# Effect of Disguise on Fundamental Frequency of Voice

**Surbhi Mathur[1]\*, Choudhary SK[2] and Vyas JM[1]**

[1]*Gujarat Forensic Sciences University, Gandhinagar, Gujarat, India*

[2]*Raksha Shakti University, Ahmedabad, Gujarat, India*

**\*Corresponding author:** Surbhi Mathur, Assistant Professor, Gujarat Forensic Sciences University, Gandhinagar, Gujarat-382007, India, Tel: 079-23977171; E-mail: surbhi.mathur@gfsu.ac.in

## Abstract

Although the rapid development of speaker recognition technology is happening, there are still many problems to be solved. The biggest problem arises when the cases of disguised voice samples are come across for the purpose of identification. The samples of disguise are frequently encountered in the forensic scenario like in case of anonymous calls, ransom calls and threatening calls where the speaker makes a deliberate effort to change their voice in order to hide their identity due to the fear of being caught. Voice disguise complicates the process of speaker identification by causing damage to the normal vocal parameters of the speakers; especially fundamental frequency (F0) which is the basic frequency with which the vocal cords of individuals vibrates. The aim of this paper is to study the amount of variation occurring in F0 of disguise and normal speech samples of speakers. This will also aid in determining the validity and reliability of F0 parameter of voice under different disguise conditions.

**Keywords:** Forensic; Disguise; Voice; Speech; Fundamental frequency

## Introduction

The science of crime investigation relies on one basic principle known as the "Principle of exchange". According to this principle when two objects come in contact with one another, there will be transfer of substance between the two. Similarly for the crime site it is believed that no matter where a criminal goes or what a criminal does, by coming into contact with things, a criminal can leave all sorts of evidence, including DNA, fingerprints, footprints, hair, skin cells, blood, bodily fluids, pieces of clothing, fibers and more. At the same time, they will also take something away from the scene with them.

Phil Rose and James R Robertson mentioned that voices of individuals are complex in nature. They aid in providing vital information related to sex, emotional state or age of the speaker. Although evidence from DNA always remains headlines, DNA can't talk. It can't be recorded planning, carrying out or confessing to a crime. It can't be so apparently directly incriminating. Perhaps, it is these features that contribute to interest and importance of forensic speaker identification.

The scenario of speaker recognition reverses in the situation where there is no immediate crime scene like cases involving blackmailing, kidnapping, extortion, threatening, anonymous calls, ransom calls, hoax calls, obscene calls, harassment calls, match fixing etc., where the criminals resort to the aid of telephones and mobiles in order to maintain their anonymity for fear of detection [1-3]. In these circumstances, the voice of an individual is an important clue for identification. This reminds the famous dialogue to a Bollywood comedy "KABIRA SPEAKING", where again telephone plays a vital role in commission of offence.

With the advancement of crime, the criminals are now capable of imparting deliberate change in their voice characteristics to prevent recognition and for misleading the investigation. For example: "A criminal make use of a simple handkerchief over the speaker with the intention to modify his voice". This is the biggest limitation faced by the voice experts all over India.

The voice disguise is considered as a deliberate action taken by the caller to falsify or to conceal his/her identity [4]. This problem is most frequently encountered in Indian forensic scenario, where the suspects always create difficulty for experts by such mischievous activities. Such actions impose potentially serious consequences on personal identification.

Lots of possibilities are offered to a speaker to change his voice and to false a human ear or an automatic system [5]. It is considered as one of the most limiting factor which may seriously affect both expert and lay speaker identification. It is perhaps to be predicted that disguise occurs in certain types of crime more frequently than others. Speakers are more likely to attempt a disguise in situations where the suspect tries to maintain his/her anonymity from the listener who is familiar with their voice [6]. Disguise is most likely in cases such as blackmail, kidnap, extortion or abusive phone call.

Campbell et al. said that forensic applications of speaker recognition should still be taken under a necessary need for caution. Disseminating this message remains one of the most important responsibilities of speaker recognition researchers.

Disguise may take many forms, ranging from simple covering of mouth by cloth, modifying pitch level to mimicry, adoption of a different accent, the use of external objects to affect vocal tract dynamics, modification of the position of the articulators like lips or tongue which affect the formant frequencies or the use of electronic devices.

