The Effect of Diverse Recording Devices on Forensic Speaker Apperception System

Margi D. Vasan
Ph.D. Research Scholar
Institute of R&D, Raksha Shakti University, Ahmedabad, Gujarat, India

Mrs. Surbhi Mathur Choudhary
Assistant Professor Jr.
Gujarat Forensic Sciences University, Gandhinagar, Gujarat, India

Dr. M.S. Dahiya
Director
Gujarat Forensic Sciences University, Gandhinagar, Gujarat, India

Abstract

In the contemporary consequence, for proclaiming the crimes in forensic discipline voice of a person has become a paramount and utilizable evidence, but in our judicial system its consequentially has always been a matter of controversy because of the circumscriptions imposed by sundry factors in the process of identification like background clamour, channel variations, distortions etc. Forensic speech samples customarily differ due to the variation among the acoustic parameters in their recording mode and conditions, affecting the findings. In existing study, to see the nature and amount of variations in acoustic/ perceptual parameters due to transmute of recording media the voice recording of speakers were taken under different recording conditions. The perceptual cues are studied along with their utility in different conditions that the human subjects use to perceive differences in voices. In this study we solely rely in terms of frequency, pitch and energy on recorders concrete variations, to facilitate the process of recognition and will differentiate such features from speaker concrete acoustic parameters.

Keywords: Forensic Science, Forensic Voice, Acoustic/perceptual parameters, Voice

I. INTRODUCTION

In the current scenario, voice of a person has become a paramount and subsidiary evidence for proclaiming the crime in forensic science. Voice (or vocalization) is the sound engendered by humans and other vertebrates utilizing the lungs and the vocal folds in the larynx, or voice box. It is the result of an intricate interplay of physical and emotional events by brain for recognition of a stimulus to communicate [1, 2]. The reason is, voice is additionally a physiological trait because every person have different vocal tract, but voice recognition is classed as behavioural as it is affected by a person's mood. Biometric voice recognition is discrete and distinct from speech recognition with the latter being concerned with precise understanding of speech content and not identification or recognition of the person verbalizing. This is the rudimental principle behind voice identification that a person's voice is his or her individual characteristics that uniquely identify a person.

Our judicial system typically relies on the scientific evidences to apprehend criminals. The well-liked category of physical evidences within judicial system is: DNA, fingerprints, ballistics, footprint, explosives, narcotics, queried documents etc. Others however are still a subject matter of debate about their scientific importance and reliability for evidentiary use. Voice print evidence is one of such controversial evidence [3, 4].

Speaker recognition [5] is the identification of a person from characteristics of voices (voice biometrics). It is withal called voice recognition. It comprises all those activities which attempt to link a speech sample to its speaker through its acoustic or perceptual properties [6, 7, 8, 9]. Forensic speaker identification is the application of widely different techniques for the purport of solving the quandaries cognate to identifying the author of an incognito recording in sundry scenario [10, 11, 12].

In forensic speaker recognition, the voice examination techniques are predicated on different parameters. These parameters as a whole avail the expert to make an opinion on a concrete case to compose a report. But above all the positive aspects poorly recorded/ noisy samples, physical and emotional variability of speaker, analog to digital conversion [13], voice disguise, channel mismatch, mismatch in recording conditions and its position, device mismatch and technology used are certain factors which impose great limitation in the process of identification [6, 14]. Due to the challenges faced by experts while analyzing the voice evidence; in the court most of them feel reluctant to admit such evidence in legal proceedings.

In forensic scenario, generally as per the standards it is expected that the questioned and control samples are recorded in the same recording device, if the device is recuperated from the scene as a law described that "like shall be compared with like" but in forensic cases generally in India, it is personally encountered that same recording devices are not found for recording control and question so in that case there is observed variation in the speech sample causing a drawback in speaker recognition.

