# Diagnosis of Disordered Speech using Automatic Speech Recognition

Arpitha V
Department of TCE
GSSSIETW
Mysuru, India

Samvrudhi K
Department of TCE
GSSSIETW
Mysuru, India

Manjula G
Associate Professor
Department of TCE
GSSSIETW
Mysuru, India

Sowmya J
Department of TCE
GSSSIETW
Mysuru, India

Thanushree G B
Department of TCE
GSSSIETW
Mysuru, India

***Abstract*—The purpose of the project is to help the speech disordered people communicate effectively. Speech is the effective way of communication. Many speech disorders are caused due to stroke. In this paper speech disorders in adults and various software available in the market has been described. The people suffering from speech disorders produce a 17.6% of mispronunciations in the phonetic level. The study of speech signals and the methods to process them is called speech processing. In order to study the speech signals methods such as speech coding, speech synthesis, speech recognition and speaker recognition plays a vital role. Among all these, speech recognition is considered in this paper. Speech recognition is converting the acoustic signals obtained from the speaker which is given as an input to microphone. The obtained input is then generated as a set of words. There must be electronic circuits in order to extract linguistic properties obtained from the speaker. This project is designed using MFCC as a feature extractor and support vector machine as the classifier.**

***Keywords*— ASR,Disorders,MFCC, Mispronunciations,Speech recognition,SVM**

## I. INTRODUCTION

Speech is considered to be the natural and effective mode of communication. Speech relies on language capabilities. The verbal communication consists of articulation, voice and fluency. Language is a set of rules which consists of phonetics, syntax, semantics, morphology and pragmatics. The difficulties faced during comprehending and expressing while speaking and writing is identified to be a language disorder. Issues faced while communicating hinders the power of information shared by the individual. Speech disorders are caused due to data flow restriction and also due to neurological changes within the brain. Speech disorders occur irrespective of age around the world which affects the standard of living. According to National Institute on Deafness and Communication (NIDCD), in 2016 approximately 7.5 million people have language impairments. And the fact is that all those people live in United States of America. So we can imagine in the whole world how many would be suffering from language impairments. In order to overcome any speech and language impairment, right diagnosis plays a remarkable role. Thus, the procedure taken to assess must be particular and this may lead to a faster recovery rate.

## II. TYPES OF DISORDER

Symptoms of speech disorders can be classified based on the severity of disorder. There are many causes for speech disorders which include damage of brain or injuries in head after the stroke, weaknesses in muscles, damages in vocal cords, dementia, cancer, autism, down syndrome, hearing loss and so on.

The types of speech disorder include:

A. Stuttering
B. Dysarthria
C. Apraxia
D. Cluttering
E. Lisping

Let us study the different types of disorder in detail.

### *A. Stuttering*

Stuttering is a type of speech disorder that interrupts the flow of speech. People suffering from stuttering experiences: Repetitions, blocks, prolongations. Repetitions are the involuntary repetitions of sounds, vowel or words. Blocks are known what to say but have difficulty making the speech sounds. Prolongations are stretching or drawing of particular words or sounds. Depending upon the situation such as stress, excitement or frustration the symptoms of stuttering can be determined. Stuttering can cause both behavioral and physical symptoms that occur at the same time which includes: tension in the face and shoulders, rapid blinking, lip tremors, clenched fists and sudden head movements.

Types of stuttering includes: developmental stuttering and Neurogenic stuttering. Most of the young children in learning speech and language skills stage experiences developmental stuttering Genetic factors may also cause increase in persons developing this type of stutter. Neurogenic Stuttering are the ones that occurs when there is damage to the brain which prevents the coordination between the regions which play a role in speech.

### *B. Dysarthria*

Dysarthria is very similar to apraxia but due to the damage in the brain muscle weakness can be seen in a person’s face, lips, tongue, throat or chest. Because of the muscle weakness in those body parts speaking is very difficult for the people suffering from dysarthria. The symptoms of dysarthria includes slurred speech, mumbling, slow or quick speaking, quiet or soft speech.