Voice disguise causes serious damage to speaker's acoustics and phonetic parameters like quality, delivery & flow of speech, degree of phonation, intonation pattern, speech rate, dynamic loudness etc. It also causes variation in the values of fundamental frequency (F0) of

voice, the basic frequency with which the vocal cords of individuals vibrates.

Masthoff conducted a study in which 20 subjects provided samples of their modal as well as disguised voices. The goals were to determine the preferred forms of disguise and possible relations between the modal voice properties and the chosen alterations. It was found that the majority of the disguises were made by an alteration of phonation. Also, the disguises were based on an alteration of a maximum of two phonetic parameters, leaving broad aspects of the vocal behaviour undisguised and thus available for forensic examination.

## Materials and Methods

For conducting the study, the voice samples were collected from 200 subjects of different sex, religion and age groups, mostly of Gujarat origin. Efforts were made to select the collect the voice samples from different parts of Gujarat in order to study the effect of different Gujarati dialects and accents in personal identification. All the voices samples were collected using high quality digital recorder. An accurate transcript was prepared and each speaker was asked to recite the provided transcript four times i.e. one in disguise and three in control conditions. Therefore, a total four voice samples were collected from 200 different subjects, which were then analyzed and compared using voice spectrograph (i.e. CSL-4500), to study the degree of variations in values of F0 among disguise voice and natural voice samples of each speaker. The disguise conditions on which we focused were:

- Keeping hand/cloth on mouth
- Variations in the vocal pitch
- Simulating anger

- Condition of extreme cold
- Condition of bad throat
- Chewing pan or tobacco
- Constriction of vocal tract
- Pinching nostrils
- Pulling cheeks
- Changing the accent and talking style
- Mimicry

### Spectrographic approaches adopted for speaker profiling

The spectrographic method for speaker recognition makes use of an instrument that converts the speech signals into a visual display [7]. In 1941, an electro mechanical acoustic spectrograph was developed by Dr. Raleph Potter, Bell Telephone Laboratory, with an idea to convert sounds into pictures. It is an instrument which is able to give a permanent record of changing energy-frequency distribution throughout the time of a speech wave [8]. Much like fingerprints, voiceprint identification uses the unique features in the spectrographic impressions of people's utterances [9]. They also help the law enforcement in identification of suspicious callers.

Since 1962, it was considered as a fool- proof method of personal identification. Voice identification by spectrographic analysis, the "voiceprint" technique has been in a legal limbo [10]. In this method, a trained examiner may be able to give an opinion about the similarity between the two samples on the basis of voice characteristics like: fundamental frequency; formant frequency, formant patterns (Figure 1), amplitude (Figure 2), energy patterns, pitch patterns, loudness, transitional characteristics, bandwidth.
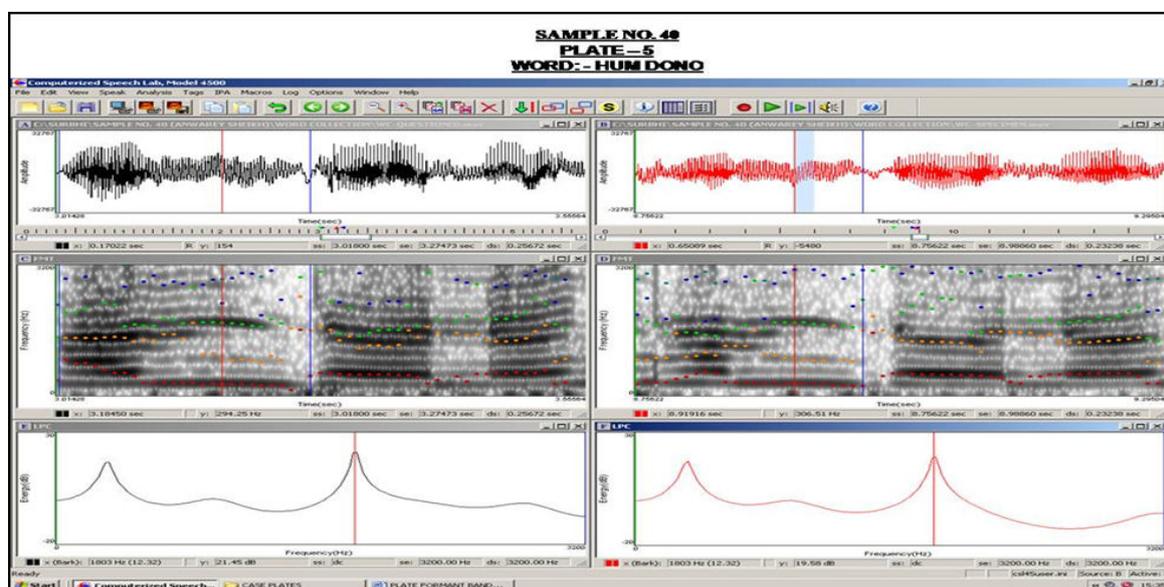


