



International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

www.iasir.net

Tamil Phonemes Recognition for Normal and Hearing Impaired: A Survey

Janakiraman.S

Pondicherry University, Puducherry, INDIA

Pavithra.S

Research Scholar, Sathiyabama University Chennai, Tamil Nadu, INDIA

Abstract: This paper presents a survey on speech recognition models and Tamil Phonemes Recognition Model. Speech is human's most efficient mode of communication. Beyond efficiency, humans are comfortable and familiar with speech. Research work on Tamil speech recognition, although lagging than other languages, is becoming more intensive than before and several researches have been published in the last few years. In spoken language, a phoneme is a basic, theoretical unit of sound that can change the meaning of a word. Depending on the language and the sounds used, a phoneme may be written consistently with one letter; however, there are many exceptions to this rule. The range of the possible applications is wide and includes: voice-controlled appliances, fully featured speech-to-text software, automation of operator-assisted services, and voice recognition aids for the handicapped. In this paper it is proposed to develop a Speech Recognition System for both Normal Hearing and Hearing Impaired children with Visual Feedback for Tamil phonemes. The proposed Phoneme Recognition System use Radial Basis Function Networks and visual feedback is a Graphical User Interface to provide the correct (or) incorrect articulation of phoneme visually.

I. Introduction

Speech recognition is one of the key tools for the next-generation user-friendly computer application that aims to simplify the lives of everyday users. Speech recognition is a difficult problem, largely because of the main sources of variability associated with the signal. The various sources of variability were phonetic identity, pitch, speaker, microphone and environment. In this Literature Review, it was understood that most of the Automatic Speech Recognition systems were developed for English and other languages like Arabic using HMM. In Indian Languages, Automatic Speech Recognition systems were developed for Tamil, Hindi, Oriya and Tamil. Those models were using HMM, Wavelet, Neural Network or Bootstrapping. It is found that there is no model found for Hearing Impaired in Tamil language using Radial Basis Function Networks. Tamil Language is a historical Dravidian language spoken by around 52 million people in countries like India, Sri Lanka, Malaysia, Vietnam, Singapore, Canada, USA, UK and Australia. And it is a phonetic language, that is, "words are written as they are pronounced, and pronounced as they are written". The proposed speech recognition system is to recognize the Tamil phonemes such as Stops, Laterals and Trills with vowel combination and responds to user with Visual Feedback based on the accuracy level of the user's speech. Using Neural Network Models, number of phonemes and accuracy level can be further improved. The main contribution of the speech recognition system is for early Tamil language learning and articulation training for Normal Hearing and Hearing Impaired children.

II. Related works

Speech Recognition Model for Tamil Stops was developed by Rathinavelu A, Anupriya G, and Muthanantha Murugavel A S [1]. This paper presents the work on articulatory acoustic data analysis and training of a Neural Network model for recognition of Tamil Phonemes. Acoustic study on Tamil laterals, trills and fricatives was carried out with Normal Hearing and Hearing Impaired children. The visual articulatory model is useful for training and improving the articulation of people who are not familiar with the articulation of phonemes. Previous studies have focused on feed forward neural network which is a biologically inspired classification algorithm. The test data for the system has been acquired from 10 subjects (5 boys and 5 girls in the age group of 4-7 years) for 20 Tamil phonemes. The Recognition % for each phoneme is calculated by taking the ratio of the number of successful recognitions to the total number of test samples. An average accuracy level of 81% has been achieved in the experiments conducted using the trained neural network. The number of phonemes for recognition can be increased by using the feed forward neural network model with back propagation algorithm and the accuracy level can also be further improved by giving more training.

Future work will focus on gender selection in the speech user interface, option for more phonemes and large corpus size. This preliminary study will help to develop Automatic Speech Recognition system for Normal Hearing and Hearing Impaired children by providing visual feedback to improve their articulation.