In the process of recording, storing and playing back the original analog sound waves (in the form of an electronic signal), it is inevitable that some signal degradation will occur. This degradation is in the form of distortion and noise. These recordings can also vary in their quality depending on their inception, which can be a police interview room, landline or mobile

phone network, etc [15. 16] as the recording medias have their own specific parameters such as frequency of recording, microphone, range of recording, sound to noise ratio, etc . These distortions could transmute formant shapes and positions in the speech spectrum and, in exceptional cases, could cause the appearance of an erroneous formant or a missing formant. Device mismatch or the analog to digital conversion brings relative performance degradation and adversely affect speaker verification [17, 18, 19].

The aim of this paper is to reflect the importance of choosing a felicitous recording media in criminal investigation, the steps taken by the expert to surmount all the limitations and achieve precise and accurate results. The study was conducted by voice samples of person in different recording media to observe the difference in the transform in parameters and form an opinion about the one which could be used for the recording of specimen sample.

II. METHODOLOGY

A. Recording Environment (Room):

As mentioned above, equipment choice may be driven partly by the characteristics of the obtainable recording environment may be a quiet laboratory/studio versus a noisy, reverberant clinic room. Normally, consistent data may be tricky to accumulate from an acoustically uncontrolled room because of extraneous noise as well as reverberation. When recording research level voice data, the most prevalent choice is a high-quality sound isolation booth with cumbersome hefty construction, double walls, a floating floor, quiet lighting ventilation with painted isolations, both of which effectively cut the absorption and reverberation control.

B. Sample Collection:

We invited subjects from the college community (age ranging in early 20's) to attend the sessions of data collection. The subjects speak English. We have 20 subjects (10 males and 10 females) who completed all two sessions. They form the multisession speaker set. Each candidate recited an introductory phrase about themselves was used in the experiment work.

C. Sample Size and Sampling Method:

Each subject recited the assigned phrase twice involving the same background conditions and 4 devices for recording arranged in a series but at different angles from the articulators of the speaker at the same instance for a minimum duration of 10-15 seconds. There individuals were requested to give 8 voice samples in two types of discourse:

1) sitting near the recorder
2) standing far from recorder

The convenience sampling method was used in both experiments, and informed consent was obtained from all participants; the subjects had no problems with speech or hearing. They did not have any diseases or personal habits that affected these two abilities, such as hoarseness, colds, or smoking. While they were participating in the research, they did not use any drugs that could have influenced their abilities to pay attention and concentrate.

Table – 1
Different Recording Conditions

SAMPLES	INSTRUMENT	AV. DISTANCE	AV. ANGLE
Sample 1	Analog recorder	32 cm	30°
Sample 2	Analog recorder	110 cm	27°
Sample 3	Digital recorder	32 cm	90°
Sample 4	Digital recorder	110 cm	90°
Sample 5	Samsung phone	32 cm	45°
Sample 6	Samsung phone	110 cm	42°
Sample 7	Goldwave	32 cm	30°
Sample 8	Goldwave	110 cm	25°

D. Recording Devices:

For conducting this study, the 4 instruments used are

1) Analog recorder
2) Digital recorder
3) Mobile phone
4) Voice recording software.

Table – 2

Specifications of Recording Device [20, 21]

Specifications	*Analog*	*Digital*	*Samsung*	*Gold wave*
File type	*Analog*	*Digital*	*Digital*	*Digital*
Formant (General)	-	*Windows Media*	*MPEG-4*	*Wave*
Formant (Audio)	-	*WMA*	*AAC*	*PCM*
Max. Overall bit rate	-	*4-98 Kbps*	*285 Kbps*	*1411 Kbps*
OLYMPUS	-	*Binary*	-	-
Channel (s)	*Mono*	*Mono*	*Mono*	*Stereo*
Sampling rate	*500-3500 Hz*	*8-48 KHz*	*48 KHz*	*8-96 KHz*
Bit depth	-	*16 bits*	*8 bits*	*16 bits*
Battery	*DC 3V*	*DC 3V*	*V*	-

E. Preparing Files for Analysis:

48,000 Hz/ 24-bit PCM audio files are not suitable for acoustic analysis, though they should be stored master, preservation copies of the recordings. Such files should be described in simple metadata terms (such as Dublin Core or OLAC) and stored on a dependable optical storage medium, such as CD-R or DVD.