**Figure 1:** Representation of formant bands in frequency versus time graph in CSL software (Courtesy CSL software for spectrographic analysis of voice, DFS, Gandhinagar).
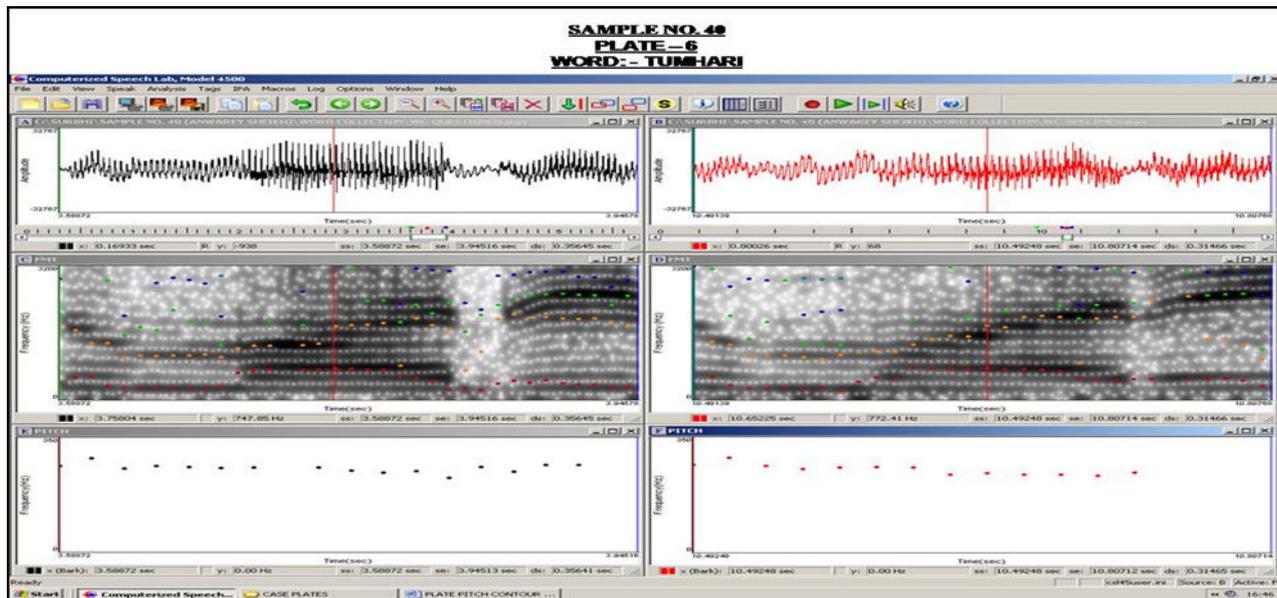
**Figure 2:** Representation of variation in amplitude with respect of time of the sound wave in CSL Software (Courtesy DFS, Gandhinagar).

## Comparison and analysis of voice samples using CSL-4500

- Divide the screen in six windows tiled horizontally.
- Take any word collected from disguised speech of any individual on one side and same word from the control sample of same individual on another side.