Interactive multimedia tool to help vocabulary learning of hearing impaired children by using 3D VR objects as visual cues was developed by Rathinavelu A, Hemalatha T, Savithri S R and Chitra [2]. Lip sync model is one of the aspects in computer facial animation. To create realistic lip sync model, a facial animation system needs extremely smooth lip motion with the deformation of the lips synchronized with the audio portion of speech.

The data for lip sync model was collected from video image and magnetic resonance imaging (MRI) techniques they are /p/, /b/, /m/. This IMM interface was used to teach small vocabulary of hearing impaired (HI) children. Virtual reality objects used to increase the cognitive process within HI children. The bilabial speech sounds were differentiated by using appropriate visual cues. In order to minimize the time involvement of teacher, computer aided sub skills were suggested for speech training needs of hearing impaired (HI) children. The speech of hearing impaired children differs from speech of Normal Hearing (NH) children in all aspects. Six hearing impaired children were selected between the ages 4 and 7 and they were trained for 10 hours across 2 weeks on 18 meaningful words such as puli, pul, uppu, kappal, appaa, palam, bimbam, kambam, bakti, ambu, cembu, balam, maram, makkal, miin, buumi, paambu, ammaa. According to deaf researchers, hearing children have consistently demonstrated the ability to perform such tasks between the ages of 4 and 5 years. But hearing impaired children are delayed by approx. 3 years in this cognitive developed milestone. It is also found that hearing impaired children are significantly delayed in the development of language skills. In this it has four modules, they are Basic research module, Development module, Training module, Evaluation module. In the basic research module, a new concept was intimated to design and develop computer aided lip sync model using VR object to help HI children. In development module, data of video image and MRI was used to develop 3D lip sync model. The lips were modeled using 36 control points. In the training phase, 18 meaningful words were presented to 6 HI children for 10 hours over 2 weeks. 3 boys and 3 girls were selected as subjects. In evaluation module, after the 2 weeks of training, children were asked to articulate the speech sounds and syllables on 10th day. The result shows, even after two weeks training, few words were not articulated properly by some of the hearing impaired children in the final test. For example, 'uppu' (salt) was substituted with 'o' and articulated as 'oppu'. This was due to similar kind of lip protrusion perceived by most of the children. Few of them didn't have any control over short vowels. For example, 'pu-li' (Tiger) was articulated as 'puu-li' and 'pu-l' (grass) was articulated as 'puu-l'. The results indicated that 65-75% of given words were articulated well and 75-85% of words were identified well by all children. The articulatory movements of the lip sync model have been presented along with virtual reality (VR) objects in an interactive multimedia (IMM) interface.

Speech Reading of Words and Sentences by Normally Hearing and Hearing Impaired Chinese Subjects: The Enhancement effect of Compound Speech Patterns was developed by Xinghui Hu, Adrian Fourcin, Andrew Faulkner and Jianing Wei [3]. Speech perception requires a receiver to make decisions both about trends and also about the language patterns of a speaker. Hearing people use the speaker's speech signal as the direct sensory evidence. Lexical tones are to be perceptually distinguished from each other by the patterns of voice fundamental frequency. It shows that the auditory supplement with voice fundamental information could enhance speech reading by Chinese speakers. Speech pattern element hearing aids are believed to be able to give substantial information to tone language users. Another missing speech feature in facial gestures is friction. Friction information is largely present in the frequency range above 2000Hz and it is generally inaccessible to profoundly deaf people's residual hearing even when speech intensity is compensated for according to their hearing losses. The low frequency range benefits hearing-impaired people, when they are speech reading. Voice fundamental frequency detection is implemented by using a spectral-balanced method. In the first experiment, 10 normally hearing subjects and 1 hearing impaired person were asked to recognize the isolated words by speech reading using a counter balanced design. 2 NH people were asked to speech read the words with the addition of three different combinations of the voice fundamental frequency, frication and amplitude envelope. 3 HI subjects were required to speech read the word by seeing the face alone. In testing, 40 words, 20 sentences and 20 short passages were developed and recorded for training purpose. 240 Chinese words were selected from the most frequently used 1000 words in Chinese newspapers, books and magazines. A word list consists of 15 monosyllabic words, 40 disyllabic words and 15 trisyllabic words. A female native Chinese speaker from Beijing, China spoke the word and sentence list. Each list was read and the words or sentences in the list came in different order each time. The speaker was recorded by a SONY video camera to a SVHS Hi-Fi NiCam recorder with speech and laryngograph signal recorded to the two audio channels. All the NH subjects were able to extract substantial enhancement effect from the auditory supplement of compound speech pattern element signal, regardless of speech reading ability. For the isolated word recognition, the enhancement effect between supplemented and unsupplemented speech reading was 35.8% and for isolated sentence recognition, the difference was 44.8%. The HI people in this study showed large diversity. Subjects SYC achieved 40.4% for word recognition and 43.1% for sentence recognition. Subject HXF has better amplified speech than SiVoIII but subject GXY gained more help from the supplement of SiVoIII than that from amplified speech. More HI people with profound Hearing losses have to be tested in order to evaluate the speech pattern element method for Chinese deaf people.

