PROJECT REPORT

AUTOMATION OF MALAYALAM ARTICULATION TEST USING AUTOMATIC SPEECH RECOGNITION TECHNIQUES

ARF Ref No: SH/CDN/ARFELECT1/2017-18

All India Institute of Speech and Hearing

Manasagangothri, Mysuru - 570006

April 2021

Automation of Malayalam Articulation test using Automatic Speech Recognition Techniques

Project under AIISH Research Fund (ARF)

Year: 2017-18

Ref: SH/CDN/ARFELECT1/2017-18

Total grants: 6.32 lakhs

Dr. Ajish K Abraham (PI)

Professor of Electronics & Acoustics

All India Institute of Speech and Hearing

Mysore– 570 006, Karnataka, India

E-mail: ajish68@aiishmysore.in

&

Dr. Leena Mary (PI)

Professor,

Department of Electronics & Communication Engineering,

Rajiv Gandhi Institute of Technology

Kottayam, Kerala, India.

E-mail: leena.mary@rit.ac.in

Acknowledgments

The investigators would like to thank All India Institute of Speech and Hearing (AIISH), Mysore for providing financial support to carry out this work. We thank all participants for their cooperation at various stages of this project. Our special thanks to Shri. P. Sameer, Professor of Audiology and Speech Language Pathology, Government Medical College, Kozhikode for providing all support during data collection. We also acknowledge the support provided by the Department of Clinical Services and Department of Electronics, AIISH, Mysore. We also would like to thank all members of Center for Advanced Signal Processing (CASP), Department of Electronics and Communication Engineering, Rajiv Gandhi Institute of Technology, Kottayam for the seamless support and encouragement without which the work would not have been completed.

Contents

Abstract

Early intervention has brought about significant impacts on the outcomes of speech therapy in Children with Hearing Impairment (CHI). Conventional method to assess articulation disorders in CHI may be demanding for the speech-language pathologists as the approach is time-consuming. The present study has developed a Graphical User Interface (GUI) based automatic assessment system for identifying articulation error at the phoneme level using Automatic Speech Recognition (ASR) techniques in the speech of Malayalam speaking CHI. The system uses an ASR module to identify the no error condition as well as Substitution, Omission, Distortion, and Addition (SODA) errors at the phoneme level. The developed system will also compute the severity through the Percentage of Consonants Correctness (PCC) score. Besides, the similarity measure between normal and disordered phones has been measured through dynamic time warping, followed by the computation of objective measures to identify the articulation errors. The effectiveness of the system was evaluated on the dataset of 30 Malayalam speaking CHI. The scores, estimated by the system, have been compared with perceptual evaluation scores to confirm the effectiveness of the system. Results show that the system output shows a significant correlation with the perceptual evaluation. The system requires improvement in performance in evaluation of children with severe impairment. GUI of the developed system provides facilities for recording, analysis, display of results and generation of reports, which makes it readily deployable for clinical trials.

CHAPTER I

Introduction

Articulation disorders are commonly found in children with hearing impairment (CHI). Errors in speech production of CHI can be attributed to the degradation in the signal they hear. The extent of impact can vary according to the type and degree of hearing loss. The frequent articulatory errors exhibited by CHI are substitutions, omissions, distortions and additions of phoneme. A comprehensive assessment is required for accurate diagnosis, which will further help to device the protocol for speech therapy. The conventional clinical techniques used for the diagnosis of articulation disorders include administration of articulation test executed manually referring to a printed document/ picture based form. Perceptual evaluation involves time and leads to errors in assessment.

1.1 Conventional method of assessment of articulation disorder in children

Children are expected to acquire speech and language skills within stipulated ages, the ones who fail might have a speech disorder. Articulation disorder involves difficulties in producing specific phonemes accurately, due to the substitution, slurring, or indistinct speech. Children who cannot hear soft speech may have significant trouble in understanding and speaking. Mostly, production is the best indicator of perception as the children produce phonemes as how they hear them. Hence, it is necessary to assess the perception when production is disordered, unintelligible, or absent, whereas articulation disorders may arise from poor production or poor perception. A Speech Language Pathologist (SLP) will assess the children with articulation disorders and find out if there are any concerns about the speech quality, whether the error pattern is specific to one phoneme, or whether the substitution that they make is the same in all word positions. This perceptual assessment is done with the help of a standardized articulation test which is language specific.

In general, the articulation test is made up of two tasks: naming and imitation. In the naming task, the child is asked to name the pictures displayed. If the participant is unable to do so, the examiner names the picture and asks him to repeat the same. Meanwhile, the child is requested to utter the words spoken by the evaluator in the imitation task. The test contains 54 words in total, which include all Malayalam phonemes in different positions. The evaluation of articulation disorders is usually performed through perceptual evaluation by following a protocol for annotating the child's response. The speech samples are recorded and stored to ensure that they can be analyzed as many times as necessary.

The Percentage of Consonants Correct (PCC) score is one of the components of the diagnostic classification system of articulation disorder for quantifying the severity. SLPs use the PCC score extensively in the diagnosis and management of the patient. PCC score can be calculated as follows:

PCC = (Number of consonants correct / Total number of intended consonants)*100

Based on the PCC score, the severity rating can be categorized as mild, mild-moderate, moderate-severe, and severe. While PCC of 85–100% is considered as mild, 65–85% signifies a mild-moderate severity. PCC score ranging from 50–65% is characterized as a moderate-severe case whereas, less than 50% exhibit a severe degree of articulation disorder.

1.2 Automated assessment of articulation errors

A few attempts were made in the past to automate the assessment of articulation errors. For example, Computerized Assessment of Phonological Processes in Malayalam (CAPP-M) is a semi-automated system developed at AIISH, Mysore (Sreedevi et al., 2013). In CAPP-M, pictures depicting words is displayed on the screen. The child will respond by naming the picture and the response is manually entered by the SLP. The system will compare the entered response with the correct one and the errors are identified. This is time consuming and requires human intervention.

Another one is Vagmi Therapy Picture-Word-Articulation Module (Voice and Speech Systems, n.d.) which also aims to provide computerized speech therapy for children with articulation disorders, but requires manual intervention. Instead, if a system can automatically recognize the speech uttered by the client and decide which phoneme is having the articulation error, the human effort and the associated judgmental errors can be reduced.

The development of automated assessment systems based on automatic speech recognition may help to improve the accuracy of the assessment and saves the time of SLP. Automated systems can also act as support systems for primary health care providers who are frontline workers in identifying persons with communication disorders. In addition, these tools can also be used by caregivers to evaluate the effectiveness of the speech therapy protocol. However, there is a lack of automated tools for evaluating articulation disorders. Hence, there is a need to develop automatic methods for identification of articulation errors.

Various signal processing techniques proposed in the literature suggests that automated evaluation using various signal features such as Teager Energy Operator (TEO), Linear Predictive Coding (LPC), Mel Frequency Cepstral Coefficients (MFCC), Pitch, Jitter, Shimmer and the first three formants together with the bandwidth of the first formant etc. is effective. The system requires training and testing with normal and disordered speech samples. The perceptual evaluation techniques require skilled and trained personnel, also several man hours and efforts. Although it is efficient and accurate, the effort shouldered by these people may be reduced by using objective evaluation techniques that make use of certain objective measures such as Itakura Saito (IS) measure, Log Likelihood Ratio (LLR), Log Area Ratio (LAR), Segmental SNR measure, and Weighted Spectral Slope (WSS).The purpose of the work is to develop a system for automated articulation test in Malayalam using automatic speech recognition system for assessing the articulation errors.

1.3 Definition of the problem

Speech development in CHI is hampered by inadequate auditory input. Speech of CHI has reduced intelligibility compared to typically developing children (TDC), mainly because of articulatory errors. SLPs assess these errors through perceptual evaluation and accordingly device the protocol to correct them through several sessions of speech therapy. Automatic methods need to be developed to reduce the time and enhance the accuracy of assessment. The problem is to identify the articulation error in the speech of the children with articulation disorder using automatic speech recognition methods. Speech recognition converts speech into the corresponding text irrespective of the speaker characteristics. It requires training the system with hours of speech recordings and corresponding transcribed text data in Malayalam. A standard set of phonemes is used for transcribing the speech data. In this work, the possibility of using an Automatic Speech Recognition (ASR) system in Malayalam for the automatic identification of articulation errors is explored.

Early assessment and rehabilitation of individuals with articulation impairment are essential to prevent long-standing communication disorders. SLPs devote inordinate amounts of time and effort during their assessment sessions, which may hamper the quality and schedule of assessment. Apart from this, the frequency and duration of therapy sessions tend to be a financial burden to families having a child with articulation disorders. These problems demand a more systematic solution to meet the needs of patients and to improve the quality of speech therapy. With these in view, this work is intended to build an automatic articulation assessment system to identify the articulation errors in the speech of children with articulation disorder, using automatic speech processing techniques.

1.4 Objectives

The study was conducted with the following objectives:

a. To develop a software-based fully automated system for articulation test in Malayalam for children with hearing impairment by evaluating the response through digital speech processing and automatic speech recognition methods.

b. To conduct clinical trials of the developed software and validate the test results by comparing with the conventional articulation test administered by experienced speech language pathologists.

c. To deploy prototype of the developed automated articulation test system in Malayalam.

CHAPTER II

Review of Literature

2.1 Automated assessment of articulation errors

Several efforts have been made in the past to automate the evaluation of articulation errors. These attempts may be categorized under three headings: 1. Semi-automatic systems, 2. Studies to identify features for automatic assessment of articulation errors and 3. Studies regarding automatic assessment of articulation errors. A review of these attempts is presented in the following subsections.

2.1.1 Semi-automatic systems

These systems make use of technology as an aid to automate only the assessment procedure. For example, Computerized Assessment of Phonological Processes in Malayalam (CAPP-M) is a system developed at AIISH, Mysore (Sreedevi et al., 2013). Using CAPP-M, SLPs can assess phonological process in Malayalam speaking children in the age group of 2.0 to 3.6 years. Each target word will be displayed on the screen in the form of the text for a picture along with its correct text and different erroneous productions. The child has to name the target word. The SLP who is administering the test has to listen to the child's response carefully and then manually mark the response on the screen. At the end of the test, the report of the phonological processes will be summarized and displayed by the system.

Vagmi Therapy Picture-Word-Articulation Module is another system which also aims to provide computerized assessment for monitoring the progress of speech therapy (Voice and Speech Systems, n.d.). This module is currently available in Malayalam, English, Kannada, Telugu, Hindi, Oriya, and Arabic and is again semi-automatic. The system will show the picture of the target word along with a back play of the spoken word. The child will listen to the word and then repeat. The SLP has to enter the word in the system, as uttered by the child. After entering the response of the child, the system will show the score (on a scale of 0 to 5) and also will give a feedback on the required improvement. Several software tools were found to be very supportive to the pathologists in training their patients (Hutchins et al., 1992, Maier et al., 2006.

2.1.2 Studies to identify features for automatic assessment of articulation errors

Story and Bunton (2017) have developed a method for assessing the characteristics of vowel production by measuring the local density of normalized formant frequencies (F1 and F2). The resulting three-dimensional plot called the vowel space density (VSD) indicated the regions in the vowel space most heavily used by a talker during speech production. They concluded that VSD plots provide a more comprehensive view of vowel characteristics during speech production and hence may be explored to identify articulation errors.