### *C. Apraxia*

The brain is the organ that controls every single action that people make which also includes speaking. The involvement by the brain can be unconscious and automatic. When the person decides to speak signals are sent by the brain to the different organs of the body that works to produce a speech. The brain gives the instructions to those organs how and when to move to form the appropriate sounds. For example, speech signal produced bythe brain opens or closes the vocal cords, movement of the tongue and shaping of the lips and also controlling the movement of air through the throat and mouth.

Apraxia occurs when there is damage in the brain which impairs a person's motor skills and it's affected in any part of the body. Specifically apraxia can be termed into apraxia of speech or verbal apraxia. This type of apraxia causes impairment of motor skills that affects the person's ability to produce sounds of speech correctly even if a person has the knowledge of right word usage.

### *D. Cluttering*

Cluttering is a problem that makes difficulty in understanding a person's speech. Cluttering is same as stuttering which affects the fluency of person's speech. The difference between stuttering and cluttering is that cluttering is a language disorder whereas stuttering is a speech disorder. People who clutter can communicate what they are thinking but it becomes disorganized as they are about to speak. So person affected by cluttering may pause in unexpected places. The rhythm of cluttered speech may sound jerky and speaker may not be aware of this problem.

### *E. Lisping*

Lisping is one among the articulatory issue whose causes are indistinguishable. For one, in different assortments stuttering happens including tongue tip distending between front teeth along with a slurping commotion in the cheek pocket. In a few nasal and grunting subtypes happens just like the substitute murmuring in the throat or inside the larynx. The reasons for lisping also includes anomalous number or position of teeth, impersonation of different lispers, lack of palatal conclusion, slight hearing problems in the higher frequencies, few psychologic problems. Lisping is less effective and may persevere into grown up life if not amended.

## III. LITERATURE SURVEY

The author says, in disordered speech recognition lexical adaptation approach is used in Automatic Speech Recognition technique (ASR). Speech disorder in individual occurs because of the physiological disorder in the vocal tract which affects the speech of individual. The aim of the author is to achieve lexical adaptation for the young speakers who are affected with different speech disorder. By making use of ASR technology a average accuracy of 31.96% was obtained in Word Error Rate (WER) and 48.25% was obtained in Phoneme Error Rate (PER). The author took a data sample from 14 young speakers whose age is between 11 to 21 years (7 boys and 7 girls). The classifier is Hidden Markov Model (HMM). The improvements can be inclusion of new lexical variants in vocabulary since the vocabulary considered here is only 55 words. [1]

The paper deals with the speech recognition of Dysarthric speech. Dysarthric in individual is caused due to paralysis, weakness or because of weakness in muscles. Since making use of Acoustic model gives less accuracy and less intelligibility they have adopted "metamodels" which overcomes some of the problems of acoustic model. Recognition of dysarthric speech is bit challenging because the pronunciation made by the patient is different from normal speakers, rate of speech is lower and the words pronounced will be different. To overcome this Maximum Likelihood Linear Regression technique (MLLR) along with Artificial Neural Networks (ANN) and Hidden Markov Model (HMM) is used. For speech recognition, a recognizer was built using HTK Toolkit by collecting the data from 92 speakers. The overall accuracy obtained was around 90%. Since only small set of data is validated here the improvement can be made so that a large set of data can be validated. [2]

According to this paper disordered speech analysis can be done by making use of Automatic Speech Recognition (ASR) technique. Here 6 different types of speech disorder are analyzed to that automatic recognition is difficult. For the detection of five speech signals Mel-frequency cepstral coefficients (MFCC) along with Gaussian mixture model (GMM) / Hidden Markov model (HMM) is used. Discrimination between normal speech and disordered speech was given by correlation. For correlation a data set of 50 Arabic speech affected speakers and 1 normal speaker was considered. The accuracy obtained in recognition lies between 56% to 82.50% in case of disordered speech and 100% for normal speech. [3]