Recognition of Voice signals for Oriya Language using wavelet Neural Networks was developed by S. Mohanty et al [4]. Speech recognition is both speech oriented and speaker oriented. Both have fuzzy effect on them which adds to the hardness. They tried with the wavelet neural network model to make an effective analysis for the recognition of speech signals. The main characteristics of speech/voice like frequency, intensity, accent and quality were analyzed for better output limiting the associated noise. Using the control parameter of the network

they have designed the wavelet network for two characters at the beginning and also they were extending for the rest of the character. The wavelet transform theory and its implementation were limited to the wavelets of the small dimension, whereas ANN is a powerful tool for handling the problems of higher dimensions. Combination of both results in Wavelet network. The weakness of each other is compensated and it handled problems of larger dimension and it shown efficient network construction methods. It resulted towards an accurate output signal by eliminating noise to an acceptable level from the desired signal. It took little bit longer time than other processes, but supported non-linear mapping between the input-output pairs.

A speaker recognition system, using a modified form of feed-forward neural network based on radial basis functions, is presented [5]. Each person to be recognized has their own neural model that is trained to recognize spectral feature vectors representative of their speech. The approach is evaluated here with a 40 speaker database for the task of speaker verification. The modified neural structure is shown to significantly outperform both a vector quantization based system and a standard multilayer perception classifier. Additional advantages include extremely quick adaptive training, with at least an order of magnitude improvement over the standard multilayer perception approach. The database used in these experiments consists of 40 speakers (27 male and 13 female). The speakers are divided into three groups. Group 1 has 10 male speakers (both true talkers and imposters), Group 2 has 5 male and 5 female speakers (both true talkers and imposters) and Group 3 has 12 male and 8 female speakers (additional imposters). Each speaker has five versions of digit set spoken in five different manner, they are loud, soft, questioning, angry and normal. The loud and questioning versions are used for training while normal, angry and soft versions are used to evaluate verification performance. The best performance is obtained from RBF network with 384 RBF nodes in the hidden layer, giving an 8% true talker rejection rate for a fixed 1% imposter acceptance rate. Additional advantages include a substantial reduction in training time over an MLP approach and the ability to readily interpret the resulting model. Here it has been shown that voiced sounds play a major role in the classification process.