Van Son et al. (2018) experimented whether vowel space can be used as a tool to evaluate articulation problems. They investigated whether vowel space parameters reliably relate to the individual changes in persons with articulation disorders and also whether the vowel space parameters are related to clinically relevant aspects of speech. They found that vowel space parameters contain relevant information about vowel articulation in normal speakers as well as in persons with articulation disorder.

Automatic speech recognition systems using various features such as Teager Energy Operator (TEO), Linear Predictive Coding (LPC), Mel Frequency Cepstral Coefficients (MFCC), Pitch, Jitter, Shimmer and the first three formants together with the bandwidth of the first formant etc. (Nieto et al., 2015) are effective. The system requires training and testing of various speech recordings including both normal and disordered speech to produce a result.

The effort shouldered during perceptual evaluation may be reduced by using objective evaluation techniques that make use of certain objective measures such as Itakura Saito (IS) measure, Log Likelihood Ratio (LLR), Log Area Ratio (LAR), Segmental SNR measure, and Weighted Spectral Slope (WSS) (Harris et al., 2005).

2.1.3 Studies regarding automatic assessment of articulation errors

Novotný et al. (2014) developed a system for fully automatic evaluation of acoustic features related to articulation attributes in Parkinson's disease (PD), based on DDK utterances. Results showed that the system was able to provide reliable automatic assessment, allowing the examination of a wide range of articulatory deficits connected with hypokinetic dysarthria. Moreover, the combination of the presented acoustic features accurately predicted speech impairment. However, this system considered features describing voice quality, coordination of laryngeal and supralaryngeal activity, precision of consonant articulation, tongue movement, occlusive weakening, and speech timing. It did not identify the articulation errors, as the focus was on estimating the severity of impairment.

Bhat et al. (2015) developed an automatic assessment method for articulation errors at phone level in Hindi and classifying a patient utterance as either correct, substitution, omission, distortion or addition (CSODA). Identification of the error at phone level is essential to provide the patient with actionable feedback for correction. The objective of the work was to improve the recognition ability of the ASR to identify articulation errors through improved classification of consonants. An automatic speech recognition (ASR) based method was proposed to identify substitution errors for consonants, using a rule based language model (LM) as well as tuning of acoustic models (AM) for consonants under consideration. But the researchers could validate their systems on normal speech only.

Ng et al. (2018) developed a system for automated assessment of speech sound disorders for Cantonese-speaking pre-school children. The system comprised a mobile application software, a back-end automatic speech recognition (ASR) system for child speech, and a clinically informed assessment scheme. The articulation errors were detected by performing ASR decoding with a language model (LM) that is tailored for each test word.

In the study conducted by Maier et al., 2006, objective evaluation of the intelligibility of speech was performed and then compared with expert's evaluation. Through this work, it is illustrated that automatic speech recognition system trained for German could be used to objectify and quantify global speech outcome of children with cleft lip and palate (CLP). Also there are various systems and proposals which suggest that evaluation of speech disorders by signal processing methods yields same result as an expert's evaluation.

2.2 Challenges in automatic assessment of articulation errors

The challenges involved in vowel detection were observed by several researchers. [e.g., Story and Bunton (2017), Clapham (2016)]. They stressed the need to make vowel detection more robust. Difference between male and female voices was another challenge pointed out by Gold (1978) and Lousada et al. (2013). They suggested that the normalization between male and female voices might be improved.

Shahin et al. (2020) have analyzed the challenges involved in automatic detection of speech disorders in children. The limited availability of disordered speech corpora pertaining to children with speech disorders, in the public domain, is a major constraint. Any researcher who attempts to take up this problem has to collect the data himself, which is expensive and time consuming. The large variability in acoustic characteristics across different ages, dialects, types and severity of speech disorders, will add to the challenges. The second challenge highlighted by Shahin et al. (2020) is the issues with speaker diarization. Most of the available speech corpora collected from children with speech disorders are recordings of complete therapy sessions. Speaker diarization with minimal diarization error rates is essential before automatic analysis, which will be a challenging problem. The third issue is with the acoustic modeling of child speech. Linguistic and acoustic mismatches between adult and child speech is a matter of concern. Hence, the automatic speech recognition systems built for adult speech perform poorly with child speech.

2.3 Summary

Previous researchers have highlighted the need for tools for evaluating articulation and pronunciation beyond perceptual assessments (for e.g., Van Son et al., 2018). In the course of therapy, there is a need to quantify and document the quality of speech so that both patients and SLPs can evaluate the progress. For voice, there are tools that can give an automatic and objective assessment (Barsties & De Bodt, 2015). However, there is a lack of tools for evaluating articulation and pronunciation beyond perceptual assessments. Till now in Malayalam language, signal processing techniques have not been implemented for the evaluation of articulation disorders. The present study is an attempt to develop an automatic speech recognition based system for identifying the articulation errors in Malayalam speaking children with hearing impairment.

CHAPTER III

Method

The aim of the present study was to develop a software-based fully automated system for articulation test in Malayalam for children with hearing impairment by evaluating the response through digital speech processing and automatic speech recognition methods. Validation of the developed system was done by comparing with the conventional articulation test administered by experienced speech language pathologists.

3.1 Participants

Group I

Thirty children (fifteen males and fifteen females) with moderate to severe degree of hearing impairment between 3 and 6 years of age satisfying the following inclusion criteria participated in the study:-

a. Children with bilateral moderate to severe sensori-neural hearing loss.

b. Children who were aided by at least two years.

c. No presence of any other co-morbid disorders.

d. Native speaker of Malayalam language and exposed to the same.

Group II

Thirty normal healthy children (Fifteen males and fifteen females) of the same age group who are native Malayalam speakers also participated in the study. It was ascertained from a structured interview that the participants doesn’t have any neurological, psychological, speech, language or hearing disorders.

A written consent was taken from the participants before recording. Ethical clearance was also obtained from the AIISH Ethics Committee.

3.2 Material

The keywords identified for the evaluation of articulation disorder of nasals, stops, fricatives, approximants, vowels, clusters and affricates are shown in Table 1.

Table 1:

Keywords identified for the evaluation of articulation disorder

a. In IPA format

Nasals	Stops	Fricatives	Approximants	Vowels	Clusters	Affricates
ma:ŋa	kaḍa	su:rjan	Eli	aɳɳa:n	pa:t̪ram	i:ʧʧa
ɲaɳḍʌ	t̪a:kko:l	kase:ra	je:ʃU	a:na	vast̪ram	ra:ja:vʌ
ma:la	t̪at̪t̪a	fa:n	mUjal	e:ɳI	braʂʌ	ʤanal
a:ma	mo:t̪iram	ʃivan	lo:rI	u:ɲa:l	gḷa:ssʌ	ʧ^{h}a:ja
maram	uḍUppʌ	me:ʃa	alama:ra	i:ʧʧa	ʈrain	
a:na	bassʌ	ʂarṭʌ	vIral	uḷḷi	ple:ttʌ	
kiɳar	pu:mba:tta	braʂʌ	ʧeruppʌ	ila	biskattʌ	
fo:ɳ	rat^{h}am	simham	vaḷa	eli	pust̪akam	
	miṭ^{h}a:ji				spu:ɳ	

b. In Malayalam

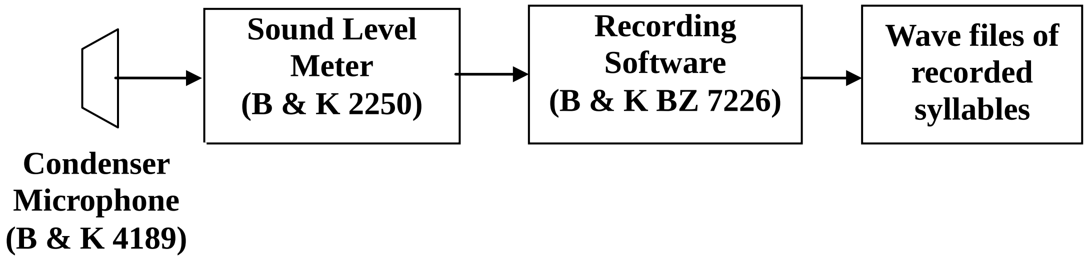
Nasals	Stops	Fricatives	Approximants	Vowels	Clusters	Affricates
മാങ്ങ	കട	സൂര്യൻ	എലി	അണ്ണാൻ	പാത്രം	ഈച്ച
ഞണ്ട്	താക്കോൽ	കസേര	യേശു	ആന	വസ്ത്രം	രാജാവ്
മാല	തത്ത	ഫാൻ	മുയൽ	ഏണി	ബ്രഷ്	ജനൽ
ആമ	മോതിരം	ശിവൻ	ലോറി	ഊഞ്ഞാൽ	ഗ്ലാസ്	ഛായ
മരം	ഉടുപ്പ്	മേശ	അലമാര	ഈച്ച	ട്രെയിൻ	
ആന	ബസ്സ്	ഷർട്ട്	വിരൽ	ഉള്ളി	പ്ലേറ്റ്	
കിണർ	പൂമ്പാറ്റ	ബ്രഷ്	ചെരുപ്പ്	ഇല	ബിസ്കറ്റ്	
ഫോൺ	രഥം	സിംഹം	വള	എലി	പുസ്തകം	
	മിഠായി				സ്പൂൺ	

3.3 Instrumentation

A high quality recorder coupled with a precision condenser microphone (Figure 1) was used in the present study to record speech samples spoken by participants in Group I and Group II.

Figure 1

The recording set up



3.4 Procedure

3.4.1 Recording of normal and disordered speech

The participants were seated comfortably in a sound treated room with the recording microphone placed in front of the participant at a distance of 15 cms. The participants were asked to read the words (Table 1) in a normal conversational voice at a comfortable pitch and loudness. Each word was repeated three times maintaining a constant interval. The words uttered by the participants were recorded with the recording set up shown in Figure 1.

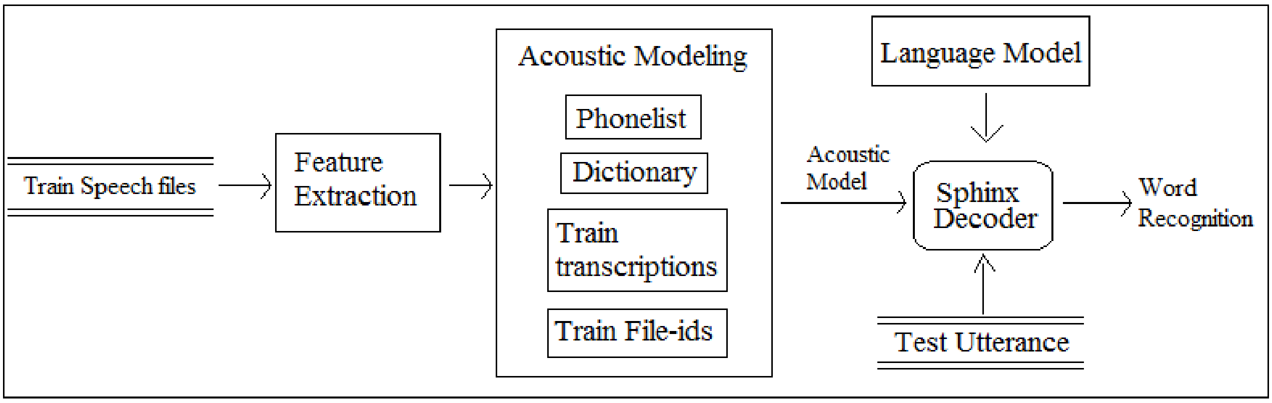
3.4.2 Speech processing

3.4.2.1 Development of ASR using CMU Sphinx

The first step is to develop an Automatic Speech Recognizer for the identified words (given in Table 2) in Malayalam using CMU Sphinx tool kit. CMU Sphinx is a project developed by the researchers at Carnegie Mellon University as part of their research works focusing on practical application development with ASR system especially for low-resource platforms. In order to develop a speech recognition system, some text files regarding the speech input needs to be prepared in toolkit format. As shown in Figure 2, Sphinx files to be prepared includes, file-ids corresponds to train and test speech, transcriptions of speech contents in Malayalam itself, phone-list used for acoustic modeling, dictionary showing the mapping between word and phones etc.

