning which is used for all kinds of noise (Table 1). In order to obtain more accurate results care was taken to avoid repetition of the same sentences for the same subject [18, 19]. ICA is useful in all noisy situations when signals are two dimensional or more can be seen by the results (table 1 and by figure 4, 5, 6 and 7). It can be seen from the results that -6dB is giving good results as compared to -5dB mostly in all cases. Here multiple of 3 mean the sound in 3dB, 6dB, -3dB, -6dB. The multiple of 5 means 0dB, 5dB, 10dB, -5dB. When sounds are in the form of 0dB, 5dB, -5dB and 10dB, in such cases CIS performance is little better than FAME at some place. Thus, we can say CIS is more sensitive when sounds are in the form of -5, 0, 5, 10dB SNR but FAME is sensitive for all kinds of SNR.

Thus, it can be justifying from the results that multiples of 3 in noisy environment giving good results as compared to multiple of five [20, 21]. In Airport, Bus, Car, Babble signal is two dimensional, so results are good but in train, traffic, restaurant, submarine signal is one dimensional then results are not good can be seen by figures 4 to 7. Thus minimum two microphones are required for good results. From the above experiment it is clear that FAME with pre-processor gave best results for the normal hearing participants so will be best for the CI user too. In the case of CI users only ICA output is played. In this case two users response are neutral who are below 10 years old. Another 5 users response was 20 percent to 80 percent from +5dB to +10dB for Airport, Bus, Car, and Babble type noises. Some users also performed well in -5dB and -3dB too. So the practical implementation of FAME is good for cochlear implant as no company has so far used FAME as a coding technique for CI. English word recognition in quiet results is between 20 to 30 percentages only. CI user feels difficulty in understanding single word can be seen in table3 and figure8 too.

Table III: Clarity of vowels and consonants in percentage

User Type	Clarity in %
Pre-lingual implant at late stage	25
Post-lingual & Pre-lingual implant early age	27.5

A. *Limitation with CI users:*

This study has a few limitations. Initially data size is less due to unavailability of CI users as required for the study. Mostly CI users are kids below 10 years. Kids are not perfect enough for the experiment, as they have small vocabulary and less concentration for the experiment. For the sake of good results from the CI users proper training is required which need minimum 3-4 visits with CI users. It was very difficult to take appointments once from the CI users. Mostly CI users start doing lip reading as they go for higher classes, so their concentration remain only to writing. Due to these reasons, results are not as good as it should be for CI users. CI user does not feel sound as natural like what one feels when a person is speaking in front of him.

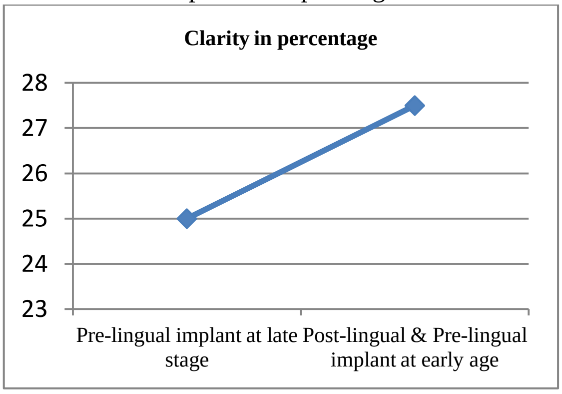

Fig. 8 Difference in percentage for English word recognition

VII. CONCLUSION:

In pre-processor, ICA algorithm gave best results when signal is minimum two dimensional. ICA works well with more than two dimensional data too. Users do feel better when they use FAME instead of CIS. It was also found that FAME is useful when used with pre-processor. The more clarity of these results can be seen by the above tables and figures too. In respect of vowel and consonants, words recognition performance is very less of CI users. Hence FAME can be used with pre-processor to meet the requirement of the user in terms of speech quality and intelligibility.

VIII. ACKNOWLEDGEMENT:

The authors are thankful to the participants for their active participation in these experiments. Our grateful thanks are also due to Dr. Chandrasekhar institute of speech and hearing and Ashadeep ENT center for providing us data of CI users and opportunity to work with them. We are also thankful to DEBEL-DRDO and SBMT (Society for Biomedical Technology) Bangalore for giving us necessary help and support in the form of guidance, infrastructure and facilities for carrying out the work in this area.

REFERENCES:

1. Kostas Kokkinakis and Philipos C. Loizou, Using blind source separation techniques to improve speech recognition in bilateral cochlear implant patients, © 2008 Acoustical Society of America. DOI: 10.1121/1.2839887 PACS numbers: 43.66.Pn, 43.72.Kb, 43.72.Qr DOS Pages: 2379–2390
2. Ningping Fan, Radu Balan, Justinian Rosca, Comparison of Wavelet and FFT Based Single Channel Speech Signal Noise Reduction Techniques,http://www.math.umd.edu/~rvbalan/PAPERS/MyPapers/waveletSPIE04.pdf
3. M.Mehrzad, M. D. Abolhassani,A. H. Jafari, J. Alirezaie, and M. Sangargir, Cochlear Implant Speech Processing Using Wavelet Transform, International Scholarly Research Network ISRN Signal Processing Volume 2012, Article ID 628706, 6 pages doi:10.5402/2012/628706
4. Malcolm Slaney, An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank, Apple Computer Technical Report #35,Perception Group—Advanced Technology Group © 1993, Apple Computer, Inc.
5. P. Vijayalakshmi, T. Nagarajan and Preethi mahadevan, Improving Speech Intelligibility in Cochlear Implants using Acoustic Models, wseas transactions on signal processing, Issue 4, Volume 7, October 2011, ISSN: 1790-5052.
6. Ginger S. Stickney, Peter F. Assmann, Janice Chang, Fan-Gang Zeng, Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences, © 2007 Acoustical Society of America. DOI: 10.1121/1.2750159PACS numbers: 43.66.Ts, 43.71.Ky, 43.71.Bp, 43.66.Hg _KWG_ Pages: 1069–1078
7. Cong-Thanh Do, Acoustic Simulations of Cochlear Implants in Human and Machine Hearing Research, Chapter from the book Cochlear Implant Research Updates Downloaded from: http://www.intechopen.com/books/cochlear-implant-researchupdates
8. Michael Nilsson, Sigfrid D. Soli, and Jean A. Sullivan, Development of the Hearing In Noise Test for the measurement of speech reception thresholds in quiet and in noise, Acoustical society of America, February 1994.
9. Aapo Hyvärinen and Erkki Oja, Independent Component Analysis: Algorithms and Applications, https://www.cs.helsinki.fi/u/ahyvarin/papers/NN00new.pdf.
10. David M. Groppe, Scott Makeig, Marta Kutas, Independent Component Analysis of Event-Related Potentials, cognitive science online, 2008, http://cogsci-online.ucsd.edu/6/6-1.pdfx

11. Dominic Langlois, Sylvain Chartier, and Dominique Gosselin, An Introduction to Independent Component Analysis: InfoMax and FastICA algorithms, Tutorials in Quantitative Methods for Psychology 2010, Vol. 6(1), p. 31-38.
12. Malcolm Slaney, An Efficient Implementation of the Patterson-Holdsworth Auditory Filter Bank, Apple Computer Technical Report #35,Perception Group—Advanced Technology Group © 1993, Apple Computer, Inc.
13. P. Vijayalakshmi, T. Nagarajan and Preethi mahadevan, Improving Speech Intelligibility in Cochlear Implants using Acoustic Models, wseas transactions on signal processing, Issue 4, Volume 7, October 2011, ISSN: 1790-5052.
14. Ginger S. Stickney, Peter F. Assmann, Janice Chang, Fan-Gang Zeng, Effects of cochlear implant processing and fundamental frequency on the intelligibility of competing sentences, © 2007 Acoustical Society of America. DOI: 10.1121/1.2750159PACS numbers: 43.66.Ts, 43.71.Ky, 43.71.Bp, 43.66.Hg _KWG_ Pages: 1069–1078
15. Cong-Thanh Do, Acoustic Simulations of Cochlear Implants in Human and Machine Hearing Research, Chapter from the book Cochlear Implant Research Updates Downloaded from: http://www.intechopen.com/books/cochlear-implant-researchupdates
16. Fan-Gang Zeng, KaibaoNie, Kong, Michael Vongphoe, Ashish Bhargave, Chaogang Wei, and Keli Cao, Speech recognition with amplitude and frequency modulations, PNAS February 15, 2005 vol. 102 no. 7 2293–2298.
17. Kaibao Nie, Amy Barco, and Fan-Gang Zeng, Spectral and Temporal Cues in Cochlear Implant Speech Perception, Ear & Hearing Copyright © 2006.
18. Deepti Gupta, Pratistha Mathur, Peeyush Tewari, Existing Noise Removal Techniques Used In Cochlear Implant, International Journal of Advanced research in Science and engineering, vol. No.5,Issue No. 3,March 2016.
19. Deepti Gupta, Pratistha Mathur, Peeyush Tewari, Signal processing and neural network techniques used in Cochlear Implant for different types of noises, International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 03 Issue: 06 | June-2016 , www.irjet.net p-ISSN: 2395-0072
20. Alan V. Oppenheim, Signals and System, Paperback 2015, Second Edition.
21. Audio wave files for different environmental sounds retrieved from http://www. findsounds.com/.

AUTHORS FROFILE

Deepti Gupta is currently pursuing PhD. She worked 7years in a college affiliated to Indraprastha University as an assistant professor. Her area of Interest is signal processing techniques related to Cochlear Implant.

Pratistha Mathur is currently working in Dept. of Information Technology, Manipal University Jaipur Since July 2016. She served over 18 years at BanasthaliVidyapith, Rajasthan. Her area of interest includes Image Processing, Soft computing, Artificial Intelligence and Indian Language Computing. She is the author of more than 40 research articles and 2 books. She can be reached at mathurprati@yahoo.com

Peeyush Tewari, did MSc, Ph.D (Applied Mathematics), from IIT Roorkee, India. Working at BIT Mersa Ranchi, India since 1991. Worked at BIT International centre in Sultanate of Oman and united Arab Emirates. Research Area includes Moving Boundary Problems: applications in geophysical Modelling, Level set methods in Image processing and Supply Chain Management. BookChapter in Nano-technology 2009: Fabrication, Particles, Characterization, MEMS, Electronics and Photonics, Editorial board member of International Journal of Intelligent Information Processing.