Speech samples were down-sampled to 8 kHz and stored as 16 bit PCM wave files to align with the input speech requirements of both mobile codecs.

First, for the analysis the files were down sampled to

- Sampling rate : 11025 Hz
- Bit rate : 172 Kbps
- Bit depth : 16 bits
- Channel : Mono
- File Format : Wave with the help of Goldwave Software and saved.

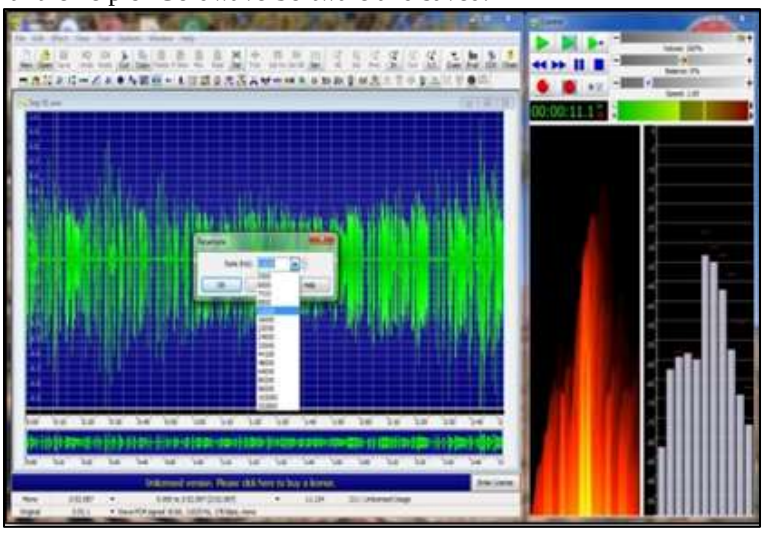


Fig. 1: Resampling in Goldwave

III. ANALYSIS

For lending on accurate and reliable conclusion, it is not advisory to rely only upon the visual comparisons [22]. Hence, we have analysed all the samples collected from the different subjects both by 2 different complementary methods:

1) Auditory analysis
2) Instrumental analysis

A. Auditory Analysis:

This method of speaker recognition by listening is the oldest amongst all. In this situation a person attempts to recognize a voice by its familiarity [23]. We listen to each and every sample repeatedly and carefully the phonetics and linguistic features were the following features are observed [24]:

- Quality of Speech
- Articulation rate
- Flow of speech or rhythm
- Degree of formation of vowels and consonants
- Pauses or stops

- Pronunciation
- Speech time (S/T) rate
- Dynamic Loudness
- Nasality
- Voice Impairment

B. Instrumental Analysis or Spectrographic Method:

Today voice analysis has matured into a sophisticated identification technique, using the latest technology science has to offer. This involves the semi-automatic measurements of particular acoustic parameters such as vowel formants, articulation rate, which is sometimes combined with the results of auditory phonetic analysis [25].

In the instrumental analysis, there are 2 different software's – Goldwave and CSL-4500 (Computerized Speech Lab) been used to approach the spectrographic analysis of the voice samples. The Low-level features refer to the information like fundamental frequency (F0), formant frequency, pitch, intensity, rhythm, tone, spectral magnitude and bandwidths of an individual's voice [26].

Fig. 2: CSL-4500 Software and Hardware

The CSL-4500 software was used to analyse all the voice samples and the following features are to be taken into consideration for their instrumental analysis:

- Spectrograph
- Fundamental and Formant frequency
- Pitch contours
- Energy contours

A sound spectrograph is an instrument which is able to give a permanent record of changing energy-frequency distribution throughout the time of a speech wave (Figure 3) [27]. The spectrograms are the graphic displays of the amplitude as a function of both frequency and time are sometimes called spectral voice prints or voicegrams [28].

Fundamental frequency is the frequency of vibration of vocal cord produced during the rapid opening and closing of vocal cord. The fundamental frequency for speech (f0) is typically 100 to 400 Hz [29].

Formants are spectral peaks of the sound spectrum [29].