Figure 2

Block diagram showing the development of ASR using CMU Sphinx



CMUSphinx project already contains several acoustic models in foreign languages suchas US English, Chinese or French. But for small vocabulary applications in Malayalam, it is necessary to train an acoustic model using enough speech data. In this work, new acoustic models are trained using normal speech data recorded from Group II. The training database contains recordings of 30 speakers containing variations of linguistics and acoustics. All the speech files are with 16kHz sampling frequency and mono channel recordings in MS WAV format. Now, in order to label the sound unit in the acoustic model, a transcript file is prepared which shows the sequence of words and non-speech sounds in the same order as they occur in speech signal. Once all the training files are prepared, Hidden Markov Modelling (HMM) can be done by extracting 13 dimensional Mel Frequency Cepstral Coefficients (MFCC).

The statistical language model is an essential part of configuring the decoder which tells probabilities of the sequences of words that are possible to recognize. The probabilities are estimated from a large sample of data for single words, two words and three words commonly known as 1-grams, 2-grams and 3-grams respectively.. Once the text data is cleaned, several language modeling tools can be used to develop a language model such as SRILM - The SRI Language Modeling Toolkit , CMU SLM - CMU Statistical Language Modeling Toolkit etc.

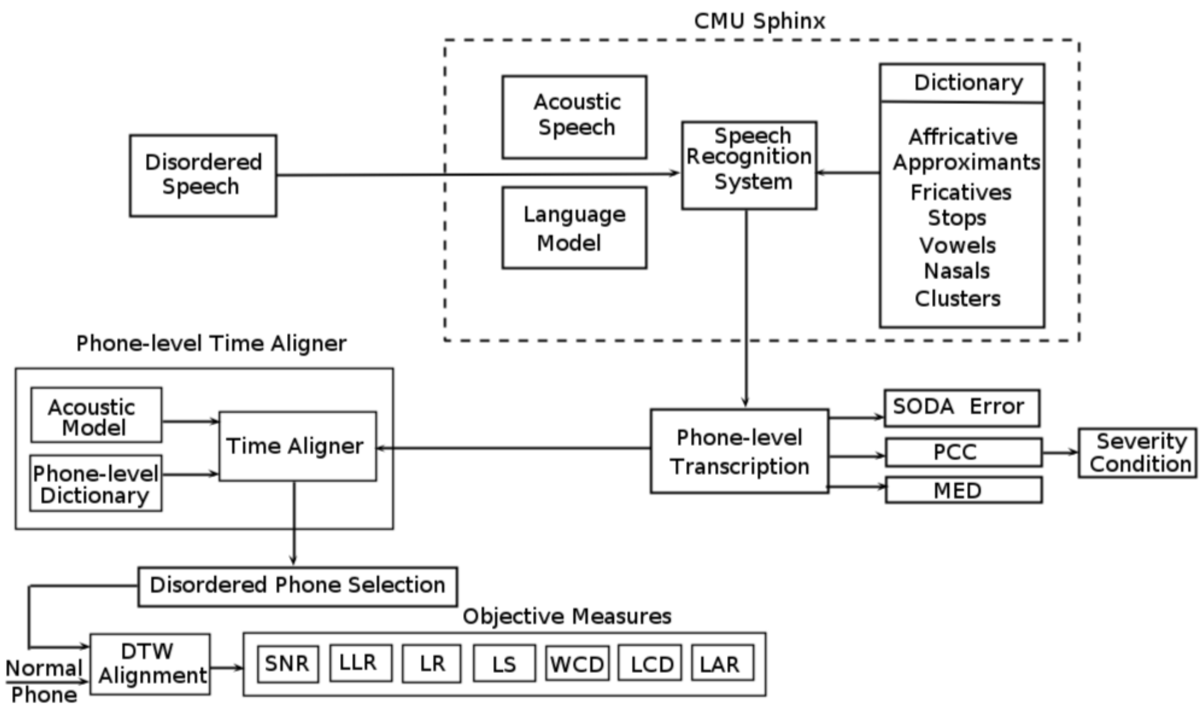
After training an acoustic model, a decoder is initialized to check the training results, which takes the model, test utterances and reference transcriptions to calculate Word Error Rate (WER) of the model. Also it uses the language model which tells the decoder the sequence of words in terms of probabilities of words and word combinations in the language. Decoding is done on the test data to find the performance of trained model and to understand the level of performance of application that uses the model.

3.4.2.2 Development of system for automation of articulation test using the ASR

Figure 3 depicts an overview of the methodology adopted for speech processing and identification of errors. The clusters from words, broad phoneme classes such as affricate, approximants, stops, fricatives, vowels, and nasals, are considered in the proposed work. Initially, an HMM-based ASR system has been built for the selected set of words as per the standardized articulation test in Malayalam. The disordered speech, given as input to the ASR provides a string of broad phoneme-like units as an output. Then, Minimum Edit Distance (MED) has been calculated from the semantic meaningful text output to quantify the similarity of strings. Likewise, articulation errors in terms of Substitution, Omission, Distortion, and Addition (SODA error) have been identified followed by the severity of speech disorder that has been measured using the PCC score. Besides, the target phoneme has been selected and cut from the speech signal with reference to the time aligner result. Furthermore, the similarity is measured by matching the corresponding phoneme features (such as MFCC, LPC) of normal as well as the disordered speech using Dynamic Time Warping (DTW). Objective speech quality measures such as Itakura-Saito (IS), Log-Likelihood Ratio (LLR), Log Area Ratio (LAR), Likelihood Ratio (LR), Weighted Cepstral Distance (WCD), Log Cepstral Distance (LCD) and Signal to Noise Ratio (SNR) have been computed for normal and disordered phoneme. Objective scores are calculated as per the protocol followed in the articulation test. Then the scores have been compared with subjective evaluation scores to confirm the effectiveness of objective measures.

Figure 3

Block diagram showing the methodology adopted for speech processing



3.4.3 Perceptual Evaluation of Articulation Errors

The perceptual evaluation of the recorded speech samples were done by five experienced speech language pathologists. The errors identified by them and the PCC scores obtained formed the reference for analyzing the performance of the automated system.

3.5 Analyses

The correlation between the articulation errors identified for each word by perceptual analysis and automatic method was analyzed for substitution error, omission error, distortion error, addition error, no error condition and PCC.

CHAPTER IV

Results

The objective of the current study was to develop an Automatic Speech Recognition (ASR) system in Malayalam for the automatic assessment of articulation disorder. The results of the study are provided under the following subsections.

4.1 Characteristics of participants

In Group I, thirty children (Fifteen males and fifteen females) with moderate to severe degree of hearing impairment between 3 and 6 years of age and who are native Malayalam participated in the study. Mean age of the participants was 4.11 (SD = 1.01) years. In Group II, thirty normal healthy children (Fifteen males and fifteen females) of the same age group who are native Malayalam speakers participated. Mean age of the participants was 4.13 (SD = 0.7) years.

4.2 Development of Graphical User Interface (GUI)

A Matlab based GUI has been developed with four tabs 'demography', 'record', 'recognizer', and 'result', as shown in Figure 3, Figure 4, Figure 5 and Figure 6, respectively. Using the 'demography' tab (Figure 4), data pertaining to the child such as the case number, date, demographic information etc. ought to be entered. The 'record' tab (Figure 5) consists of a drop-down menu that displays a list of broad phoneme classes, which allows the user to make a selection. When the user chooses the phoneme class the image and text of the word to be uttered, will be displayed. A 'play' button is provided for playing back the recorded sound of the word. The user can record (at a sampling frequency of 16000 Hz), the word spoken by the child, playback the recording, plot the waveform and store the recorded audio signal by pressing the record button. The 'recognizer' tab (Figure 6) directs us to the speech recognition module, which processes the audio input. The results are displayed graphically using the 'result' tab (Figure 7). The SODA error, MED score and objective measures are calculated with the 'recognizer' tab, whereas the 'result' tab displays the PCC score and severity condition. The 'result' tab is provided with a 'report' button. A report can be generated based on the system assessment findings, by clicking the 'report' button.

