Effects of Alcohol Intoxication on Human Body and Changes in Speech Production before and after Consumption of Alcohol

Amol R. Gaikawad[1], Ratnadeep R. Deshmukh[2], Swapnil D. Waghmare[3]
[1,2,3] *Department of Computer Science & Information Technology, Dr. Babasaheb Ambedkar Marathwada University, Aurangabad 431004 MS India*

***Abstract*— Speech is the most prevalent mode of human communication, allowing people to express, state, and share their emotions, feelings, ideas, and information with one another. According to the research, drinking alcohol produces alterations in glottal excitation, which can lead to a bad impact on a professional and, in some cases, deadly accidents. Additionally, alcohol intake causes major alterations in their vocal cord coordination, such as making extended noises, needlessly long pauses, or creating sounds, as well as breaking words and sentences in their speech, which can lead to speech disorder and disfluencies.**

***Index Terms:* Speech, Marathi Language, MFCC, LPC, Alcohol Intoxication.**

1.INTRODUCTION

Speech is one of the most popular ways for people to express their thoughts, feelings, and emotions. The lungs, trachea, pharyngeal cavity, oral cavity, and nasal cavity are the key components of the human voice system. Muscles press air from the lungs through the larynx, then the vocal cord vibrates and interrupted the air to produce pitch impulses, which are quasi-periodic pressure waves. Impulse production differs from person to person, gender to gender, and age to age[1].

Speech is made up of a detailed acoustic and linguistic backdrop description that contains a set of data. Speech is used to communicate information or a message in a linguistic format, which can be incomprehensible owing to certain diseases[2].

Alcohol causes poor muscular coordination, which affects speech impulses, resulting in speech dysfluency, noise in words or sentences, repeating phrases, incorrect pronunciation, drowsiness, vomiting, unconsciousness, and memory loss, among other things[3].

Speech analysis for intoxication detection was first used in 1989 in the aftermath of the Exxon Valdez, a US oil tanker that ran aground in Alaska. It was discredited that the captain was influenced by alcohol during the flight / accident, but it was impossible to judge owing to late testing of alcohol in blood. However, the sole proof was a tape of the captain's chat with the communication station, which was sufficient to assess the circumstances. Although it took two years to analyse the captain's speeches before and after consuming alcohol, significant differences were discovered[4].

More than normal drunken people are impacted and lose their mental ability to grasp things, which leads to drunk-driving accidents, sexual assault, murder, and/or self-injury. Although there are numerous changes that may be seen or observed before and after drinking alcohol, we can check for changes in voice pitch, which is particularly sensitive to body changes, and after stressing it, we can easily test and know the level of intoxication[5].

"Alcohol Intoxication" is defined as "abnormally high alcohol consumption" that lowers blood supply to the cerebellum and causes euphoria/excitement, diminished inhibition, erratic behaviour, impaired balance, slow reaction time, and gradual muscle atrophy (ataxia). Drinker repetition, incorrect starting, breaking of words or phrases, and unnecessarily creating sounds in speech may all influence overall alcohol consumption[6].

2. EFFECT OF ALCOHOL ON SPEECH

As we all know, listeners can easily judge someone who is under the influence of alcohol by judging unnecessarily pauses, /s/ sound creation in words or sentences, facial actions, and so on, because intoxicated people are unable to control their speech coordination with cognitive function and motor functions, as well as by judging unnecessarily pauses, /s/ sound creation in words or sentences, and so on[7].

1) Other Short-term effects of Alcohol

The short-term effects of alcohol vary depending on how much alcohol is consumed and an individual's physical state[7].

- Speech that is slurred (slow)
- Drowsiness or a drowsy sensation
- Vomiting
- Diarrhoea
- Upset stomach
- Headaches
- Breathing difficulties
- Vision and hearing are distorted (unable or less able functioning)
- Unable to judgment
- Perception and coordination are impaired
- Anemia (loss of RBC)
- Memory losses

2) Other Long-term effects of alcohol

Several health concerns can arise as a result of using alcohol on a regular basis, including the following[7].

- Unintentional injuries such as automobile accidents, fires, and falls
- Sexual assault and domestic violence are examples of intentional injuries.
- Loss of job or loss of productivity on job
- Broken relationship and family problems
- High blood pressure, stroke, and alcohol poisoning are all risks associated with drinking.
- Diseases of the liver and nerve injury
- Sexual problem and intestinal ulcers

3.OBJECTIVES

Blood tests, urine tests, BRAC tests, and other procedures can be used to determine whether a person is under the influence of alcohol.

These tests must be completed on time since late sample collection may cause the subject to be misjudged as intoxicated; for example, in a BRAC test, the suspect may suck air in instead of being blown out. However, the suspect will not be able to deceive in speech production because alcohol affects the speech production system and its muscles. This can involve things like delayed speech, word repetition, visual and hearing impairment, memory loss, and so on.