The Author, has used voice-input voice-output communication aid (VIVOCA) which recognizes the disordered speech and builds a message which is then converted into synthetic speech. Previously voice-output communication (VOCA) was used which consists of keyboards and relays where the user needs to give the input by typing which was tiring many users to overcome this problem VIVOCA is used. For speech recognition Automatic Speech Recognition (ASR) along with Hidden Markov model (HMM) and for the training of acoustic models Mel-frequency cepstral coefficients (MFCCs) were used. By making use of VIVOCA around 96% efficiency was obtained. The further improvements could be to design VIVOCA device that supports a better-quality internal microphone. [4]

The author, says that the dysarthric speakers often cannot get the satisfactory performances by the off-shelf automatic speech recognition hence the experiments described in this paper uses UA speech which is one of the largest dysarthric databases. This study is based on the feature like LVCSR by using MAP and MLLR techniques.

The study is carried out by considering data set of 15 speakers. The speakers in total produced 70 minutes of speech. Research of this paper is supported by the study of use of speech-driven assistive technology for disabled and elderly persons .The experiment resulted in the efficiency of 50.9%. [5][16]

The author has conducted experiments that suggested frequency band of 1-1562 Hz which exhibits the differences in normal and pathological subject spectrum. The detection rate obtained in this band is 91.28%. The fractal dimensions are calculated using Katz and Higuchi algorithms and are provided to SVM for classification. The results of the experiment shows that the fractal dimension is successful in capturing the transient behavior of disordered speech which further leads to the characterization of the signals. The accuracy is enhanced by concatenating fractal dimensions with MDVP parameters which provided an accuracy of 94.71%. However the obtained accuracy is 94.07%.[6]

In previous works, software was designed and made applicable only in the speech dependent applications. Accuracy achieved from the system was comparatively less, i.e. it was ranging from 31.96% to 91.28%. The software supported only one language and recognized only one type of speech disorder. The recognized disorders were corrected and text outputs were given. In order to overcome all the above drawbacks, we are designing the software which works as both speech dependent and speech independent. The designed software is capable of recognizing more than one type of speech disorder. In this software the speech is corrected and the output is given as both text and speech which can be used in multiple applications.

## IV. METHODOLOGY

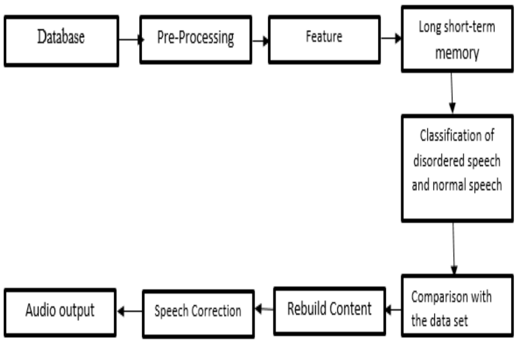


Fig. 1. Block Diagram

In the proposed disordered speech recognition model we have considered a dataset of 30 speakers of different age groups who were suffering from stuttered speech. The speech signal extracted from the dataset is processed in order to convert it to text format. Later, Pre-processing will be done in order to remove the noise from the speech signal and to compress the wavelength. In Pre-processing it involves two steps Denoising and Wavelength Compression and it is performed using Dubachies Wavelet Transform which is the built-in wavelet function in python. In pre-processing bandpass filter is used in order to allow the signals of selected frequencies that is, from 1Khz-30Khz.

Feature extraction is done in order to illustrate a speech signal by a predetermined number of components of the signal. In order to extract features from speech signal Mel Frequency Cepstral coefficients (MFCC) is used. Block diagram for MFCC is as shown in the figure 2:

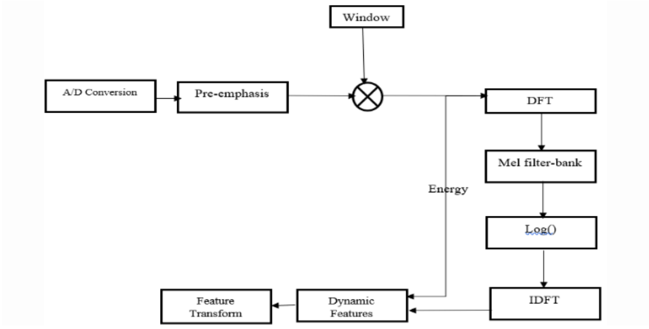


Fig. 2. Block Diagram of MFCC

### A. *Pre–emphasis*

x[n] is the speech signal which is then sent to high-pass filter. The equation is given below,

$$y[n] = x\,[n] - a \times x\,[n - 1]\ (1)\ (1)$$

Where y [n] = output signal.