Pitch is the variations of the fundamental frequency (pitch) during the duration of the utterance. It is the rate of vibration of vocal folds [6].

Energy is the strength of sensation received through the ear.

1) For analysis procedure:

Open CSL-4500 software and make four window configurations as shown in figure and open the sample file in first window. Mark it as source. This is the file to be analysed.

▼

The second window will have a spectrogram and formants history of all data, the third will contain spectrogram and pitch contour of all data and the final window will take the spectrogram and energy contour of all data.

▼

Finally save the result plates of all the 3 windows and print screen the CSL page in word file as observation.

▼

The analysis performed by CSL-4500 to extract out the features to check the authenticity of the ideal parameters in the voice of the speaker that are individual to the speaker alone.

▼

The results of the analysis made in CSL-4500 were expressed in the way shown in the figure

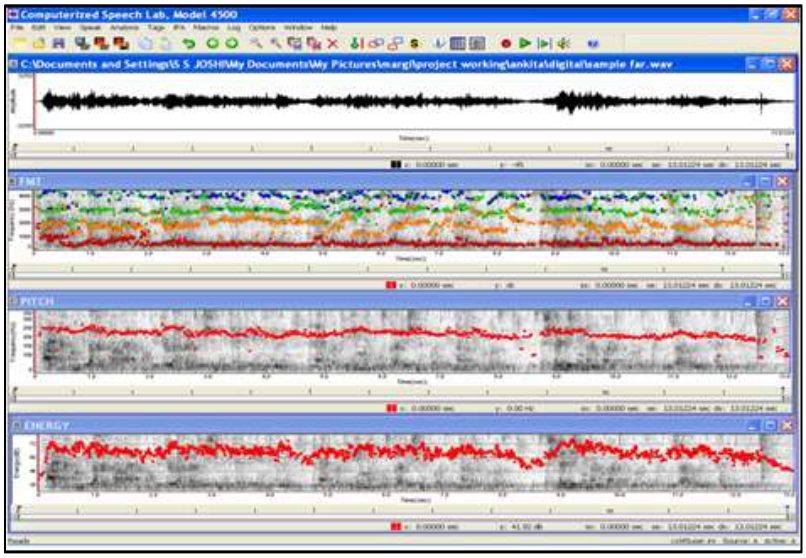


Fig. 3: Analysis of Speech in CSL-4500

IV. OBSERVATIONS

A. Observation Sheet for Auditory Analysis of Voice Samples:

1) Recording Conditions:

Sample 1	*Analog- Sitting, Distance: 33 cm*
Sample 2	*Analog- Standing Distance: 111cm*
Sample 3	*Digital- Sitting, Distance: 30 cm*
Sample 4	*Digital- Standing Distance: 110cm*
Sample 5	*Mobile - Sitting, Distance: 35 cm*
Sample 6	*Mobile- Standing Distance: 113cm*
Sample 7	*Goldwave- Sitting, Distance: 32 cm*
Sample 8	*Goldwave- Standing, Distance: 111cm*

2) The Speaker begins with:

Sample 1	*Hi, I am……..Science University*
Sample 2	*Hi, I am……..Science University*
Sample 3	*Hi, I am……..Science University*
Sample 4	*Hi, I am……..Science University*
Sample 5	*Hi, I am……..Science University*
Sample 6	*Hi, I am……..Science University*
Sample 7	*Hi, I am……..Science University*
Sample 8	*Hi, I am……..Science University*

3) Recording Mode:

Sample 1	*Analog recorder*
Sample 2	*Analog recorder*
Sample 3	*Digital recorder, WMA, 32kbps, 44KHz, Mono*
Sample 4	*Digital recorder, WMA, 32kbps, 44KHz, Mono*
Sample 5	*Samsung phone, 3GA, 96kbps, 48KHz, Mono*
Sample 6	*Samsung phone, 3GA, 96kbps, 48KHz, Mono*
Sample 7	*Gold wave, WAV, 1411kbps, 44 KHz, Stereo*
Sample 8	*Gold wave, WAV, 1411kbps, 44 KHz, Stereo*

4) Quality of Speech Samples:

Sample 1	*Nor mal*
Sample 2	*Nor mal*
Sample 3	*Nor mal*
Sample 4	*Nor mal*
Sample 5	*Nor mal*
Sample 6	*Nor mal*
Sample 7	*Nor mal*
Sample 8	*Nor mal*

5) *Articulatory Speech:*

Acoustic Parameters → / *Sample* ▼	*Speech Duration (sec)*	*Flow of Speech*	*Speech rate (words/ 10 sec)*
Sample 1	*14*	*Normal*	*25 words*
Sample 2	*12*	*Normal*	*26 words*
Sample 3	*14*	*Normal*	*24 words*
Sample 4	*13*	*Normal*	*26 words*
Sample 5	*14*	*Normal*	*25 words*
Sample 6	*13*	*Normal*	*26 words*
Sample 7	*14*	*Normal*	*25 words*
Sample 8	*13*	*Normal*	*27 words*

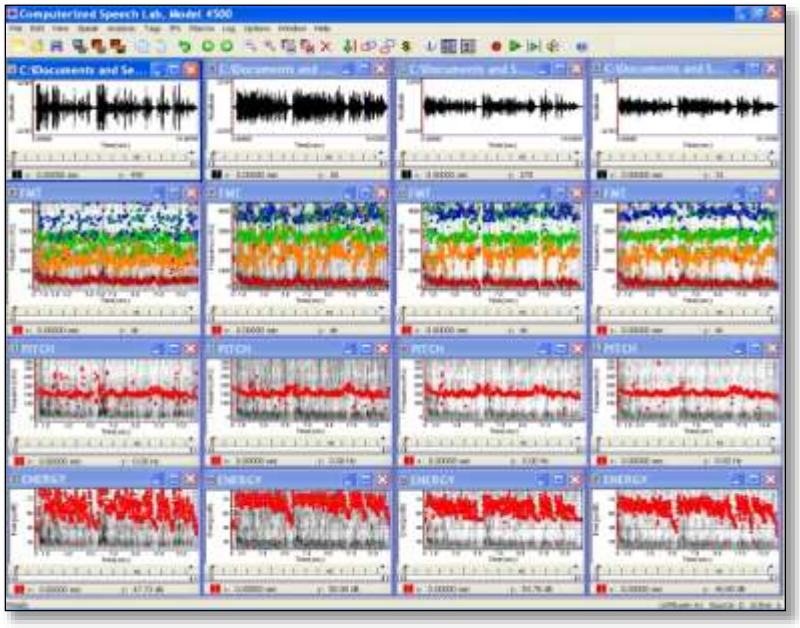
6) *Prosodoic Analysis:*

Acoustic Parameters → / *Sample* ▼	*Intonation Pattern*	*Dynamic Loudness*
Sample 1	*Normal*	*Normal*
Sample 2	*Normal*	*Low*
Sample 3	*Normal*	*Normal*
Sample 4	*Normal*	*Low*
Sample 5	*Normal*	*Normal*
Sample 6	*Normal*	*Low*
Sample 7	*Normal*	*Normal*
Sample 8	*Normal*	*Low*

B. Observation Sheet for Spectrographic Analysis of Voice Samples:

Name:
Age:
Sex:
Educational qualification:
Language

C. Near Sample:



D. Observations:

1) *Formants:*

Formants	*Analog*	*Digital*	*Mobile*	*Goldwave*
F1	*723.43*	*767.37*	*779.42*	*763.69*
F2	*1814.44*	*1632.64*	*1619.13*	*1575.27*
F3	*2517.53*	*2796.57*	*2779.92*	*2740.96*
F4	*3682.26*	*3724.81*	*3529.55*	*3714.96*
F5	*4059.76*	*4454.83*	*4634.16*	*4599.26*

2) *Energy:*

Statistics (dB)	*Analog*	*Digital*	*Mobile*	*Goldwave*
Min. Energy	*43.21*	*47.69*	*53.25*	*41.19*
Max. Energy	*86.09*	*82.02*	*78.78*	*76.45*
Mean Energy	*66.63*	*71.35*	*66.34*	*66.44*
Std. Deviation	*8.80*	*5.33*	*5.62*	*5.61*
Med. Energy	*67.00*	*72.47*	*67.22*	*67.34*