Although similar study has been done for languages such as English, Korea, Russia, and others, it has not been done for phonetic languages like as Marathi and Hindi. As a result, we're constructing and developing a database to analyse digital speech signals for alcohol-intoxicated Marathi speakers with little resources.

4. METHODOLOGY

The major goal of this study is to determine whether a person is under the influence of alcohol by recording a speech sample (e.g., a consonant or a vowel a word or a sentence), which in our case has the following sequence:

1 Gathering samples
2 Cleaning (background) noise, if any
3 Extracting features
4 Comparing with recorded dataset & Concluding result

i) Gathering samples

For training purposes, we must take two samples of each, one before (in a sober condition) and one after intoxication. However, in a real-life scenario, we may be required to take an intoxicated sample of a person after misbehaving or accusing illegal activity, and then a sober sample for analysis. In a quiet room, gather samples using PRAAT, an open-source software using a microphone or headphones. This application may be obtained from www.fon.hum.uva.nl/praat and is compatible with a variety of operating systems including Windows, Linux, and Mac-OS. We must examine the standards provided by the Linguistic Data Consortium for Indian Languages (LDC-IL) on their official website http://www.ldcil.org/download/SamplingStandards.pdf. For our purposes, we will use a data rate of 16kHz (16bit) utilising a desktop microphone or in a studio environment[8].

ii) Cleaning Noise

Cleaning the noise is the most important element of speech processing since background noise can make speech difficult to hear. For this reason, we utilise Audacity, which is free and open-source software that can be downloaded from the official website at https://www.audacityteam.org. Audacity is compatible with a variety of operating systems, including Windows, Linux, and Mac OS[8].

iii) Extracting Features

This is the most critical and crucial phase because it necessitates a lot of attention. LPC, MFCC, AMFCC, RAS, DAS, MFCC, Higher lag autocorrelation coefficients, PLP, MF-PLP, BFCC, and RPLP are some of the feature extraction approaches or algorithms for speech[9].

Mel-Frequency Cepstral Coefficient:

The most accurate and resilient algorithm that works on acoustic aspects of speech, abbreviated as MFFC, takes into account human perception sensitivity and frequency for the best recognition. The MFCC processor's block diagram illustrates five steps: Frame-Blocking, Windowing, FFT, Mel-Frequency Wrapping, and Cepstrum[9].

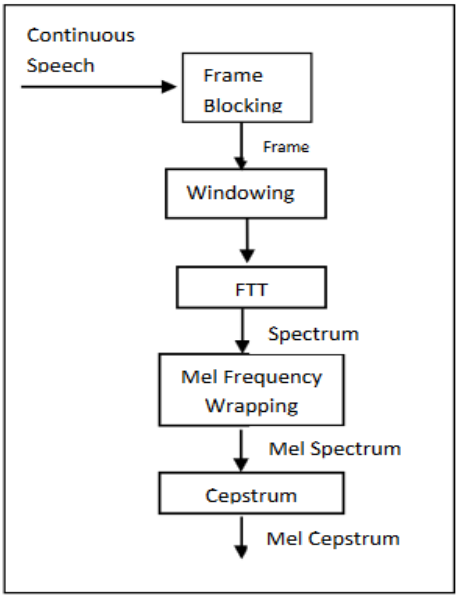

Fig. 1: Block Diagram of MFFC Processor

- Frame Blocking

The recorded continuous speech sample is divided into N-samples, which are separated by M (MN) frames, where N and M are the number of samples in each. N is made up of first samples, whereas M is made up of second samples, with N-M samples overlapping[9].

- Windowing

Each individual frame is windowed to reduce signal discontinuities at the start and end of each frame. The primary goal of windowing is to reduce spectral distortion by tapering the signal to zero at the beginning and end of each frame[9].

- Fast-Fourier Transform

This is how each frame is converted from time-domain to frequency-domain. It is a quick approach for implementing the Discrete Fourier Transform (DFT), which is defined on N Samples xn[9].

Here, j denotes the imaginary unit i.e. $j = \sqrt{-1}$ In

$$Xn = \sum_{k=0}^{N-1} xnke^{-2\pi jkm/N}$$

general Xn's are complex numbers. After this, the result is referred as spectrum or periodogram[9].