A novel approach for implementing Arabic isolated speech recognition using Neural Network was developed [6]. While most of Automatic Speech Recognition (ASR) work is based on Hidden Markov Models (HMM), their system was implemented by using modular recurrent Elman neural networks (MRENN). The SR system was modular, i.e. for each word in the vocabulary set; there was a separate Elman network. Modularity adopts a “divide-and-conquer” approach by dividing the complex problem at hand into many smaller and simpler problems. The vocabulary set used 6 Arabic words: “manzel” (house), “hirra” (cat), “chajara” (tree), “tariq” (road), “ghinaa” (singing), “zeina” (zeina). The function of each network was to recognize its dedicated word only and to reject other words. Back-propagation with momentum and variable learning rate was used as training algorithm. Consistent training was performed after discriminative training, since the recurrent networks inherently remembered the most recent training utterance applied to it. After each training pass (100 epochs), the network was simulated on a validation set composed of 40 new utterances of the dedicated word and 80 new utterances from the remaining 5 words. The output obtained from the simulation of the network for an utterance was a non-linear curve. The decision-making criterion was the slope of the line obtained from the linear fitting of this curve. The classification of an utterance, other than the dedicated word, has been based on the comparison of its resulting slope with the minimum slope among the slopes obtained from all the utterances of the dedicated word. The speech recognizer described in the paper has been fully implemented in MATLAB, and was subjected to several test inputs. The promising results obtained through their system have shown that the new neural networks approach can compete with the traditional HMM-based speech recognition approaches.

A large-vocabulary continuous speech recognition system for Hindi”, by M Kumar, N Rajput, and A Verma [7] was trained on 40 hours of audio data and has a trigram language model that has been trained with 3 million words. For a vocabulary size of 65000 words, the system gives a word accuracy of 75% to 95%. They built a trigram language model to capture the statistics of Hindi language. They used Hindi text data from various sources to assign bigram and trigram probabilities to facilitate a probability distribution over all possible sentences. A language model vocabulary was used and probability was assigned to unknown words also. Since the amount of data was not huge, they had assigned probabilities to all the bigrams and trigram occurring even once. The English data used to generate the initial phone models for Hindi consisted of 3000 utterances of 30 different speakers. The English training sentences were chosen from a vocabulary consisting of 96,000 words. This constituted about 7 hours of continuous speech. Hindi data has been collected for 7 speakers having a total of 350 utterances totaling around 30 minutes of continuous Hindi speech. They used a Hindi phonetic vocabulary containing 900 Hindi words. A 24-dimensional mel-cepstral coefficient feature vector has been formed for each audio frame of duration 10 ms. Linear Discriminant Analysis has been used to transform this vector to a 60-dimensional space. The recognition system performs well over a test set that has not been used in training, emphasizing the speaker independence of the system. A full-fledged Hindi phone set that covers all different sounds that occur in Hindi has been defined in this paper.

Language Models For Tamil Speech Recognition System was developed by S Saraswathi and T V Geetha[8]. In this paper, language models are used in different phases of speech recognition such as phoneme segmentation, recognition and error correction.

An Artificial Intelligence based speech recognition approach which is a hybrid of knowledge based approach and statistical approach is designed to recognize speech in Tamil language. Language models can be applied from all levels of granularity, from the phoneme level, to syllable level, to morpheme level and till the word level in the proposed Tamil speech recognition system. The type of language model to be applied depends on the language being used. Current research in speech recognition focuses on methods to improve the performance of the speech recognition system using improved language models. The characteristic of a language decides the type of language model designed for it. Most languages are generally assumed to have a basic word order. For example, in English SVO (subject-verb-object) is the commonly used fixed word pattern for normal type of sentences. Different types of language models such as category-based phoneme language models, inter and intra word language models, syllable and morpheme based language models were used to improve the performance of Tamil speech recognition system. In category-based phoneme language models, This work performs phoneme recognition based on the acoustic features of the phonemes and uses language models to decrease the errors in phoneme recognition phase.