Figure 4

Screenshot of demography tab in GUI

Figure 5

Screenshot of record tab in GUI

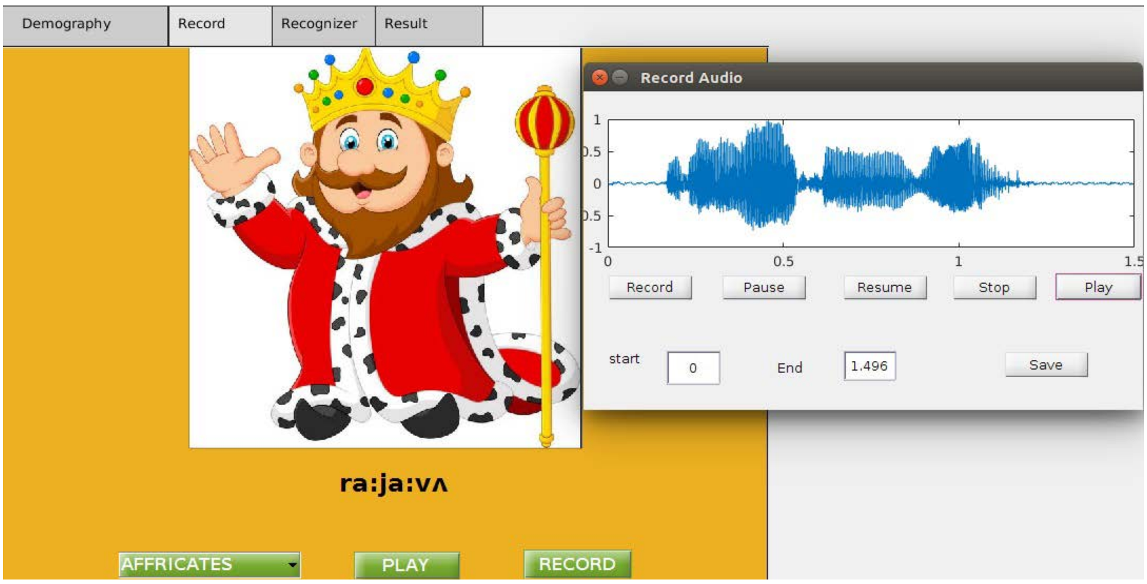


Figure 6

Screenshot of recognizer tab in Matlab GUI

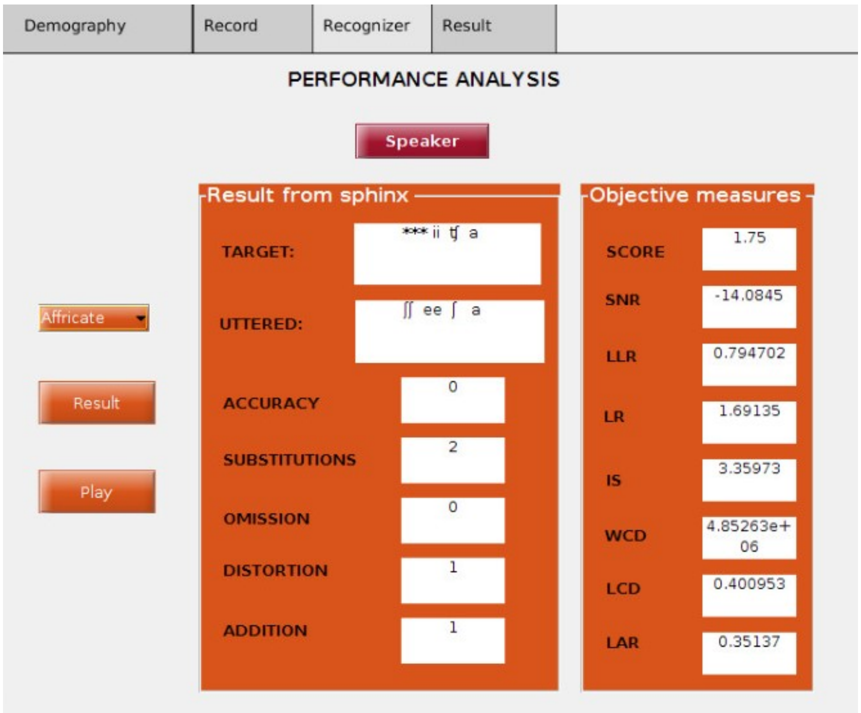
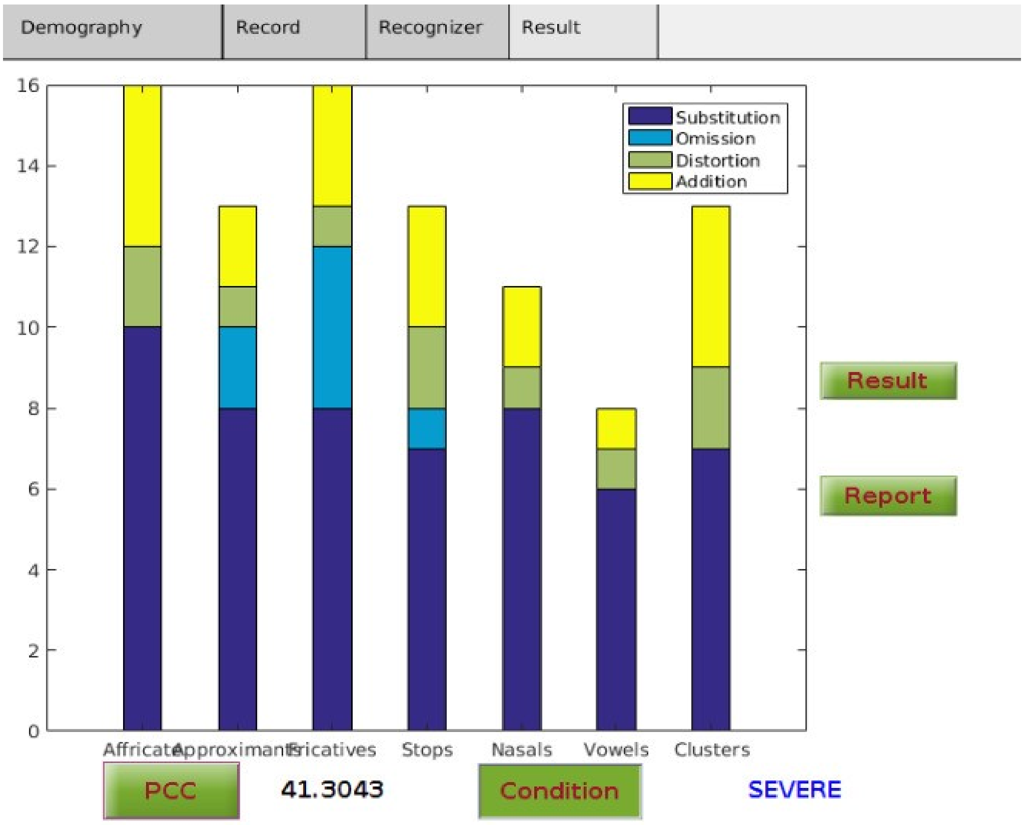


Figure 7

Screenshot of result tab in GUI



4.3 Framework:

The system works in Ubuntu operating system and to run the software, it is mandatory to install the following in the system

1. Matlab R2016b
2. CMU-Cam_Toolkit_v2
3. SRI language modelling toolkit
4. Sphinxtrain-master
5. Pocketsphinx-master
6. Sphinxbase-master
7. Apache NetBeans IDE 11.0
8. Sphinx4

4.4 Operating procedure of the developed system

(a) Run the matlab code “SimpleOptimizedTabs3.m”

(b) A GUI opens consisting of 4 tabs - Demography, Record, Recognizer and Result

(c) In the ‘demography’ tab, enter the details of the child to be tested. It is mandatory to enter all the details, otherwise the code will show an error message.

(d) In ‘record’ tab, select the category of the target sound first, and then click the play button to hear the word. Then, click on the record button, a GUI will open. In the new GUI, click the record button to start recording and click stop button to stop it. Play button is used for playing back the recorded sound. There is an option to select the desired portion of the recorded speech by entering start and stop values. The edited recorded speech can then be saved in the name of the participant in the record_wav folder.

(e) In ‘recognizer’ tab, select the speaker from record_wav folder, choose the category, click the result and play button to display the results and play the audio file. Click the play button only after clicking the result button to hear the word which is displayed in the result. Click the result button one by one to view the results of each word.

(f) In result tab, click the result button to view the result sheet and PCC button to view the PCC score and severity. Click the 'Report' button to generate the report. Generated report will be in record_wav folder.

4.5 Implementation of Speech Recognition module using CMU Sphinx

It can often be challenging to make an accurate conversion of speech to a stream of phones as a consequence of contextual constraints. Researchers rely on various tools such as HTK, Kaldi, and Sphinx for developing an ASR system. In this work, the CMU sphinx-4 has been used to build a phone model Malayalam speech recognizer as it is flexible for any language. A text file has been created to represent the words with their corresponding phonetic transcription, which helps to verify the correctness of the pronunciation of the word. Apart from the speech signals and transcription file, the trainer requires to access a language phonetic dictionary as well as a filler dictionary. Non-speech sounds in the filler dictionary are mapped to non-speech or speech-like sound units, whereas, legitimate words are mapped to corresponding sequences of sound units in the phonetic dictionary. A set of 53 Malayalam phonemes generated for the phonetic dictionary is used to build the language model. The model is trained with SRI Language Modeling Toolkit (SRILM), which is the easiest and advanced toolkit.

Training of ASR was carried out using speech recorded from 22 speakers(11Male and 11Female). It was then tested using speech recorded from 6 other normal speakers (3 Male and 3 Female). Testing was carried out separately of each category of sounds such as affricates, approximants, vowels, fricatives, nasals, stops and clusters. For some of the categories, training using additional training examples were carried out to obtain the required accuracy. The following measures used for the evaluation of ASR:

i. Word error rate (WER) = (I + D + S) / N where N be the length of the text. I is the number of inserted words, D is the number of deleted words and S represents the number of substituted words. The WER is usually measured in percentage.

ii. Accuracy = (N - D - S) / N

It is almost the same as the word error rate, but it doesn't take insertions into account.

Some of the results are appended below in Table 2.To obtain a better model, human tuning was performed on the training data. From Table 2, it has been clear that the system provides 100% accuracy when tested using normal speech for all cases except nasals, vowels, and clusters.

Table 2: Result obtained while testing with the normal samples for each category

Category	Total percent correct (%)	Error (%)	Accuracy (%)
Affricates	100	0	100
Approximants	100	0	100
Fricatives	100	0	100
Stops	100	0	100
Nasals	97	4	96
Vowels	93.40	7.55	92.45
Clusters	98	3	97

4.6 Minimum Edit Distance

Minimum edit distance (MED) is a widely used string similarity measure which transforms one string into another through a sequence of operations such as insertion, deletion, or substitution. Table 3 specifies the cost assigned to different operations for computing the edit distance. When target and uttered phones are identical, the edit distance assigns zero scores; conversely, the score increases according to the values in the scoring matrix as the similarity reduces.

Table 4 displays the computation of minimum edit distance between the target text and recognizer output for the word i:ʧa. Initially, all the characters are aligned equally, resulting in the edit distance of 0. When the second case is taken into account, the MED score is 1, as the phone i: is deleted on the recognizer text. While aligning the strings i:ʧa and ʃe:ʃa in the third example, making an insert operation of character ʃ, substitution operation of ʃ and distortion operation of e: on the latter transforms it into the former. Therefore, the edit distance between the strings turns out to be 1.75.

Table 3

Cost assigned to different operations for MED calculation

Operation	**Cost**
Correct Response	0
Substitution	0.5
Omission/Addition	1
Distortion	0.25

Table 4

Illustrations of minimum edit distance calculation for the word 'Eecha'

Target text	i: ʧ a	i: ʧ a	i: ʧ a
Recognizer output	i: ʧ a	* ʧ a	ʃ e: ʃ a
Number of matches	3 (i:i:=0,ʧʧ=0,aa=0)	2	1
	-	(ʧʧ=0,aa=0)	(aa=0)
Number of mismatch	-	-	2(i:e:=0.25, ʃ=0.5)
Number of gap	0+0+0	1(i:=1)	1 (ʃ=1)
MED score calculation	0	1+0+0	0+0.25+0.5+1
MED score		1	1.75

4.7 Analyses

Analyses were carried out to determine the correlation between the automated and conventional articulation tests in rating the severity of articulation disorders. Table 5, 6, 7, 8, 9, 10 and 11 illustrate these comparisons for affricates, approximants, fricatives, nasals, stops, vowels and clusters, respectively. CHI who was included in the study had moderate to severe articulation disorder. While the target indicates the target phoneme to be considered, ASR out and perceptual indicates the phoneme recognized by the system and the phoneme recognized by the SLP during perceptual evaluation, respectively. MED indicates the similarity measure of strings between the target and system output. The results obtained from the ASR system shows an improvement in the recognition accuracy when compared with the perceptual evaluation.

Table 5

Comparison of target phoneme recognition of affricates by the developed system and by the participant no.10.

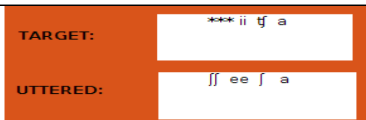
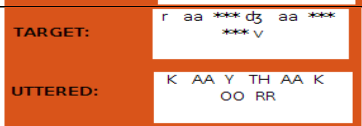
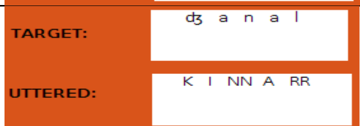
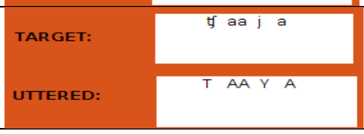
Word	**Target**	**ASR out**	**Perceptual**	**MED**	**Screenshot of result obtained**
i:ʧa	ʧ	ʃ	ʃ	1.75	
ra: ʤ:vʌ	ʤ	th	tʃ	4	
ʤanal	ʤ	K	OM	0.75	
ʧha:ja	ʧ	t	T	0.25	

Table 6

Comparison of target phoneme recognition of approximants by the developed system and by the participant no.10.