3) *Pitch:*

Statistics (Hz)	*Analog*	*Digital*	*Mobile*	*Goldwave*
Min. Pitch	*71.69*	*73.15*	*75.62*	*73.86*
Max. Pitch	*317.13*	*256.25*	*278.66*	*334.46*
Mean Frequency	*163.81*	*161.69*	*162.35*	*159.41*
Mean F0	*160.04*	*159.00*	*160.63*	*157.32*
Mean Period	*6.25*	*6.29*	*6.23*	*6.36*
Std. Deviation	*27.35*	*18.94*	*16.57*	*17.79*
Median Pitch	*158.93*	*159.73*	*160.42*	*158.34*
Geometric Mean	*161.84*	*160.46*	*161.51*	*158.42*

The similar procedure was followed for the analysis of speech samples taken from far distance as well as voice samples from the other 19 individuals and the observations were noted in the similar manner.

V. RESULTS AND DISCUSSION

On conducting the auditory analysis and simultaneously instrumental analysis by Goldwave and Computerised Speech Lab (CSL-4500) of the intra-speaker as well as the inter-speaker variation it was concealed that each of the 4 recording instrument

1) Analog recorder
2) Digital recorder
3) Samsung phone
4) Goldwave software showed instrument specific variation in features of voice – Formants, pitch and Energy.

The results obtained on analyses of samples collected from all individuals by each recording device are shown in table – 3

Table – 3
Results of the analysed samples

CHARACTERISTIC	*ANALOG RECORDER*	*DIGITAL RECORDER*	*SAMSUNG PHONE*	*GOLDWAVE*
Formants	*The values of F2, F3 and F4 overlap with each other. Hence, they are coming almost in the same range.*	*At the point of pauses, whenever the speaker has a pause in the conversation the first formant F1 is gaining the value and is overlapping with the values of third F3 and fourth formant F4.* *The range of formant frequency is higher in this particular digital recorder indicating high range of recognition.*	*There is clarity in visualization of the formants because each and every formant is going at a particular frequency level. Hence, we can clearly distinguish between all the 4 formants.* *There exists a remarkable artificial upward shift of the centre frequency of the first formant (F1).* *Also the 3rd and 4th formant are coming in almost same range showing less difference.*	*The formant values are not very dense, there is lower amount of overlapping than in other recorders.* *Some F2 formants fall in the range of F4 values.*
Pitch	*Some of the pitch values at 4, 8 and 12 seconds of interval time interval rise to a frequency value upto 250 Hz. These are the places of pauses taken by the speaker.*	*Some of the pitch values show a linear value all over the graph in the range of 100-150 Hz frequency.*	*Some pitch values show decrease at point at 15 second on the graph to a frequency value of 150 HZ.*	*The pitch value falls to the frequency range of 100 dB at the time interval of 10 seconds in every sample*
Energy	*The energy values are seen constantly fluctuating within the range of 40-75 dB.*	*This particular recorder gives the highest values of energy for all the samples taken on it.*	*The energy value of this particular recorder remains constant in a particular range of values about 40-72 Hz.*	*The values of energy are always higher than 60 dB in this particular recorder.*

		The values reach to 75 dB as seen in the entire sample collected.		

On examining the near and far samples of all 20 individuals, the instrument specific variations are same but the energy contour showed high probability of variation. The samples are taken from near distance to the recorder and from far distance; the near sample will be of more energy than of far samples due to distance variation.

From the previous sections, we can observe that aural recognition is robust to changes in recording conditions Such recorder specific features were observed and examined so that to distinguish the voice sample taken at the same time on different recorder.

As we all know, the work procedure of analysis of voice samples and recognition of an individual in CSL and other automatic instruments in a tedious, tiring and time taking which needs an ample amount of time to frame a proper conclusion.

As we were in our study period so there was a shortage of time duration for work completion to form a proper conclusion. Hence, we have taken a limited amount of sample for a short duration and from limited amount of peoples. Therefore, it could be that results may not be as specific as required.