The value of ‘a’ normally lies in between 0.9 to 1.0. The Z transform of this equation is given by:

$$H(z) = 1\text{-}\ a/(z\text{-}1)$$

In order to make continuous frame from first to the last point, each frame is multiplied with the window. If the signal in a frame is denoted by

$$x\,[n],\ n = 0,\ldots N\text{-}1$$

The signal after hamming windowing is,

$$x[n] \times w[n]$$

Where w[n] = hamming window which is defined by

$$w[n]=0.54\text{-}0.46\times\cos(2\pi n/(N\text{-}1))$$

Where $0 \leq n \leq N\text{-}1$

### B. *Mel Filter Processing*

The Mel frequency scale is linearly spaced below 1000Hz and has logarithmic spacing above 1000Hz. The mesh is computed from the given formula for a given frequency f is in Hz:

$$Mel(f)=2595 \times \log10(1+ f/700)$$

Speech signals which are produced by humans are not linear, so, one approach is to use filter bank in order to represent the spectrum. A filter bank uses each desired Mel frequency component. Mel spectrum coefficients, K, is typically taken as 20.

### C. *Signal delay*

$$y(t) = x(t) + decay{*}x(t\text{-}delay)$$

Raw Data: Analog data is converted to digital data
Sampling rate 8000 samples/sec
Storage space for one sample: 8bit

*Total data size = Number of samples * Storage space for one sample*
*= Samples/sec * Number of seconds * Storage space*
*= 8000 * 60 * 8 bits*

*= 3840 Kbits*
*Bit rate = Samples/sec * Storage space for one sample = 64 Kbits/sec*

### *D. Raw waveform input*

Before we dig deeper into acoustic features that we can use to build speech recognition models, we have to understand what we are recording. A microphone records sound commonly at a sampling rate of 44100 Hz. But what exactly is sound? Well, we can define sound as some sort of elastic pressure waves that are caused by some vibration and are travelling through a transmission medium such as air or water. A microphone samples the differences of air pressure to ambient pressure to record these waves.

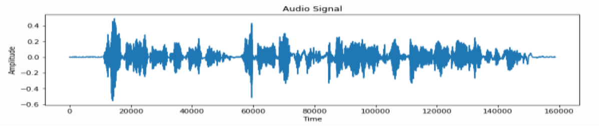


Fig. 3. Raw input waveform

### *E. Windowing*

Using spectrograms has the disadvantage that it covers frequencies we humans cannot hear as well. To reduce dimensionality, we could reduce the spectrogram to what we can hear. This does not only include minimum and maximum frequencies that we can hear but differences between frequencies as well. This concept is used by lossy audio encoders as well.

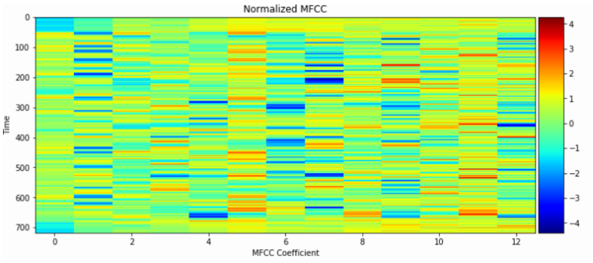


Fig. 4. MFCC output waveform

### *F. Long Short-Term Memory (LSTM)*

It is a type of recurrent neural network which is capable of learning order dependence in sequence prediction problems. This deportment is required in complex problem domains such as speech recognition, machine translation and many more. LSTMs are a complex area of deep learning.