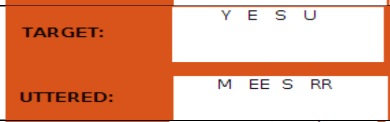
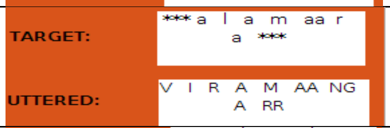
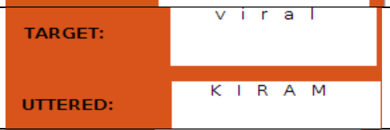
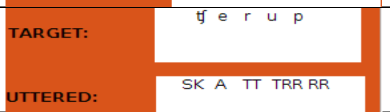
Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
Eli	l	l	d	0	
je:ʃU	j	m	j	1	
mUjal	j	OM	OM	3	
lo:rI	l	r	j	3	
alama:ra	l	r	ɳ	3	
vIral	l	m	n	1	
ʧeruppʌ	r	tt	d	2.25	
vaḷa	ḷ	ḷ	ḷ	1	

Table 7

Comparison of target phoneme recognition of fricatives by the developed system and by the participant no.10.

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
su:rjan	s	sp	s	3	TARGET: ʃʃ uu r j a n UTTERED: SP UU *** *** CH EE
kase:ra	s	tt	tʃ	3.75	TARGET: k a ʃʃ ee r a *** UTTERED: *** *** TH I R A RR
fa:n	f	f	p	0	TARGET: f aa n UTTERED: F AA N
ʃivan	ʃ	ʃ	ʃ	3.5	TARGET: *** ʃ i *** v a n UTTERED: EE S I K U D A
me:ʃa	ʃ	ʃ	ʃ	0	TARGET: m ee ʃ a UTTERED: M EE S A
braşʌ	ş	v	-	3.25	TARGET: *** *** br a ʃh UTTERED: PL EE S I V
Simham	h	v	m	2	TARGET: ʃʃ i m h a m UTTERED: S I *** V AA RR

Table 8

Comparison of target phoneme recognition of nasals by the developed system and by the participant no.10.

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
ma:ŋa	ŋ	tt	ŋ	1.25	TARGET: m aa ng a UTTERED: PL EE TT A
ɲaɳd̪ʌ	ɲ	U:	t	3.5	TARGET: *** *** ndʒ a nd̪ UTTERED: CH E UU CH A
ma:la	m	m	m	0.25	TARGET: m aa l a UTTERED: M AA R A
a:ma	m	m	m	0.5	TARGET: aa m a UTTERED: AA M M
Maram	m	OM	m	1.25	TARGET: m a r a m UTTERED: M AA Y A ***
kiɳar	ɳ	ɳ	ɳ	0	TARGET: k i ɳ a rr UTTERED: K I NN A RR
fo:ɳ	ɳ	ɳ	OM	1.5	TARGET: f o: ɳ *** UTTERED: SP UU NN I
a:na	n	l	ɳ	0.75	TARGET: aa n a UTTERED: OO L A

Table 9

Comparison of target phoneme recognition of stops by the developed system and by the participant no.10.

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
ta :ko:l	k	u	t	2.75	TARGET: ṭh aa *** k o: l UTTERED: P AA M U LL I
tatta	tt	tt	t	1	TARGET: ṭh a ṭh a *** UTTERED: TH A TH A RR
mo:titra m	tt	tt	t	1	TARGET: m o: ṭh *** i r a m UTTERED: M OO TH K I R A M
uḍUpʌ	p	t	p	1.75	TARGET: u ḍ u p UTTERED: *** PL EE TT
bassʌ	b	br	p	0	TARGET: b a ʃʃ UTTERED: BR A SH
pu:mba: tta	tt	thr	tt	3.75	TARGET: p uu m b aa ṭṭ a *** UTTERED: SK UU *** P AA THR A M
Ratham	th	OM	t	2.5	TARGET: r a ḍh a m UTTERED: T A *** RR T
miṭha:ji	th	th	th	1	TARGET: m i ṭ aa j i UTTERED: U D T AA Y I
kaḍa	k	f	ɳ	2.25	TARGET: k a ḍ a *** UTTERED: F OO NN A RR

Table 10

Comparison of target phoneme recognition of vowels by the developed system and by the participant no.10.

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
i:ʧa	i:	e:	i:	1.75	TARGET: *** ii ʧ a UTTERED: SS EE S A
a:na	a:	o:	a:	0.75	TARGET: aa n a UTTERED: OO L A
Ila	I	I:	I	0.5	TARGET: i l a UTTERED: II CH A
Eli	E	E	E	0	TARGET: e l i UTTERED: E L I
aɳɳa:n	a	pl	a	2	TARGET: a ɳ aa n UTTERED: PL EE S A
e:ɳI	e:	e	i for e	0	TARGET: ee ɳ i UTTERED: E NN I
u:ɲa:l	u:	u:	u:	1.25	TARGET: *** uu nʤ aa l UTTERED: SP UU NN A RR
ul̪l̪i	u	u	u	0	TARGET: u ll i UTTERED: U LL I
i:ʧa	i:	e:	i:	1.75	TARGET: *** ii ʧ a UTTERED: SS EE S A

Table 11

Comparison of target phoneme recognition of cluster by the developed system and by the participant no.10.

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
braʂʌ	br	CR	CR	3.25	TARGET: *** *** br a ʃh UTTERED: PL EE S I V
pa:trt am	trt	trt	CR	0	TARGET: p aa t̪hr a m UTTERED: P AA THR A M
vastrt am	strt	CR	CR	1	TARGET: v a ʃt̪hr a m UTTERED: BR A SH A RR
gl̩a:ssʌ	gl̩	gl̩	gl̩	2.5	TARGET: gl aa *** *** ʃʃ UTTERED: GL AA SH A RR
t̪rain	t̪r	CR	CR	3.25	TARGET: t̪rr e j i n UTTERED: *** *** K OO RR
ple:ttʌ	pl	CR	CR	1.25	TARGET: pl ee t̪t̪ UTTERED: E M K
biskattʌ	sk	CR	CR	1.75	TARGET: b i ʃk a *** t̪t̪ UTTERED: M EE S A RR T
pust̪ka m	stt	CR	CR	2.25	TARGET: p u ʃt̪ a k a m UTTERED: *** M CH AA TT A M
spu:ɳ	sp	sp	ʃp	0.25	TARGET: ʃp uu ɳ UTTERED: SP UU M

4.8 Performance evaluation of the system based on correlation

Table 12 shows the procedure followed for finding the correlation between perceptual method and ASR method, considering target words i:ʧa, ra: ʤ:vʌ, ʤanal and ʧha:ja as examples. The correlation between the total number of articulation errors identified for each target word by perceptual analysis and ASR method is estimated for substitution error, omission error, distortion error, addition error and no error conditions. The correlation analysis is conducted on the data of 30 children with hearing impairment. Data indicated in the ‘no error’ row shows the number of target words articulated without any errors. Data indicated in the ‘overall’ row shows the overall correlation between ASR and perceptual evaluation in identifying no error as well as error conditions such as substitution error, omission error, distortion error and addition error.

Table 12

Estimation of correlation between perceptual method and ASR method, considering the target words i:ʧa, ra: ʤ:vʌ, ʤanal and ʧha:ja *as examples.*

Word	Analysis			
		ASR	Manual	Correlation
i:ʧa	Overall	22	30	0.73
	Normal	7	12	0.58
	Substitution	12	15	0.80
	Omission	0	0	-
	distortion	3	3	1.00
	Addition	**0**	**0**	**-**
ra: ʤ:vʌ	Overall	24	30	0.8
	Normal	2	3	0.67
	Substitution	6	9	0.67
	Omission	0	1	0.00
	distortion	17	17	1.00
	Addition	0	0	-
ʤanal	Overall	17	30	0.57
	Normal	0	4	0.00
	Substitution	4	8	0.50
	Omission	0	2	0.00
	distortion	13	16	0.82
	Addition	0	0	-
ʧha:ja	Overall	18	30	0.6
	Normal	3	8	0.38
	Substitution	9	15	0.6
	Omission	3	3	1
	distortion	3	4	0.75
	Addition	0	0	-

Table 13

Correlation between perceptual method and ASR method in identification of articulation errors for all in target words

Type	word	*Phone*	*Position*	Correlation					
				No error	Substitution	Omission	Distortion	Addition	Overall
Affricatives	chaya	ch	initial	0.38	0.6	1	0.75	-	0.6
	eecha		middle	0.58	0.8	-	1	-	0.73
	janal	j	initial	0	0.5	0	0.82	-	0.57
	raajav		middle	0.67	0.67	0	1	-	0.8
Approximants	je:ʃU	j	initial	0.00	1.00	0.50	1.00	-	0.77
	mUjal		middle	0.17	0.25	0.36	0.92	-	0.57
	lo:ri	l	initial	0.00	0.92	0.50	1.00	-	0.80
	Eli		middle	0.60	0.71	0.50	1.00	-	0.73
	alama:ra		middle	1.00	0.60	1.00	1.00	-	0.93
	viral		end	1.00	0.92	0.80	1.00	-	0.93
	ʧeruppʌ	r	middle	0.50	0.67	0.10	0.92	-	0.57
	vaḷa	ḷ	middle	0.50	0.75	-	1.00	-	0.77
Stops	kaḍa	k	initial	0.50	0.48	0.00	0.89	-	0.57
	t̪a:kko:l		middle	0.63	0.55	0.67	0.60	-	0.60
	t̪at̪t̪a	t	initial	0.44	0.80	0.00	1.00	-	0.63
	mo:t̪iram		middle	0.00	0.75	0.20	0.91	-	0.57
	uḍUppʌ	p		0.00	0.58	0.13	1.00	0.00	0.50
	bassʌ	b	initial	0.33	0.67	0.00	1.00	-	0.63
	pu:mba:tta	tt		0.83	0.73	0.33	1.00	-	0.73
	ratham	th	middle	0.00	0.73	0.00	1.00	-	0.73
	miṭha:ji	ṭh	middle	0.50	0.37	0.67	0.75	-	0.53

Fricatives	su:rjan	s	initial	0	0.5	0.6	1	-	0.60
	kase:ra	s	middle	0.5	0.53	0	0.88	-	0.57
	fa:n	f	initial	0.4	0.5	0	0.72	-	0.50
	ʃivan	ʃ	initial	0.5	0.22	1	1	-	0.73
	me:ʃa	ʃ	middle	0.67	0.5	0	0.75	-	0.60
	ʂarṭʌ	ʂ	initial	0.33	0.45	0.5	0.83	-	0.60
	braʂʌ	ʂ	middle	0.4	0.91	-	0.85	0	0.77
	Simham	h	middle	0	0.58	0	1	-	0.67
Nasals	ma:ŋa	ŋ	middle	0.33	0.69	0.25	1	-	0.60
	ɲaɳḍʌ	ɲ	initial	0.33	1	0	0.93	-	0.87
	ma:la	m	initial	0.64	0.93	1	-	-	0.80
	a:ma	m	middle	0.61	0.7	0	-	1	0.63
	Maram	m	final	0.4	0.45	-	1	-	0.50
	a:na	n	middle	0.4	0.9	0.25	1	-	0.67
	kiɳar	ɳ	middle	0.25	0.67	0	0.93	-	0.67
	fo:ɳ	ɳ	final	0.67	1	0.08	1	-	0.60
Vowels	aɳɳa:n	a	initial	0.16	-	-	1	1	0.47
	a:na	a:	initial	0.73	-	-	2	0.5	0.73
	e:ɳI	e:	initial	0.65	0.67	0	0.67	1	0.67
	u:ɲa:l	u:	initial	0.61	0.4	-	0.33	0.5	0.53
	i:ʧʧa	i:	initial	0.38	1	1	0.88	0.75	0.67
	uḷḷi	u	initial	0.44	1	1	1	1	0.70
	Ila	I	initial	0.37	1	0	1	-	0.57
	Eli	E	initial	0.53	0	-	1	1	0.63

Table 13 shows the correlation between perceptual evaluation and ASR method in identifying the no error condition and errors such as substitution error, omission error, distortion error, addition error and overall correlation. Table 14 computes the correlation between ASR output and perceptual evaluation for cluster reduction. In this case, the target phone is the cluster, while its mispronunciation leads to cluster reduction.