In inter word language model, in this section first describe the use of preceding sequence of phonemes of the word as the context to determine the probability of occurrence of the phoneme under consideration. Since the length of the word in which the phoneme under consideration occurs and the position of this phoneme in the word varies, the sequence length of the preceding phonemes taken as context for the probability prediction of phoneme under consideration varies. Thus the varying length of the preceding phoneme sequence results in an intra-word variable length language model. In intra word language model, This hybrid model essentially predicts the probability of occurrence of the phoneme sequence used as context in variable length language model based on the syllables at the end of the previous word and independent probability of syllables occurring as an ending syllable in a word. In syllable and morpheme based language models.

In this section move from phoneme based language models to language models based on larger units such as syllables and morphemes. In the previous section, obtained a probable sequence of phonemes corresponding to a word. In this section, Use specially designed language models for error correction based on probability of syllables and morphemes in the formation of words. Recognition rate of 87.1 % was obtained by applying language models at various phases of recognition. The speech recognition system is designed for Tamil speech corpus transcribed at the word level. The system can be further enhanced, by modifying it for continuous speech recognition. Design of language models at the sub-word level to identify wrongly segmented words can further improve the performance of continuous speech recognition for the Tamil language. A combination of the morpheme based language model and the syntactic and semantic class based models can further improve the error correction rate in the Tamil speech recognition system, due to the inflectional characteristics of Tamil.

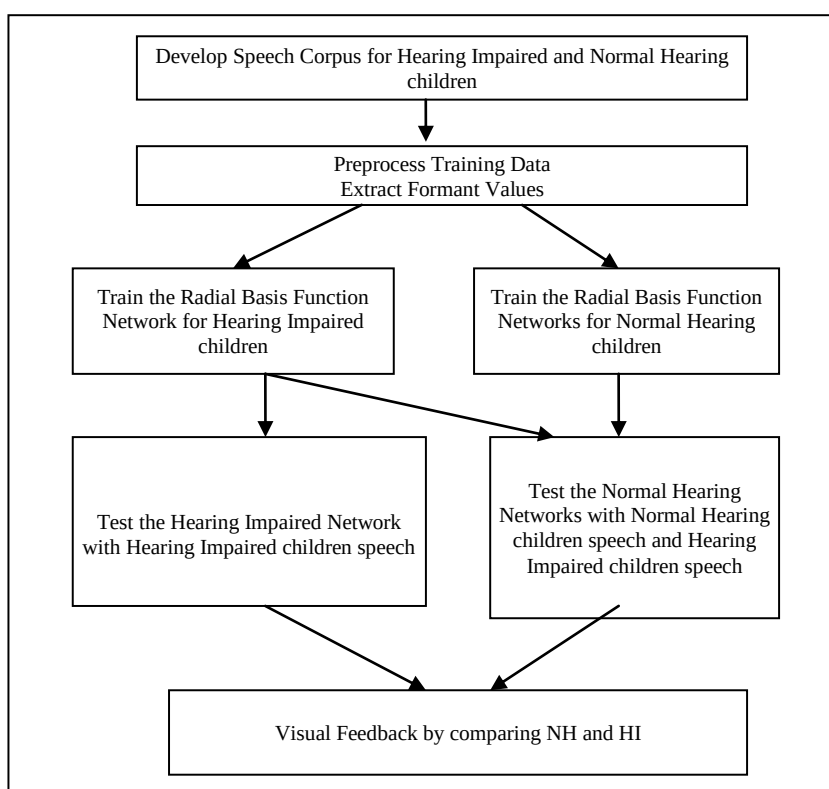


Figure 1: Arc of speech Recognition system.

III. ARICHITECTURE

The methodology of the design and development of Speech Recognition System is shown in Fig first develop the speech corpus for Normal Hearing children and Hearing Impaired children. Then these network gets the Formant values for each speech sounds in Preprocessing phase.