There are several architectures of LSTM units. The most commonly used architecture includes a cell and three "regulators", usually called gates, which is used in the flow of information inside the LSTM unit: an input gate, an output gate and a forget gate. By removing one or more gates variations can be made in LSTM unit.

### *G. Support vector machine (SVM)*

In machine learning, Support vector machine(SVM) are supervised learning models with associated learning algorithms that analyze data used for classification and multivariate analysis. It is mostly employed in classification problems. In this algorithm, each data unit is plotted as a point in n-dimensional space where n is number of features, with the value of each feature being the value of a particular coordinate. Then, classification is performed by finding the hyper-plane that best differentiates the two classes. Linear classification, non-linear classification and implicitly mapping their inputs into high dimensional feature spaces can be efficiently performed by SVMs.

SVM can be considered as a best classifier when compared to others because of the following reasons,

1) It is suitable when the number of features and number of training data is very large (say millions of features and millions of instances (data)).
2) When sparsity in the problem is very high, i.e., most of thefeatures have zero value.
3) It is best for document classification problems where sparsity is high and features/instances are also very high.
4) It also performs very well for problems like image classification, genesclassification, drug disambiguation etc. where number of features is high.

Good range of algorithms is proposed which make use of problem structures and other smaller things like problem shrinking during optimization etc. Classifier makes a model from the training data and predicts the target values of the test data.

After all these steps disordered speech will be compared with the normal speech and the disordered speech will be corrected into normal speech and the output will be given in both speech and text format.

After all these steps disordered speech will be compared with the dataset of normal speech and the content is rebuild according to the compared output. Further speech correction takes place and then the corrected speech signal is given as the audio as well as text output.

## V. APPLICATIONS

### *a) Assistance to Vocally Handicapped*

A hand-held, powered artificial speech aid will be employed by vocally handicapped to specify their words. This instrument have specially designed keyboard which accepts the input and converts it to the specified speech in fraction of second.

### *b) Source of Learning for Visually Impaired*

Listening is a crucial talent for folks who are visually impaired. Blind people admit their ability to listen or gain data quickly and with efficiency. Using sensor of hearing students realize the data from books or CD, however conjointly to access what is happening around them.

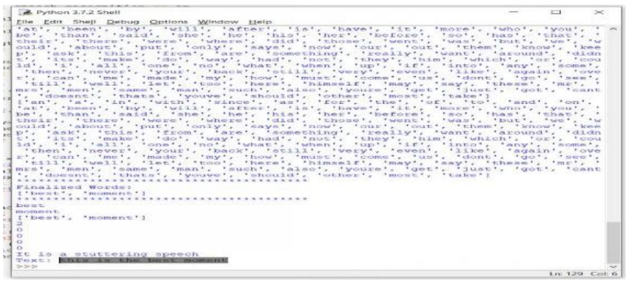
### *c) Talking Books and Toys*

Audio recording not only teaches a way to scan however conjointly has additional impact on students than text or reading within the same approach talking toys square measure an excellent supply of fun and amusement for youngsters.

## VI. RESULTS

After processing of disordered speech it will show what kind of speech disorder the person is having and the output will be given in text form and also in speech format. The accuracy obtained from the model is approximately 92%.