But this study can be extended in future as by collecting ample number of samples from different distance in different condition and by more different newly equipped and sophisticated recording devices having more speaker specific features. This in turn increases the scope of study extracting more speaker specific features on auditory analysis. There are many drawbacks of speaker recognition procedure which affect the features of voice samples of speaker e.g. Change in recording devices, change in recording conditions, the mental and physical state of speaker, environmental conditions, the microphone of speaker, place and position of the recording instrument, etc. For future work, we will also consider more conditions and additional parameters.

For future work we will also change the position of speaker's voice along with the increase in amount of recording devices during collection of speech samples and change in the angle of recording device in a controlled environment.

A. Statistical Evaluation:

Based on above observations it is clear that each recording device used for speech recording has their own characteristic features which can be clearly indicated by the spectrographic analysis of the recording. Some of these features are the key solutions to the problems like differentiating between the recorder specific and speaker specific acoustic parameters. Given below are some recorder specific features marked in voice samples of different speakers, which are statistically evaluated to determine their proficiency in identification of different recorders.

1) Analog Recorder:

Characteristics	*Chances of being analog recording (%)*	*Chances of recording using other devices (%)*
The values of F2, F3 and F4 overlap with each other	*100*	*0*
Some of the pitch values at 4, 8 and 12 seconds of interval time interval rise to a frequency value upto 250 Hz in case of pauses by the speaker.	*100*	*0*

Probability that the recording is done on analog recorder is 100% (average of all observations) Likelihood is 1

2) Digital Recorder:

Characteristics	*Chances of being digital recording (%)*	*Chances of recording using other devices (%)*
At the point of pauses, whenever the speaker has a pause in the conversation the first formant F1 is gaining the value and is overlapping with the values of third F3 and fourth formant F4.	*100*	*0*
Some of the pitch values show a linear value all over the graph in the range of 100-150 Hz frequency	*100*	*0*

Probability that the recording is done on digital recorder is 100% (average of all observations) Likelihood is 1

3) Samsung Phone:

Characteristics	*Chances of being Samsung phone (%)*	*Chances of recording using other devices (%)*
The 3rd and 4th formant are coming in almost same range showing less difference.	*100*	*0*
Some pitch values show decrease at point at 15 second on the graph to a frequency value of 150 HZ.	*100*	*0*

Probability that the recording is done on Samsung phone is 100% (average of all observations) Likelihood is 1

4) Goldwave Software:

Characteristics	*Chances of being goldwave (%)*	*Chances of recording using other devices (%)*
Some F2 formants fall in the range of F4 values.	*100*	*0*
The pitch value falls to the frequency range of 100 dB at the time interval of 10 seconds in every sample	*100*	*0*

Probability that the recording is done on Goldwave is 100% (average of all observations) Likelihood is 1.

VI. CONCLUSION

As we know, the various positions and distance between the talker's mouth and the microphone affect the quality of speech signal, especially in the spectral domain, which is consequential in forensic speaker identification. These variations are due to the reason that different recording medias have their own concrete parameters such as range of frequency of recording, types of microphone, sound to noise ratio, type of batteries, sampling rate, bit rate etc. The transfer characteristics of the communication channel may significantly affect the spectral characteristics of the speech signal; but the question of variation in the speech due to mismatched recording conditions appears as fundamental in forensic applications conducted.

This paper empirically investigates how contrivance mismatch affect the fundamental features of speech. On examining the voice samples collected on the 4 different recorders- Analog recorder, Digital recorder, Samsung phone and Goldwave software that were marked recorder specific variations which discriminate one voice sample with another voice samples recorded in terms of frequency, pitch and energy. The recording conditions have been varied in order to study the differences in the perceptual cues that human beings use to recognize different speakers.

From, the results based on the observations seen in the values of formants, pitch and energy, it was concluded that each recording device consists of its own specific feature which are unique for the device itself.