Planned to build a database of of 25 Tamil phonemes such as 4 stops (k, t, t and p) and 1 Alveolar Lateral (l) with vowel (a, e, i, o, u) combination from few children (5 to 10 boys and 5 to 10 girls from Normal Hearing and 5 to 10 boys and 5 to 10 girls from Hearing Impaired) in the age group of 4-7 years. This speech corpus will be used for training and testing the neural networks. First, training will be given to the Hearing Impaired children and Normal Hearing children for the selected Tamil phonemes for few days. Then network is trained for these features using the first two formant values constituting the input vectors to the Radial Basis Function network (RBF) used as classifier. The Neural Networks Model for Tamil Phonemes Recognition will be designed using Radial Basis Function Network. The individual networks to train both Normal Hearing and Hearing Impaired will be designed. I have also planned to do Gender wise testing for Normal Hearing network with Normal Hearing and Hearing Impaired speech sound and testing Hearing Impaired network only with Hearing Impaired speech sound for the selected Tamil phonemes & also to design Speech User Interface to load, play, record, save, view waveform and to extract formants for the speech sounds and it provides visual feedback based on the accuracy level of the recognition.

IV. MODULES DESCRIPTION

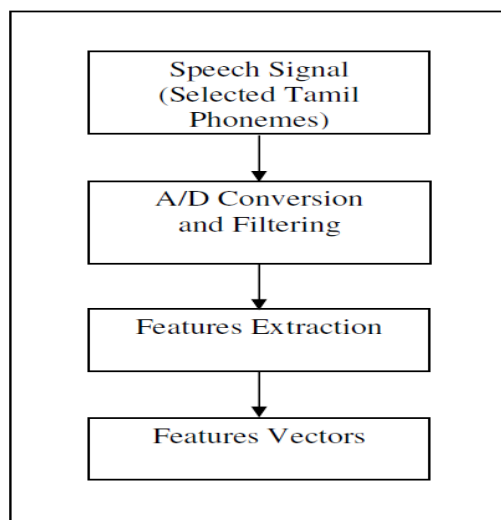
The general scheme of the Visual Feedback System for Tamil Phoneme consists of three parts

1. Preprocessing phase
2. Training phase
3. Testing phase

Preprocessing Phase: Speech signal will be digitized and also the important frequency component in the signal is filtered. Formants are the dominant acoustic components which determine the sound quality of particular vowels. They are formed by the air flowing through the vocal tract, and vibrating at different bands of frequencies as it responds to changes in the tract's shape.

Formants values will be extracted and analyzed for collection of these selected Tamil phonemes. All these Formant values will be collected and then for each phoneme these values are maintained in a database.

Figure 2: Steps in Preprocessing Phase



Training Phase: It is proposed to do the training phase using Radial Basis Function Network. After Feed Forward network, the Radi comprises one of the most used network models. The design of a RBFN in its most basic form consists of three separate layers. The input layer is the set of source nodes (sensory units). The second layer is a hidden layer of high dimension. The output layer gives the response of the network to the activation patterns applied to the input layer (Figure). The transformation from the input space to the hidden-unit space is nonlinear. On the other hand, the transformation from the hidden space to the output space is linear. Training has to be given to boys and girls of Normal Hearing and boys and girls from Hearing Impaired for the selected Tamil phonemes. Here the individual network is developed for each phoneme. These individual networks are developed and train these formant values for each phoneme.