Output from the disordered speech recognition model is as shown in the below figures.



```
'at', 'been', 'by', 'will', 'after', 'is', 'have', 'it', 'more', 'who', 'you', '
be', 'than', 'said', 'she', 'he', 'his', 'her', 'before', 'so', 'has', 'that', '
their', 'there', 'were', 'where', 'did', 'those', 'went', 'was', 'but', 'we', 'w
ould', 'about', 'put', 'only', 'says', 'now', 'our', 'out', 'them', 'know', 'kee
p', 'ask', 'this', 'from', 'are', 'something', 'really', 'want', 'around', 'didn
t', 'its', 'make', 'do', 'way', 'had', 'not', 'they', 'him', 'which', 'or', 'cou
ld', 'i', 'all', 'one', 'no', 'what', 'when', 'up', 'if', 'into', 'any', 'some',
 'then', 'never', 'your', 'back', 'still', 'very', 'even', 'like', 'again', 'ove
r', 'can', 'me', 'made', 'my', 'how', 'must', 'come', 'us', 'dont', 'go', 'see',
 'till', 'well', 'let', 'too', 'here', 'himself', 'may', 'say', 'these', 'mr', '
mrs', 'men', 'same', 'man', 'such', 'also', 'youre', 'get', 'just', 'got', 'cant
', 'doesnt', 'thats', 'youve', 'should', 'other', 'most', 'take']
['an', 'a', 'in', 'with', 'since', 'as', 'for', 'the', 'of', 'to', 'and', 'on',
'at', 'been', 'by', 'will', 'after', 'is', 'have', 'it', 'more', 'who', 'you', '
be', 'than', 'said', 'she', 'he', 'his', 'her', 'before', 'so', 'has', 'that', '
their', 'there', 'were', 'where', 'did', 'those', 'went', 'was', 'but', 'we', 'w
ould', 'about', 'put', 'only', 'says', 'now', 'our', 'out', 'them', 'know', 'kee
p', 'ask', 'this', 'from', 'are', 'something', 'really', 'want', 'around', 'didn
t', 'its', 'make', 'do', 'way', 'had', 'not', 'they', 'him', 'which', 'or', 'cou
ld', 'i', 'all', 'one', 'no', 'what', 'when', 'up', 'if', 'into', 'any', 'some',
 'then', 'never', 'your', 'back', 'still', 'very', 'even', 'like', 'again', 'ove
r', 'can', 'me', 'made', 'my', 'how', 'must', 'come', 'us', 'dont', 'go', 'see',
 'till', 'well', 'let', 'too', 'here', 'himself', 'may', 'say', 'these', 'mr', '
mrs', 'men', 'same', 'man', 'such', 'also', 'youre', 'get', 'just', 'got', 'cant
', 'doesnt', 'thats', 'youve', 'should', 'other', 'most', 'take']
****************************************
Finalized Words:
['best', 'moment']
****************************************
best
moment
['best', 'moment']
2
0
0
0
0
It is a stuttering speech
Text: this is the best moment
>>>
```

Fig. 5. Disordered speech output

The above figure corresponds to the stuttered speech output. When the disordered speech is fed into the model which recognizes which type of disorder after comparing it with the inbuilt dataset and displays it along with the corrected speech in the text form on the screen.

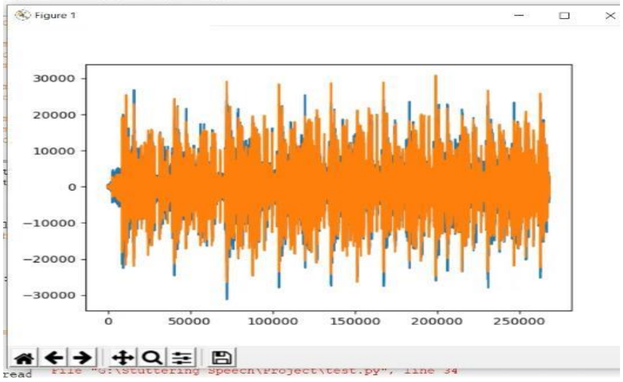
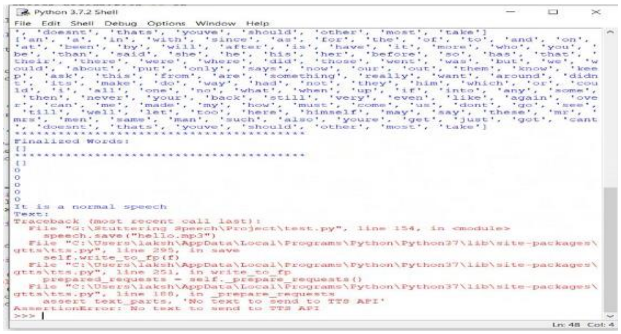