Table 14

Correlation between ASR output and perceptual evaluation in identifying cluster reduction

Type	Word	Phone	Position	Correlation	
				No error	Cluster Reduction
Clusters	pa:t̪̺ram	t̪̺r	middle	-	0.80
	vast̪̺ram	st̪̺r	middle	-	1
	gḷa:ssʌ	gḷ	initial	1	0.95
	braʂʌ	br	initial	0.11	1
	ʈrain	ʈr	initial	-	0.88
	ple:ttʌ	pl	initial	1	0.73
	biskattʌ	sk	middle	-	0.84
	pust̪̺akam	st̪̺	middle	0	0.8
	spu:ɳ	sp	initial	1	0.88

Table 17, 18, 19, 20 and 21 included in Appendix A provides correlation of articulation error identification between manual method and ASR method for clusters, vowels, nasals, fricatives, and stops, computed for all the 30 speakers.

Figure 7 displays the screenshot of the report generated by the Matlab GUI, at the end of automatic assessment. The report consists of demographic details of the subject, SODA errors of each category, PCC score, and the severity condition.

Figure 7

Screenshot of the report generated from Matlab GUI

All India Institute of Speech and Hearing
Mysore

Date : 22/07/2020 Evaluation Number : 2 Name : Ashiya Age : 5
Gender: Female Class : 1 Contact No. : 8547164760 Hometown : Kottayam

AFFRICATE
Number of substitutions : 10
Number of omissions : 0
Number of distortions : 2
Number of additions : 4

FRICATIVES
Number of substitutions : 8
Number of omissions : 4
Number of distortions : 1
Number of additions : 3

NASALS
Number of substitutions : 8
Number of omissions : 0
Number of distortions : 1
Number of additions : 2

CLUSTERS
Number of substitutions : 7
Number of omissions : 0
Number of distortions : 2
Number of additions : 4

APPROXIMANTS
Number of substitutions : 8
Number of omissions : 2
Number of distortions : 1
Number of additions : 2

STOPS
Number of substitutions : 7
Number of omissions : 1
Number of distortions : 2
Number of additions : 3

VOWELS
Number of substitutions : 6
Number of omissions : 0
Number of distortions : 1
Number of additions : 1

PCC Score = 41.304348
Condition = SEVERE

CHAPTER V

Discussion

The present study attempted to

a. Develop a software-based fully automated system for articulation test in Malayalam for children with hearing impairment by evaluating the response through digital speech processing and automatic speech recognition methods.
b. Conduct clinical trials of the developed software and validate the test results by comparing with the conventional articulation test administered by experienced speech language pathologists.
c. Deploy prototype of the developed automated articulation test system in Malayalam.

5.1 Development of automated system for articulation test in Malayalam

In the current study, a fully automatic approach to assess articulatory errors in CHI was developed. HMM-based ASR system was built first considering the target words in the standardized Malayalam articulation test. Speech samples of these target words recorded from 30 CHI formed the input to the ASR which provided a string of broad phoneme-like units as an output. Minimum Edit Distance (MED) was estimated to quantify the similarity of strings. Articulation errors were identified. Severity of the disorder was measured using the PCC score.

Automatic evaluation system developed by Novotný et al. (2014) was able to provide reliable automatic assessment, allowing the examination of a wide range of articulatory deficits connected with hypokinetic dysarthria, based on DDK utterances. However, this system considered features describing voice quality, coordination of laryngeal and supralaryngeal activity, precision of consonant articulation, tongue movement, occlusive weakening, and speech timing. The focus was on estimating the severity of impairment. In the current study, similarity was measured by matching the corresponding phoneme features (such as MFCC, LPC) of normal as well as the disordered speech using Dynamic Time Warping (DTW). Objective speech quality measures such as Itakura-

Saito (IS), Log-Likelihood Ratio (LLR), Log Area Ratio (LAR), Likelihood Ratio (LR), Weighted Cepstral Distance (WCD), Log Cepstral Distance (LCD) and Signal to Noise Ratio (SNR) were computed for normal and disordered phonemes in the present study.

An automatic assessment method for articulation errors at phone level in Hindi was developed by Bhat et al. (2015). They attempted to classify the utterance as either correct, or with substitution, omission, distortion or addition errors. They could identify substitution errors for consonants, using a rule based language model (LM) as well as tuning of acoustic models (AM) for consonants under consideration. Their system was validated on normal speech only in Hindi whereas the system developed in the present study was validated with the speech of 30 Malayalam speaking CHI.

A system for automated assessment of speech sound disorders for Cantonese-speaking pre-school children was developed by Ng et al. (2018). The articulation errors were detected by performing ASR decoding with a language model (LM) that is tailored for each test word. In the current study, errors were detected by matching the corresponding phoneme features (such as MFCC, LPC) of normal as well as the disordered speech.

5.2 Comparison of the automated system with perceptual evaluation

Analysis was carried out to determine the correlation between automated and conventional articulation tests in identifying the articulation errors. In detecting ‘no error’ condition (Table 12), the system developed in the present study showed a correlation above 0.6 for 40% of the target words, whereas the correlation was below 0.4 for only 20% of the target words. In identifying ‘substitution error’ condition (Table 12), the system developed in this study showed a correlation above 0.6 for 93% of the target words, whereas the correlation below 0.4 for only 7% of the target words. In identifying ‘omission error’ condition (Table 12), the system developed in this study showed a correlation above 0.6 for only 30% of the target words, whereas the correlation was below

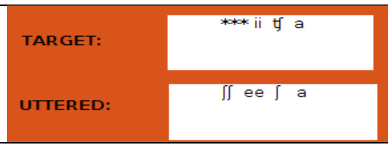
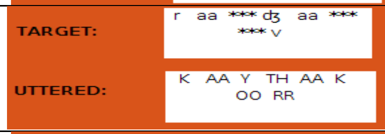
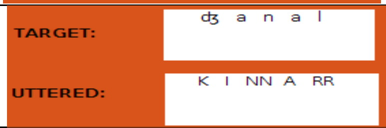
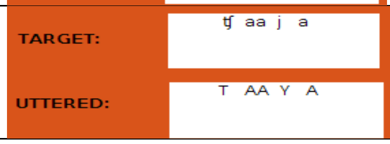
0.4 for 38% of the target words. In identifying ‘distortion error’ condition (Table 12), the system showed a correlation above 0.6 for 70% of the target words, whereas the correlation was below 0.4 for only 2% of the target words. In identifying ‘addition error’ condition (Table 12), the system showed a correlation above 0.6 for 75% of the target words, whereas for none of the target words showed a correlation below 0.4. Thus the developed system performed well in identifying the ‘substitution error’ as well as ‘addition error’. Performance was poor in identifying ‘omission error’.

5.3 Validation of prototype

In the present study, severity of articulation disorder was assessed through SODA error, MED score and PCC score. But in the case of severe articulation disorder, the ASR output was not reliable and hence computing the above values using these phoneme labels did not provide accurate results. Table 1 shows some of the results obtained for the existing system during validation with affricates. Even though first and last words (Table 1) give reasonably good ASR output, second and third words (Table 1) give unsatisfactory outputs. Results shown in the last column did not give a true reflection regarding severity of articulation in these cases. Hence incorporation of other matching techniques is needed to evaluate moderately severe and severe articulation disorders. Similarly more investigation is needed to device methods to identify the distortion error precisely.

Table 15

Comparison of automated ASR output with target and perceptual sounds during the evaluation of affricates

Word	Target	ASR out	Perceptual	MED	Screenshot of result obtained
i:ʧa	ʧ	ʃ	ʃ	1.75	
ra: ʤ:vʌ	ʤ	th	tʃ	4	
ʤanal	ʤ	K	OM	0.75	
ʧha:ja	ʧ	t	T	0.25	

The severity rating can be categorized based on the PCC score. While PCC of 85–100% is considered as mild, 65–85% signifies a mild-moderate severity. PCC score ranging from 50–65% is characterized as a moderate-severe case whereas, less than 50% exhibit a severe degree of articulation disorder. Accordingly, Subject 17 and 21 (Table 16) belong to mild to moderate category, where the perceptual scores and ASR scores are tallying. Subject 10 (Table 16) belongs to moderate to severe category, where the perceptual and ASR scores are close. In the severe category also, within the PCC range of 40% to 50%, (Subjects 6, 20 and 22) the scores do not have significant difference. For all other subjects, there is significant difference between the perceptual and ASR scores. It was noticed that, for speech of children with severe articulation disorder, the ASR generated labels are sometimes absurd, which adversely affects its accuracy. This explains the poor correlation of perceptual and ASR scores in the severe category. So, there is a need for refinement of this existing system to improve the correlation with perceptual evaluation scores in children with severe disorders also.

Table 16

PCC Final

Sl. No.	Name	Age	Sex	Perceptual PCC	ASR PCC
1	Subject 1	3	M	15.3	12.9630
2	Subject 2	4.5	M	16.9	7.4074
3	Subject 3	3.1	M	11.3	7.4074
4	Subject 4	5	M	15	9.2593
5	Subject 5	4	F	13	16.6667
6	Subject 6	5.9	M	**45.2**	**42.6296**
7	Subject 7	3.1	M	50	52.9259
8	Subject 8	4.2	F	18.8	20.3704
9	Subject 9	5	M	9.4	11.1111
10	Subject 10	4	F	**50**	**48.8889**
11	Subject 11	4	M	3.7	9.2593
12	Subject 12	4	F	15	3.7037
13	Subject 13	4.8	M	26.4	29.6296
14	Subject 14	3	M	7.5	12.9630
15	Subject 15	5	M	3.7	12.9630
16	Subject 16	4.5	F	18.6	31.4815
17	Subject 17	6	M	**73.5**	**75.1852**
18	Subject 18	3.2	M	20.7	16.6667
19	Subject 19	3.11	M	22.6	12.9630
20	Subject 20	3.5	M	**43.3**	**41.8148**
21	Subject 21	5	M	**66**	**62.7407**
22	Subject 22	3.2	F	**41.5**	**40.7407**
23	Subject 23	4	M	33.9	24.0741
24	Subject 24	4.2	F	28.3	9.2593
25	Subject 25	5	F	37.7	24.0741
26	Subject 26	3	F	5.6	11.1111
27	Subject 27	4	F	16.9	9.2593
28	Subject 28	4.3	M	1.8	3.7037
29	Subject 29	3.8	F	7.5	11.1111
30	Subject 30	3.5	F	11.3	3.7037

CHAPTER VI

Summary and Conclusions

The purpose of this study was to develop an Automatic Speech Recognition (ASR) system in Malayalam for the automatic assessment of articulation disorders. Validation of the developed system was done by comparing its results with the results of perceptual evaluation.