REFERENCES

[1] Hollien, Harry (2002) Forensic Voice Identification. London: Academic Press Forensics.
[2] Dr. Joseph Picones: Fundamentals Of Speech Recog-nition: A short Course.
[3] Decker JF, Handler J. Voiceprint identification evi-dence- out of the Frye pan and into admissibility. The American university law review www.wcl.american.edu/journal/lawrev/26/decker.pdf
[4] Moenssens AA. Admissibility of Scientific Evidence - An alternative to the Frye rule. William and Mary law review. Volume 25, issue 4.
[5] Kinnunen, Tomi; Li, Haizhou (1 January 2010). "An overview of text-independent speaker recognition: From features to supervectors" (PDF). Speech Communication 52 (1): 12–40. doi:10.1016/j.specom.2009.08.009
[6] SPEAKER RECOGNITION: Joseph P. Campbell, Jr. De-partment of Defense
[7] Ramachandran RP, Farrell KR, Ramachandran R, Mammone RJ (2002) Speaker recognition-general classifier approaches and data fusion methods. Pat-tern recognition 35: 2801-2821.
[8] Homayoon Beigi, ``Speaker Recognition, Biometrics / Book 1, Jucheng Yang (ed.), Intech Open Access Publisher, 2011, pp. 3-2
[9] Ling Feng's: Sepaker Recognition.
[10] Philip rose's (2002) : Forensic Speaker Identification .
[11] Künzel, H.J., "Current Approaches to Forensic Speaker Recognition", Tutorial paper, Proc. of ESCA workshop on Automatic Speaker Recognition, pp. 135-141, Martigny (Switzerland), April 1994.
[12] Meuwly, D., Reconnaissance de Locuteurs en Sciences Forensiques: L'apport dúne Approche Automatique, PhD Thesis, IPSC-Université de Lausanne, 2001.
[13] Advantages and disadvantages of analogue and digital by Matt ottewill
[14] Gfroerer S. Auditory-instrumental forensic speaker recognition. Eurospeech, Geneva, 2003,
[15] Künzel, H. J. (2001) Beware of the 'telephone effect': the influence of telephone transmission on the measurement of formant frequencies.Forensic Linguistics 8(1): 80–99. 4.
[16] Byrne, C. and Foulkes, P. (2004) the 'mobile phone effect' on vowel formants. International Journal of Speech, Language and the Law 11(1): 83-102. http://dx.doi.org/10.1558/sll.2004.11.1.83
[17] H. Meng, et.al, "The Multi-biometric, Multi-device and Multilingual (M3) Corpus," Proc. MMUA, May, 2006
[18] Bin ma et al: Effects of Device Mismatch, Language Mismatch and Environmental Mismatch on Speaker Verification, Acoustics, Speech and Signal Processing, 2007.ICASSP. IEEE International Conference on Speaker Recpgnition (Volume:4).
[19] A.Alexander's: The effect of mismatched recording conditions on human and automatic speaker recogni-tion in forensic applications, Forensic Science Interna-tional Journal 146S (2004), pp. 95-96
[20] Olympus Digital recorder Manual- www.olympus .com
[21] Manual of Panasonic microcassette Recorder
[22] HOLLIEN, Harry.About Forensic Phonetics, Univeristy of Florida, UDK 81'342:343.98, pg-27-53.
[23] Eeva K. Komulainen, "Subjective Voice Identification: The literal meaning of 'Talking Yourself Behind Bars", 26 Alta. L. Rev. 521 (1987-1988)
[24] Forensic Speaker Identification
[25] Michael C. Mc Dermott, "Voice identification: The Aural Spectrographic meth-od",(www.owlinvestigations.com/forensic_articles/aural../fulltext.html)
[26] Graevenitz GA About Speaker Recognition Technolo-gy. Bergdata Biometrics GmbH, Germany.
[27] Samudravijaya K (2003) Speech and speaker recogni-tion: a tutorial. Tata institute of fundamental re-search, Mumbai.
[28] Kvistedal YA (2000) A research paper in forensic sci-ence. the university of Auckland, New Zealand.
[29] Voice Acoustics: an introduction from www.phys.unsw.edu.au/jw/voice.html