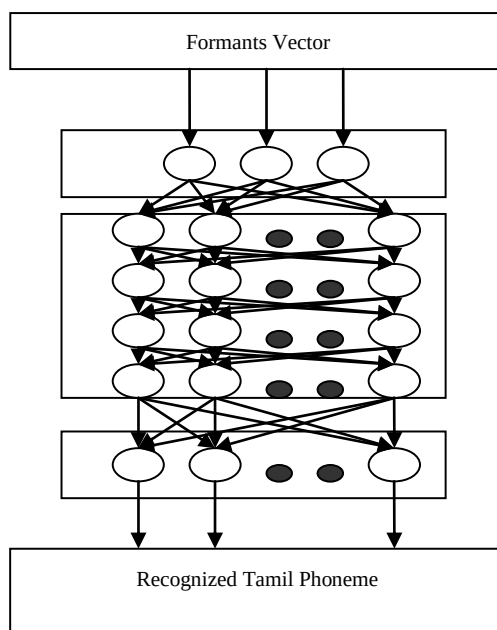


Figure 3: Network Architecture

The aim of this network is to train the net to achieve a balance between the ability to respond correctly to the input patterns that are used for training and the ability to provide good responses to the inputs that are similar. Features such as first two formant values constitute the input vectors to the Radial Basis Function network used as classifier. Multiple layers of neurons with nonlinear transfer functions allow the network to learn nonlinear and linear relationships between input and output vectors

The training phase will be implemented using Radial Basis Function network. For training the network newrb will be used. The neural networks for Normal Hearing have to be designed as 25x20 and Hearing Impaired have been designed as 25x10 neurons corresponding to the first two formant values extracted from 750 samples (25 Tamil phonemes from 30 subjects). There is only one hidden layer in the network. It produces the values in the range -1 to +1. If the target value is 1 then the corresponding phoneme is recognized, otherwise the phoneme is not recognized.

Testing Phase: The trained Networks will be tested Gender wise. Normal Hearing networks will be tested with Normal Hearing and Hearing Impaired input. Hearing Impaired networks will be tested only with Hearing Impaired children. These networks will be simple multiple networks and Visual Feedback is provided by comparing Normal Hearing and Hearing impaired.

The unknown phoneme is given to test for recognition. If the target value is 1 then the corresponding phoneme is recognized, otherwise the phoneme is not recognized. The Speech User Interface gives visual feedback like good, poor or not recognized based on the accuracy level. If the recognition result ranges between 0.4 and 1 then the visual feedback is as good, if the recognition accuracy ranges between 0 and 0.4 then the visual feedback is as poor, otherwise the visual feedback is as Not Recognized.

User Interface: The speech user interface is designed with options like NH, HI, play, save, load, zoom in, zoom out, extract formants and recognize. The other features provided by the user interface are:

- Loading the sound file to play and view waveform
- Selecting Normal Hearing or Hearing Impaired which is boy or girl
- Extracting formant values for the selected sound file
- Have to done Gender wise testing for Normal Hearing network with Normal Hearing and Hearing Impaired and testing Hearing Impaired only with Hearing impaired for the selected Tamil phonemes
- Giving visual feedback as 'Good', 'Poor' or 'Not Recognized' based on the accuracy level of the uttered phoneme.

Expected Outcome: In the Preprocessing phase, the data will be collected and also to get the Formant values for each phoneme and these values are maintained in database. In Training phase, individual network is developed for each phoneme and each network is trained well. In the Testing phase, these networks have to be tested with both Normal Hearing children and Hearing Impaired children and Visual Feedback by Normal Hearing and Hearing Impaired. The speech user interface gives Visual Feedback as recognized or not recognized based on the accuracy level of the phonemes.

V. Conclusion

Speech recognition is a difficult problem, largely because of the many sources of variability associated with the signal. The various sources of variability were phonetic identity, pitch, speaker, microphone and environment. In this Literature Review, it was understood that most of the Automatic Speech Recognition systems were developed for English and other languages like Arabic using HMM. In Indian Languages, Automatic Speech Recognition systems were developed for Tamil, Hindi, Oriya and Tamil. Those models were using HMM, Wavelet, Neural Network or Bootstrapping. It is found that there is no model found for Hearing Impaired in Tamil language using Radial Basis Function Networks. Tamil Language is a historical Dravidian language spoken by around 52 million people in countries like India, Sri Lanka, Malaysia, Vietnam, Singapore, Canada, USA, UK and Australia. And it is a phonetic language, that is, "words are written as they are pronounced, and pronounced as they are written". The proposed speech recognition system is to recognize the Tamil phonemes such as Stops, Laterals and Trills with vowel combination and responds to user with Visual Feedback based on the accuracy level of the user's speech. Using Neural Network Models, number of phonemes and accuracy level can be further improved. The main contribution of our speech recognition system is for early Tamil language learning and articulation training for Normal Hearing and Hearing Impaired children.