An automatic assessment system for articulation disorder at the phone level using Automatic Speech Recognition (ASR) techniques in Malayalam has been implemented. The proposed system was evaluated on recorded speech samples of children with different severity levels of impairment. The severity of the disorder was identified through PCC score and SODA errors at the phone level. Results suggest that the system output shows a significant correlation with the perceptual evaluation. The system performance would further improve by increasing the number of training samples. Facilities provided in the GUI for recording, graphical analysis and report generations make it directly deployable for clinical trials.

For improving the system performance in identifying the SODA error, recording of more number of speech samples belonging to the category of moderate to severe and mild to moderate is required. It was difficult to get reasonable ASR transcripts as well as accurate time alignment at the phone level for performing objective quality measures for severely disordered samples. We believe that the system can be refined further, by performing more clinical trials and by comparing the results with the existing manual/semiautomatic techniques.

6.1 Important results of the study

The important findings of the study are summarized below:-

- Words uttered by the children with articulatory disorders are automatically recognized (selected set of words for each category of sounds) and the degree of disorder is computed by matching ASR generated phoneme string to the expected phoneme string.
- Substitution, omission, distortion and addition of phonemes are identified.
- The Minimum Edit Distance (MED) was successfully used for obtaining a numerical score indicating difference between the strings.
- PCC scores were calculated by matching ASR generated labels with the expected labels.
- A reasonably good recognition accuracy was obtained for mild to moderate category.

6.2 Implications of the study

- The results of the present study will help clinicians who are dealing with children having articulation disorders.
- Existing system in Malayalam is semi-automated in nature e.g. CAPP-M (Sreedevi, Merrin & Sindhusha, 2013), whereas the developed system is fully automatic.
- By suitably selecting an objective metric that correlates with the conventional metric, it is possible to automate the evaluation of disordered speech using speech recognition system in Malayalam.
- The developed fully automated system will save the time for assessment and eliminate manual errors and biasing.

6.3 Limitations of the present study

- Improvement in recognition accuracy is required for ‘moderately severe’ to severe category.
- Severity of articulation disorder was assessed through SODA error, MED score and PCC score. While using this system for speech of children with ‘severe’ articulation disorder, it was difficult to get correct ASR transcripts. Hence, in several cases, SODA errors as well as PCC score calculation using erroneous ASR transcripts were not accurate.
- Time stamps of phonemes were not obtained in some cases especially for short words.

6.4 Future recommendations

- The design of the developed system needs to be modified using other signal processing and matching techniques to improve automated articulation test of children with severe disorders.
- The objective measures such as DTW score, Log Cepstral Distance (LCD) and Log-Likelihood Ratio (LLR) need to be combined with ASR scores to enhance the performance of the developed system in assessment of children with severe disorders.
- Matching of spectrogram images at the word level may be considered for enhancing the performance of the system.

6.5 Significance of the results of the study

The significance of the results of the present study should be seen in the following context:

Through the present study, we were able to implement a fully automated system for the evaluation of articulation disorder in Malayalam speaking children. This system will save the time for assessment, reduce the effort of the SLP and eliminate manual errors and biasing. This will also give a motivation for automating articulation tests in other South Indian Languages.

References

Barsties, B., & De Bodt, M. (2015). Assessment of voice quality: current state-of-the-art. *Auris Nasus Larynx*, *42*(3), 183-188.

Bhat, C., Vachhani, B., & Kopparapu, S. (2015). Automatic assessment of articulation errors in Hindi speech at phone level. In *TENCON 2015-2015 IEEE Region 10 Conference.*

Clapham, R. P., Martens, J. P., van Son, R. J., Hilgers, F. J., van den Brekel, M. M., & Middag, C. (2016). Computing scores of voice quality and speech intelligibility in tracheoesophageal speech for speech stimuli of varying lengths. *Computer Speech & Language, 37*, 1-10.

Georgoulas, G., Georgopoulos, V. C., & Stylios, C. D. (2006). Speech sound classification and detection of articulation disorders with support vector machines and wavelets. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society* . 2199-2202.

Gold, T. G. (1978). *Speech and hearing skills: A comparison between hard-of-hearing and deaf children* [Doctoral dissertation]. ProQuest Information & Learning.

Gordon-Brannan, M. E., & Weiss, C. E. (2007). *Clinical management of articulatory and phonologic disorders*. Lippincott Williams & Wilkins.

Gu, L., Harris, J. G., Shrivastav, R., & Sapienza, C. (2005). Disordered speech assessment using automatic methods based on quantitative measures. *EURASIP Journal on Advances in Signal Processing*.

Hutchins, Say & See (1992). Articulation Therapy Software. *Proceedings of the Johns Hopkins national search for computing Applications to assist persons with disabilities*. 37-40.

Thennattil, J. J., & Mary, L. (2016). Phonetic engine for continuous speech in Malayalam. *IETE Journal of Research*, *62*(5), 679-685.

Lousada, M., Jesus, L. M., Capelas, S., Margaça, C., Simões, D., Valente, A., & Joffe, V. L. (2013). Phonological and articulation treatment approaches in Portuguese children with speech and language impairments: a randomized controlled intervention study. *International journal of language & communication disorders, 48*(2), 172-187.

Maier, A., Hacker, C., Noth, E., Nkenke, E., Haderlein, T., Rosanowski, F., & Schuster, M. (2006). Intelligibility of children with cleft lip and palate: Evaluation by speech recognition techniques. *18th International Conference on Pattern Recognition (ICPR'06). 4,* 274-277. IEEE.

Ng, S. I., Tao, D., Wang, J., Jiang, Y., Ng, W. Y., & Lee, T. (2018). An Automated Assessment Tool for Child Speech Disorders. In *2018 11th International Symposium on Chinese Spoken Language Processing (ISCSLP).* 493-494. IEEE.

Novotný, M., Rusz, J., Čmejla, R., & Růžička, E. (2014). Automatic evaluation of articulatory disorders in Parkinson's disease. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *22*(9), 1366-1378.

Nieto, R. G., Marín-Hurtado, J. I., Capacho-Valbuena, L. M., Suarez, A. A., & Bolaños, E. A. B. (2014). Pattern recognition of hypernasality in voice of patients with Cleft and Lip Palate. *XIX Symposium on Image, Signal Processing and Artificial Vision* [Symposium]. IEEE.

Kassa, W. E., Billabert, A. L., Faci, S., & Algani, C. (2013). Electrical modeling of semiconductor laser diode for heterodyne RoF system simulation. *IEEE Journal of Quantum Electronics*, *49*(10), 894-900.

Sreedevi, N., John, M., & Chandran, S. (2013). Computerized assessment of phonological processes in Malayalam (CAPP-M). *2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite (GHTC-SAS)*. 264-269.

Story, B. H., & Bunton, K. (2017). Vowel space density as an indicator of speech performance. *The Journal of the Acoustical Society of America*, *141*(5), EL458-EL464.

van Son, R., Middag, C., & Demuynck, K. (2018). *Vowel Space as a Tool to Evaluate Articulation Problems*. Interspeech, Conference, Hyderabad, India . 357-361.

Voice and Speech Systems. (n.d.). Phonic Drill: Mis-pronunciation- Vaghmi Therapy Pronunciation Module. http://voiceandspeechsystems.com/index.php/phonic-drill-mis-pronunciation/

Appendix – A

Results of correlation of articulation error identification between perceptual evaluation and ASR method

Table 17

Correlation of articulation error identification between perceptual evaluation and ASR method for clusters, computed for all the participants (CHI)

Speakers	Age	Gender	Clusters											
			INITIAL						MIDDLE					
			NE	CR	S	O	D	A	NE	CR	S	O	D	A
Subject1	3	M	1	1	-	-	1	-	-	1	-	-	1	-
Subject2	4.5	M	0	1	-	-	1	-	0	1	-	-	1	-
Subject3	3.1	M	0	1	-	-	1	-	-	1	-	-	0.67	-
Subject 4	5	M	-	0.67	-	-	1	-	0	1	-	-	-	-
Subject 5	4	M	0	1	-	-	1	-	-	1	-	-	-	-
Subject 6	4	M	-	1	-	-	1	-	-	1	-	-	-	-
Subject 7	3.10	M	-	0.75	-	-	-	-	-	0.75	-	-	-	-
Subject 8	4.2	F	0.5	-	-	-	0.67	-	-	0.75	-	-	-	-
Subject 9	5	M	0	0.67	-	-	1	-	-	1	-	-	-	-
Subject 10	4	F	1	1	1	-	-	-	-	0.5	-	-	-	-
Subject 11	4	F	-	1	-	-	1	-	-	1	-	-	1	-
Subject 12	4	F	1	1	0	-	1	-	-	1	-	-	-	-
Subject 13	4.8	F	-	0.8	-	-	-	-	-	1	-	-	-	-
Subject 14	3	F	-	1	-	-	1	-	-	-	-	-	1	-
Subject 15	6	F	-	-	-	-	0.8	-	-	-	-	-	1	-
Subject 16	4.5	F	-	1	-	-	-	-	-	0.5	-	-	-	-
Subject 17	6	M	-	1	-	-	1	-	-	0.5	-	-	-	-
Subject 18	5	M	-	0.67	1	-	1	-	-	1	-	-	-	-
Subject 19	3.11	M	-	1	-	-	-	-	0	1	-	-	-	-
Subject 20	3.5	M	0	0.75	-	-	-	-	-	1	-	-	-	-
Subject 21	5	M	1	0.67	-	-	-	-	0	1	-	-	-	0
Subject 22	3.2	M	-	0.8	-	-	-	-	-	1	-	-	-	-
Subject 23	4	M	-	0.6	-	-	-	-	-	1	0	-	-	-
Subject 24	4.2	F	0	1	-	-	-	-	-	0.75	-	-	-	-
Subject 25	5	F	1	1	-	-	-	-	-	1	1	-	-	-
Subject 26	3	F	-	1	-	-	-	-	-	-	-	-	1	-
Subject 27	4	F	-	1	-	-	-	-	-	1	-	-	1	-
Subject 28	4.3	F	-	1	-	-	-	-	-	0.75	-	-	-	-
Subject 29	3.8	F	0	1	-	-	-	-	-	0.5	-	-	-	-
Subject 30	3.5	F	-	-	-	-	-	-	-	-	-	-	1	-

Note:- NE – No Error, CR – Cluster Reduction, S – Substitution, O – Omission, D – Distortion and A - Addition

Table 18

Correlation of articulation error identification between perceptual evaluation and ASR method for vowels, computed for all the participants (CHI)