Fig. 6. Waveform of disordered speech signal v/s preprocessed speech signal

The above figure shows the input speech signal and the pre-processed speech signal.



```
', 'doesnt', 'thats', 'youve', 'should', 'other', 'most', 'take']
['an', 'a', 'in', 'with', 'since', 'as', 'for', 'the', 'of', 'to', 'and', 'on',
'at', 'been', 'by', 'will', 'after', 'is', 'have', 'it', 'more', 'who', 'you', '
be', 'than', 'said', 'she', 'he', 'his', 'her', 'before', 'so', 'has', 'that', '
their', 'there', 'were', 'where', 'did', 'those', 'went', 'was', 'but', 'we', 'w
ould', 'about', 'put', 'only', 'says', 'now', 'our', 'out', 'them', 'know', 'kee
p', 'ask', 'this', 'from', 'are', 'something', 'really', 'want', 'around', 'didn
t', 'its', 'make', 'do', 'way', 'had', 'not', 'they', 'him', 'which', 'or', 'cou
ld', 'i', 'all', 'one', 'no', 'what', 'when', 'up', 'if', 'into', 'any', 'some',
 'then', 'never', 'your', 'back', 'still', 'very', 'even', 'like', 'again', 'ove
r', 'can', 'me', 'made', 'my', 'how', 'must', 'come', 'us', 'dont', 'go', 'see',
 'till', 'well', 'let', 'too', 'here', 'himself', 'may', 'say', 'these', 'mr', '
mrs', 'men', 'same', 'man', 'such', 'also', 'youre', 'get', 'just', 'got', 'cant
', 'doesnt', 'thats', 'youve', 'should', 'other', 'most', 'take']
****************************************
Finalized Words:
[]
****************************************
[]
0
0
0
0
0
It is a normal speech
Text:
Traceback (most recent call last):
  File "G:\Stuttering Speech\Project\test.py", line 154, in <module>
    speech.save("hello.mp3")
  File "C:\Users\laksh\AppData\Local\Programs\Python\Python37\lib\site-packages\
gtts\tts.py", line 295, in save
    self.write_to_fp(f)
  File "C:\Users\laksh\AppData\Local\Programs\Python\Python37\lib\site-packages\
gtts\tts.py", line 251, in write_to_fp
    prepared_requests = self._prepare_requests()
  File "C:\Users\laksh\AppData\Local\Programs\Python\Python37\lib\site-packages\
gtts\tts.py", line 188, in _prepare_requests
    assert text_parts, 'No text to send to TTS API'
AssertionError: No text to send to TTS API
>>> 
```

Fig. 7. Output when the normal speech is fed.

The above figure shows the output when the model is fed with normal speech. It recognizes which type of speech signal depending upon the inbuilt data set.

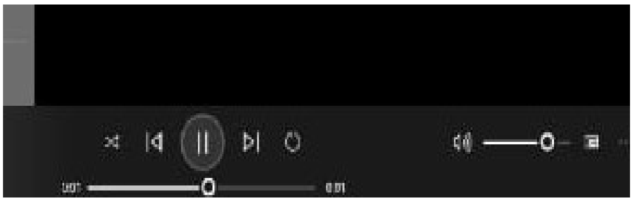


Fig. 8. Output in the audio format.

After comparison and speech correction the output will be in the form of audio which is shown in the above figure.

## VII. CONCLUSION AND FUTURE WORK

The sizeable improvements in the automatic speech recognition system have stimulated the ratification of automated software solutions to recognize speech and language problems in scientific settings. However, this perception continues to be in its early stage because of language obstacles and high-priced web program solutions. Then by, we proposed a semi-automated framework to diagnose the presence of a common language disease, named 'Stuttering'. Aiming to overcome the inefficiencies in the manual diagnosis and to lessen the subjectivity in manual analysis. The effective and efficient detection will result in better know-how of the impact of remedies on illnesses and may be useful to improve the nice of lifestyles of publish-stroke sufferers suffer from aphasia. Researchers and experts are keen on shaping a generalizable software strategy to diagnose speech and language issues regardless of the language regulations.