References

- [1] Adrin G Bors, "Introduction of the Radial Basis Function Networks", Department of Computer Science, University of York, York, YO10 5DD, UK.
- [2] Ankara, "Radial Basis Function Networks", Chapter IX, METU EEE, EE543 Lecture notes.
- [3] Carlos Gershenson, "Artificial Neural Networks for Beginners", C.Gershenson@sussex.ac.uk4.
- [4] El Choubassi M. M, El Khoury H. E, Jabra Alagha C. E, Skaf J. A and Al- Alaoui M. A, "Arabic Speech Recognition Using Radial Basis Networks", Signal Processing and Information Technology, 2012. ISSPIT 2012. Processing of the 3rd IEEE International Symposium, 14-17 Dec 2012, Page(s): 543-547.
- [5] Joe Tebelskis, "Speech Recognition using Neural Networks", School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213-3890, May 2010, CMU-CS-95-142.
- [6] Kumar M, Rajput N, and Verma A, "A large-vocabulary continuous speech Recognition system for Hindi", Internet publication October 8, 2011
- [7] Lauren Nichole Mungenast, "Speech and the Hearing Impaired Child: Evaluation and Revision of the CID Speech Skills Worksheet", Master of Science in Deaf Education, Washington University School of Medicine, Program in Audiology and Communication Sciences, April 27, 2007.
- [8] Mohanty S, Bhattacharya S, "Recognition of Voice signals for Oriya Language using wavelet Neural Networks", STRANS, IIT Kanpur 2010.
- [9] Oglesby J and Mason J S, "Radial Basis Function Networks for Speaker Recognition", Department of Electrical and Electronics Engineering, University College, Swansea, UK, CH 2977-7/91/0000-0393, 2005
- [10] Professor Don Colton – BYU Hawaii, "Automatic Speech Recognition Tutorial", June 17, 2007
- [11] Rathinavelu A, Anupriya G, Muthanantha murugavel A S, "Speech Recognition Model for Tamil Stops", International conference of Computational Intelligence and Intelligent Systems (ICCIIS '07) during 2-4 July, 2011 at London.
- [12] Rathinavelu A., et al: (2007), "Three dimensional Articulatory Model for Speech acquisition by children with hearing loss", Springer LNCS 4554, 978-3-540-73278-5.
- [13] A, Hemalatha T, Savithri S R and Chitra, "Interactive Multimedia tool to help vocabulary learning of hearing impaired children by Using 3D VR objects as visual cues", Nat J Technol 2006; 2(1), Page(s): 25-32.
- [14] Rathinavelu A., et al: (2006), "Computer aided articulation tutor using three dimensional visual cues for the children with hearing loss", J Comput Sci 2006; 2(1):76-82.
- [15] Rabiner L.R. and B. Juang (1993), "Fundamentals of Speech Recognition", Prentice Hall.
- [16] Saraswathi S and Geetha T V, "Language Models For Tamil Speech Recognition System", IETE Technical Review, Vol 24, No 5, Sep-Oct 2007, PP 375-383.
- [17] Suvidha Rupesh Kumar, B.Bharathi (2006), "Continuous Word Recognizer for Tamil Language", In Proceedings of Sixth National Conference on Convergence of Computer and Information Engineering, 22-23 December 2006, PSG College of Technology, Coimbatore.
- [18] Xinghui Hu, Adrian FOURCIN, Andrew FAULKNER and Jianing Wei, "Speech Reading of Words and Sentences by Normally Hearing and Hearing Impaired Chinese Subjects: The Enhancement effect of Compound Speech Patterns", in Speech, hearing and Language: Work in Progress (1996) Volume 9, 119-131.