Speakers	Age	Gender	Vowels					
			INITIAL					
			NE	CR	S	O	D	A
Subject1	3	M	0.43	1	-	-	-	1
Subject2	4.5	M	1	1	1	-	1	-
Subject3	3.1	M	0	1	0.5	1	0.5	-
Subject 4	5	M	0.25	0.67	-	-	0.67	0
Subject 5	4	M	0.4	1	-	-	1	-
Subject 6	4	M	0.72	1	-	-	-	1
Subject 7	3.10	M	0.43	0.75	1	-	-	-
Subject 8	4.2	F	0.58	-	-	-	-	1
Subject 9	5	M	0.33	0.67	-	-	-	0.5
Subject 10	4	F	0.71	1	0	-	-	-
Subject 11	4	F	0.25	1	1	-	1	0
Subject 12	4	F	0	1	1	-	1	-
Subject 13	4.8	F	0.86	0.8	-	-	1	-
Subject 14	3	F	1	1	-	-	1	-
Subject 15	6	F	0.33	-	-	-	1	-
Subject 16	4.5	F	0.67	1	0.33	-	0	1
Subject 17	6	M	0.5	1	-	-	0.5	-
Subject 18	5	M	0.25	0.67	1	-	1	-
Subject 19	3.11	M	0.4	1	1	-	1	-
Subject 20	3.5	M	0.625	0.75	-	-	-	-
Subject 21	5	M	0.67	0.67	-	-	-	0.5
Subject 22	3.2	M	0.75	0.8	-	-	-	-
Subject 23	4	M	0.43	0.6	-	-	1	-
Subject 24	4.2	F	0.2	1	0.5	-	-	1
Subject 25	5	F	1	1	0.5	-	-	1
Subject 26	3	F	0	1	-	0	1	-
Subject 27	4	F	0	1	0	1	1	1
Subject 28	4.3	F	0	1	0.5	0	1	1
Subject 29	3.8	F	0.33	1	-	-	1	1
Subject 30	3.5	F	1.	-	0.5	-	1	-

Note:- NE – No Error, CR – Cluster Reduction, S – Substitution, O – Omission, D – Distortion and A - Addition

Table 19

Correlation of articulation error identification between perceptual evaluation and ASR method for nasals, computed for all the participants (CHI)

Speakers	Age	Gender	Nasals														
			INITIAL					MIDDLE					FINAL				
			NE	S	O	D	A	NE	S	O	D	A	NE	S	O	D	A
Subject1	3	M	-	1	-	1	-	-	1	-	1	-	-	1	-	1	-
Subject2	4.5	M	-	1	-	1	-	-	0.66	-	1	-	-	0	-	1	-
Subject3	3.1	M	-	-	0. 5	-	-	-	0.5	0	-	-	-	-	1	1	-
Subject 4	5	M	-	1	-	1	-	-	0.66	-	1	-	-	0	-	1	-
Subject 5	4	M	1	-	-	-	-	-	-	0	-	-	-	1	-	1	-
Subject 6	4	M	-	1	-	0	-	-	1	-	-	-	1	1	-	-	-
Subject 7	3.10	M	0	1	-	-	-	-	0	-	-	-	0.5	-	-	-	-
Subject 8	4.2	F	-	1	-	1	-	-	1	-	1	-	0	1	-	-	-
Subject 9	5	M	-	1	-	1	-	-	1	-	1	-	-	0	-	1	-
Subject 10	4	F	1	1	-	-	-	-	1	-	-	-	0	-	0	-	-
Subject 11	4	F	-	1	-	1	-	-	0	-	1	-	0	-	0	-	-
Subject 12	4	F	-	1	-	-	-	-	0.66	-	1	-	-	-	0	1	-
Subject 13	4.8	F	1	1	-	-	-	-	-	0	-	-	0	-	0	-	-
Subject 14	3	F	-	1	-	1	-	-	-	-	1	-	-	0	-	-	-
Subject 15	6	F	-	1	-	1	-	-	1	-	1	-	-	1	0	1	-
Subject 16	4.5	F	0	-	-	1	-	-	0.5	-	-	-	1	-	-	-	-
Subject 17	6	M	0.5	-	-	-	-	-	1	-	0	-	1	1	-	1	-
Subject 18	5	M	1	1	-	-	-	-	0	-	1	-	0	-	0	-	-
Subject 19	3.11	M	-	0.5	-	-	-	1	0.5	-	-	1	-	-	0	-	-
Subject 20	3.5	M	0	-	-	-	-	-	1	-	1	-	-	1	0	-	-
Subject 21	5	M	1	-	-	1	-	-	1	-	0.5	-	0.5	-	-	-	-
Subject 22	3.2	M	1	1	-	-	-	-	1	-	-	-	1	-	-	-	-
Subject 23	4	M	1	1	-	-	-	-	0	-	-	-	0	-	0	-	-
Subject 24	4.2	F	1	1	-	-	-	-	-	0.5	1	-	0	-	-	1	-
Subject 25	5	F	-	1	-	-	-	-	0	-	1	-	0	-	-	1	-
Subject 26	3	F	-	1	-	-	-	-	1	0	1	-	-	1	-	1	-
Subject 27	4	F	0	1	-	-	-	-	1	0	1	-	-	-	0	1	-
Subject 28	4.3	F	-	1	-	1	-	-	1	1	-	-	-	0	-	1	-
Subject 29	3.8	F	-	1	-	1	-	-	0.5	-	1	-	-	0	-	1	-
Subject 30	3.5	F	0	-	-	1	-	-	1	-	1	-	1	-	0	-	-

Note:- NE – No Error, CR – Cluster Reduction, S – Substitution, O – Omission, D – Distortion and A - Addition

Table 20

Correlation of articulation error identification between perceptual evaluation and ASR method for fricatives, computed for all the participants (CHI)

Speakers	Age	Gender	Fricatives									
			INITIAL					MIDDLE				
			NE	S	O	D	A	NE	S	O	D	A
Subject1	3	M	-	0.5	-	1	-	0	0	0	-	-
Subject2	4.5	M	0	-	-	1	-	-	-	0	1	-
Subject3	3.1	M	-	1	1	1	-	0	-	0	1	-
Subject 4	5	M	-	1	-	1	-	1	1	0	1	-
Subject 5	4	M	-	1	0.66	-	-	0	0	-	1	-
Subject 6	4	M	-	0.33	-	1	-	-	0.5	-	-	-
Subject 7	3.10	M	-	0.33	-	1	-	-	0.5	-	-	-
Subject 8	4.2	F	-	0.33	0	-	-	-	0.75	-	-	-
Subject 9	5	M	-	0	-	1	-	-	1	-	0.66	-
Subject 10	4	F	0.5	0	-	-	-	1	1	-	-	-
Subject 11	4	F	-	0	-	1	-	0	-	-	0.66	-
Subject 12	4	F	0	0	-	1	-	-	1	-	-	-
Subject 13	4.8	F	-	0.5	0.5	-	-	-	0.5	-	-	-
Subject 14	3	F	-	1	-	1	-	0	-	-	1	-
Subject 15	6	F	-	1	0	1	-	-	-	0	0.75	-
Subject 16	4.5	F	-	0	1	1	-	-	0.5	-	1	-
Subject 17	6	M	0.66	0	-	-	-	0.5	0.5	-	-	-
Subject 18	5	M	-	0	0	1	-	1	0	-	1	-
Subject 19	3.11	M	-	0.66	-	1	-	-	0.75	-	-	-
Subject 20	3.5	M	0	1	1	-	-	0.5	-	0	1	-
Subject 21	5	M	0	0.5	-	1	-	1	1	-	0	0
Subject 22	3.2	M	-	0.5	-	0	-	-	0.5	-	1	-
Subject 23	4	M	-	0.75	-	-	-	-	0.33	0	1	-
Subject 24	4.2	F	0.5	-	1	1	-	-	1	-	-	-
Subject 25	5	F	1	0	-	-	-	1	0	-	1	-
Subject 26	3	F	0	-	0	1	-	-	0.5	-	1	-
Subject 27	4	F	-	0	-	1		-	1	-	1	-
Subject 28	4.3	F	-	1	-	0.33	-	-	1	-	1	-
Subject 29	3.8	F	-	-	-	0.25	-	-	0.5	-	0.5	-
Subject 30	3.5	F	-	-	-	1	-	-	-	-	1	-

Note:- NE – No Error, CR – Cluster Reduction, S – Substitution, O – Omission, D – Distortion and A - Addition

Table 21

Correlation of articulation error identification between perceptual evaluation and ASR method for stops, computed for all the participants (CHI)

Speakers	Age	Gender	Stops									
			INITIAL					MIDDLE				
			NE	S	O	D	A	NE	S	O	D	A
Subject1	3	M	-	1	-	-	-	0	0.75	-	1	-
Subject2	4.5	M	-	0.5	-	1	-	-	1	0.66	1	-
Subject3	3.1	M	-	-	0	1	-	-	1.000	0	1	-
Subject 4	5	M	-	1	-	-	-	-	0.333	0	1	-
Subject 5	4	M	1	1	-	1	-	-	0.5	0	1	-
Subject 6	4	M	0	1	-	-	-	0.66	1	-	-	-
Subject 7	3.10	M	0.5	0	-	-	-	0.66	0.66	-	-	-
Subject 8	4.2	F	0	0	-	1	-	0	0.33	-	1	-
Subject 9	5	M	0	-	-	1	-	-	0	0.5	1	-
Subject 10	4	F	-	0.66	0	-	-	0.5	0.66	0	-	-
Subject 11	4	F	-	-	-	1	-	-	-	1	1	-
Subject 12	4	F	-	1	0	1	-	-	0.5	0	1	-
Subject 13	4.8	F	1	0	-	-	-	0	0	0.5	1	-
Subject 14	3	F	-	0.5	-	1	-	-	1	-	1	-
Subject 15	6	F	0.5	1	-	-	-	-	0	-	0.66	-
Subject 16	4.5	F	-	0.66	-	-	-	0.5	1	-	0.66	-
Subject 17	6	M	0.66	-	0	-	-	1	0.5	0	-	-
Subject 18	5	M	-	-	-	1	-	1	0.66	1	1	-
Subject 19	3.11	M	0	0.5	-	-	-	-	0.5	0	-	-
Subject 20	3.5	M	-	1	-	1	-	0	1	0	-	-
Subject 21	5	M	1	1	-	-	-	0.25	0.5	-	-	0
Subject 22	3.2	M	0	1	-	-	-	1	0.66	0	-	-
Subject 23	4	M	0.5	0	-	-	-	0	0.5	-	1	-
Subject 24	4.2	F	-	0.33	-	-	-	1	0.75	-	1	-
Subject 25	5	F	0	1	-	1	-	0.5	0.66	-	0	-
Subject 26	3	F	-	0.5	-	0	-	1	0.66	0	0	-
Subject 27	4	F	-	0.33	-	-	-	-	0	1	1	-
Subject 28	4.3	F	-	-	1	1	-	0	1	1	1	-
Subject 29	3.8	F	-	0.33	-	-	-	0	1	-	1	-
Subject 30	3.5	F	-	-	-	1	-	-	-	-	1	-

Note:- NE – No Error, CR – Cluster Reduction, S – Substitution, O – Omission, D – Distortion and A - Addition