In future work, we can recognize more language using the automatic speech recognition system.

## ACKNOWLEDGMENT

The authors would like to express their heartfelt gratitude to GSSS Institute of engineering and technology for women and Department of Telecommunication Engineering for their valuable time and support needed through the course of our project work. With immense pleasure we express our sincere gratitude and thanks to our project guide Prof. Manjula G for her valuable support and guidance whenever it was required.

## REFERENCES

[1] Ghulam Muhammad, Khalid AlMalki, Tamer Mesallam, Mohamed Farahat and Mansour Alsulaiman, "Automatic Arabic Digit Speech Recognition and format analysis for voicing Disordered People", 2011 IEEE Symposium on computers and informatics.

[2] Heidi Christensen, Stuart Cunningham, Charles Fox, Phil Green, Thomas Hain, "A comparative study of adaptive, automatic recognition of disordered speech", 2012 INTERSPEECH ISCA's 13th Annual Conference Portland.

[3] Mark S. Hawley, Stuart P. Cunningham, "A Voice Input Voice output communication Aid for people with severe speech impairment", VOL.21, NO.1, January 2013.

[4] Zulfiqar Ali, Irraivan Elamvazuthi, Mansour Alsulaiman, Ghulam Muhammad, "Detection of voice pathology using Fractal dimension in a Multi resolution analysis of normal and disordered speech signals", 2013 King Saud University.

[5] H. Christensen, M. B. Aniol, P. Bell, P. Green, T. Hain1, S. King, P. Swietojanski, "Combining in domain and out of domain speech data for automatic recognition of disordered speech", INTERSPEECH 2013.

[6] H. Christensen, P. Green, T. Hain, “Learning speaker-specific pronunciations of disordered speech”, INTERSPEECH 2013 25-29 August, Lyon, France.

[7] H. Christensen, I. Casanueva1, S. Cunningham, P. Green, T. Hain, “Automatic selection of speakers for improved acoustic modelling: Recognition of disordered speech with sparse data”, 2014 IEEE.

[8] Ying Qin, Tan Lee, Anthony Pak Hin Kong, Sam Po Law, “Towards automatic assessment of aphasia speech using automatic speech recognition techniques”, 2016 IEEE.

[9] Norezmi Jamal, Shahnoor Shanta, Farhanahani Mahmud, and MNAH Sha’abani, “ASR based approach for speech therapy of aphasic patients”, AIP Conference Proceedings 14 September 2017.

[10] Murad Khan, Bhagya Nathali Silva, Syed Hassan Ahmed, Awais Ahmad, Sadia Din, Houbing Song. “You speak, We detect: Quantitative diagnosis of Anomic and Wernicke’s Aphasia using DSP technique”, 2017 IEEE.

[11] Biswajit Das, Khalid Daoudi, Jiri Klempir, Jan Rusz, “Towards disease-specific speech markers for different diagnosis in Parkinsonism” 2019 IEEE.

[12] Shansong Liu, Xunying Liu, Shoukang Hu, Helen M. Meng, “On the use of Pitch Features for Disordered Speech Recognition” Conference Paper · September 2019.

[13] Oscar Saz, Eduardo Lleida, Antonio Miguel, “Combination of acoustic and lexical Adaptation for disordered speech recognition," INTERSPEECH-2009.

[14] Yunbin Deng, Glen Colby, James Heaton, Rupal Patel, “Disordered speech recognition using acoustic and sEMG signals” Conference Paper · January 2009.

[15] Santiago omar Caballero Morales and Stephen J.Cox, “Modelling Errors in Automatic Speech Recognition for Dysastric Speakers”, January 2009.

[16] Prabu, S., M. Lakshmanan, and V. Noor Mohammed. "A multimodal authentication for biometric recognition system using intelligent hybrid fusion techniques." Journal of medical systems 43.8 (2019): 249.