- Mel-Frequency Wrapping

We all know that the substance of sound signals' frequency does not follow a linear scale. As a result, subjective pitch is assessed on a scale known as the "Mel" scale for each tone with an actual frequency 'f' measured in Hz. Below 1000Hz, this frequency scale has linear frequency spacing, and above 1000Hz, it has logarithmic frequency spacing. For comparison, 1000mels is the pitch of a 1kHz tone that is 40dB above the perceptual hearing threshold. As a result, we can determine the mels for a given frequency 'f' in Hz using the formula below[9].

$$mel(f) = 2595*log10(1+ f / 700)$$

- Cepstrum

The log mel spectrum is converted back to time in the final step of the process. The mel frequency cepstrum coefficient is a term used to describe the result (MFCC). For a given frame analysis, this provides a decent depiction of local spectral properties. Because the mel-spectrum coefficients are real numbers, we may use the Discrete Cosine Transform to transfer them to the temporal domain (DCT). As a result, we

$$Cn = \sum_{k-1}^{k} (\log Sk)\cos[n(k-0.5)\frac{\pi}{k}]$$

where $n = 1,2,3..k$

can calculate the MFCC's below by denoting the mel power spectrum coefficients that are the outcome of the previous step as k, k = 1, 2,... K[9].

iv) Comparing with Recorded Dataset

We must now compare the freshly recorded and cleaned data to our existing dataset (Speech Corpus). Our dataset includes both a sober and drunk corpus, i.e. samples gathered in both states. This will provide data and, if necessary, graphical representations of Energy Level, Pitch Level, Slurred audio/audio, Noisy speech/audio, and other factors that can assist us in distinguishing between sober and drunk speech.

5. FUTURE WORK

Finding the people who use alcohol is the most difficult task for our investigation. We are considering 10-15 people who drink on a regular or occasional basis, and we need to collect speech samples in both sober and intoxicated states, then remove background noise (if any) to get the words or sentences of both samples, extract the features and compare them for individual subjects or persons, and get the differences for the result to predict or judge whether the person is intoxicated.

6. CONCLUSION

While alcohol can bring people closer together, it also has a number of drawbacks, including sexual assault, domestic violence, drunk driving, and intentional or unintentional injury to oneself or others. Though, by analysing a suspect's speech before and after being under the influence of alcohol, we can detect variations in the gathered samples, which may allow us to judge in a variety of circumstances to identify a suspect of a crime.

7. ACKNOWLEDGEMENT

The University Grants Commission is funding this research through the programme Major Research Project: "Development of Database and Automatic Recognition System for Continuous Marathi Spoken Language for Agriculture Purpose in Marathwada Region." The authors would like to express their gratitude to the Department of Computer Science and IT at Dr. Babasaheb Ambedkar Marathwada University in Aurangabad for providing the research facility. The authors would also like to express their gratitude to the Dr. Babasaheb Ambedkar Research and Training Institute, Pune (BARTI) for their financial assistance.

REFERENCES

[1] Swapnil D. Waghmare1, Ratnadeep R. Deshmukh, Pukhraj P. Shrishrimal, Vishal B. Waghmare, Ganesh B. Janvale and Babasaheb Sonawane, "A Comparative Study of Recognition Technique Used for Development of Automatic Stuttered Speech Dysfluency Recognition System", Indian Journal of Science and Technology, Vol 10(21), DOI: 10.17485/ijst/2017/v10i21/106092, 2017

[2] Amandeep Singh Gill, "A Review on Feature Extraction Techniques for Speech Processing", International Journal Of Engineering And Computer Science ISSN: 2319-7242 Volume 5 Issue 10 Oct. 2016, Page No. 18552-18556

[3] Eszter Tisljár-Szabó, Renáta Rossu, Veronika Varga, Csaba Pléh, "The Effect of Alcohol on Speech Production", (2014) 43:737–748, DOI 10.1007/s10936-013-9278-y, Dec-2014

[4] Milan Sigmund, Petr Zelinka, "Analysis Of Voiced Speech Excitation Due To Alcohol Intoxication", ISSN 1392 – 124x Information Technology And Control, 2011, Vol.40, No.2

[5] Florian Schiel and Christian Heinrich, "Disfluencies in the speech of intoxicated speakers", Journal of Speech, Language and the Law, ijsll (print) issn 1748-8885, ijsll (online) issn 1748-8893 (2015)

[6] Cited on page: "https://www. drugfreeworld. org/drugfacts/alcohol/short-term-long-term-effects.html"

[7] Wiqas Ghai, Navdeep Singh, "Literature Review on Automatic Speech Recognition", International

Journal of Computer Applications (0975 – 8887) Volume 41–no.8, March 2012

[8] Seong-Geon Bae, Won-Hee Lee,Myung-Jin Bae, “A Judgment of Intoxication using Hybrid Analysis with Pitch Contour Compare (HAPCC) in Speech Signal Processing”, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 10 (2017)

[9] Risha Mal, R.K. Sharma, Naveen Kumar, "Intoxicated Speech Detection using MFCC Feature Extraction and Vector Quantization", International Journal of Electronic and Electrical Engineering. ISSN 0974-2174, Volume 7, Number 3(2014), pp.269-280.