- Compare the spectrographs and format patterns of both samples side by side and mark the points of similarities and dissimilarities in the two in terms of frequency, pitch, amplitude and energy.
- The results were recorded in terms of LPC graphs, pitch contour graphs and energy contour graphs (Figures 3-5).
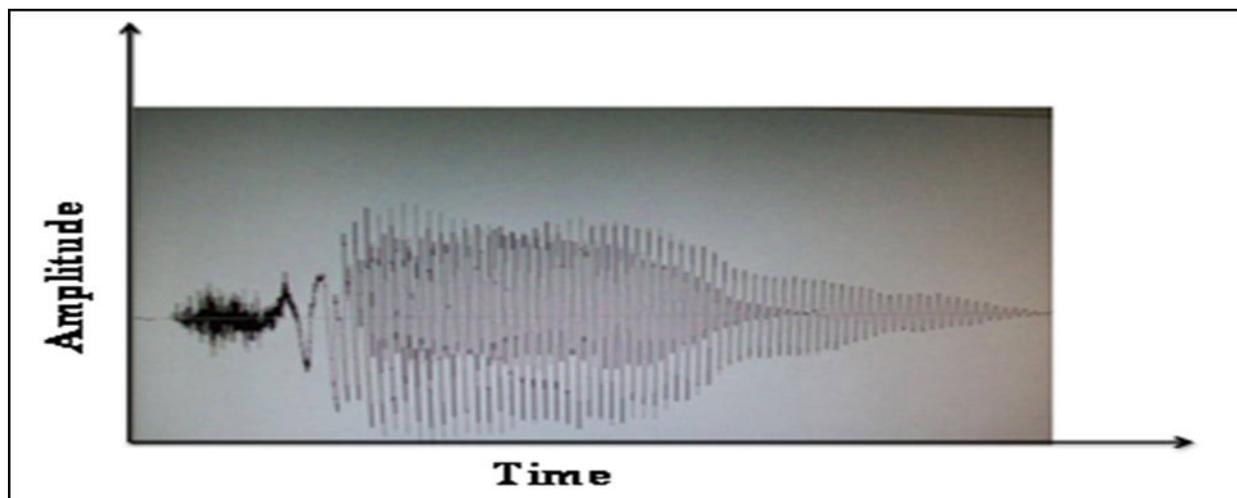


**Figure 3:** LPC graph for comparison of formant frequencies of disguise and control voice samples using CSL software (Courtesy CSL software for spectrographic analysis of voice, DFS, Gandhinagar).
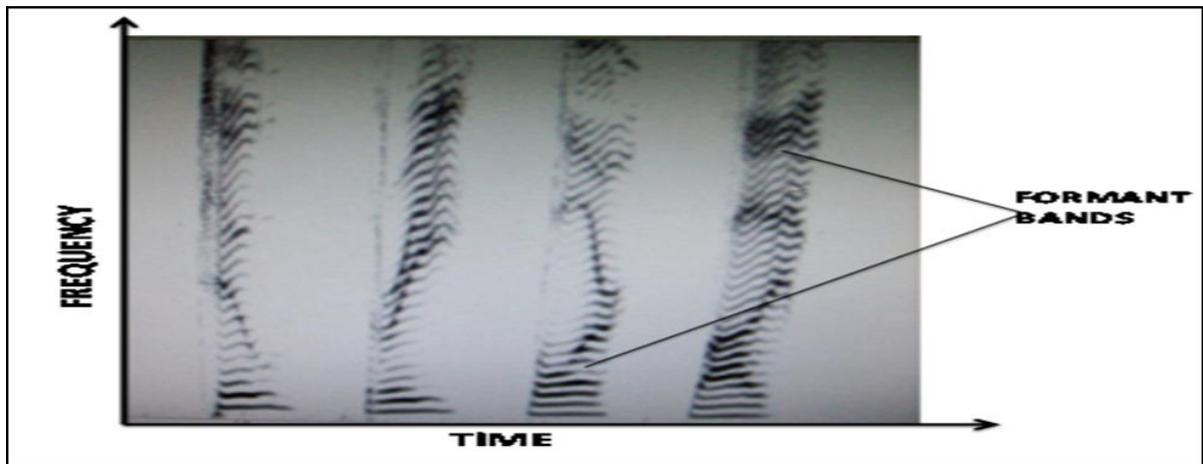
**Figure 4:** Comparison of pitch contour of disguise and control voice samples in terms of frequency versus time graph using CSL Software (Courtesy CSL software for spectrographic analysis of voice, DFS, Gandhinagar).
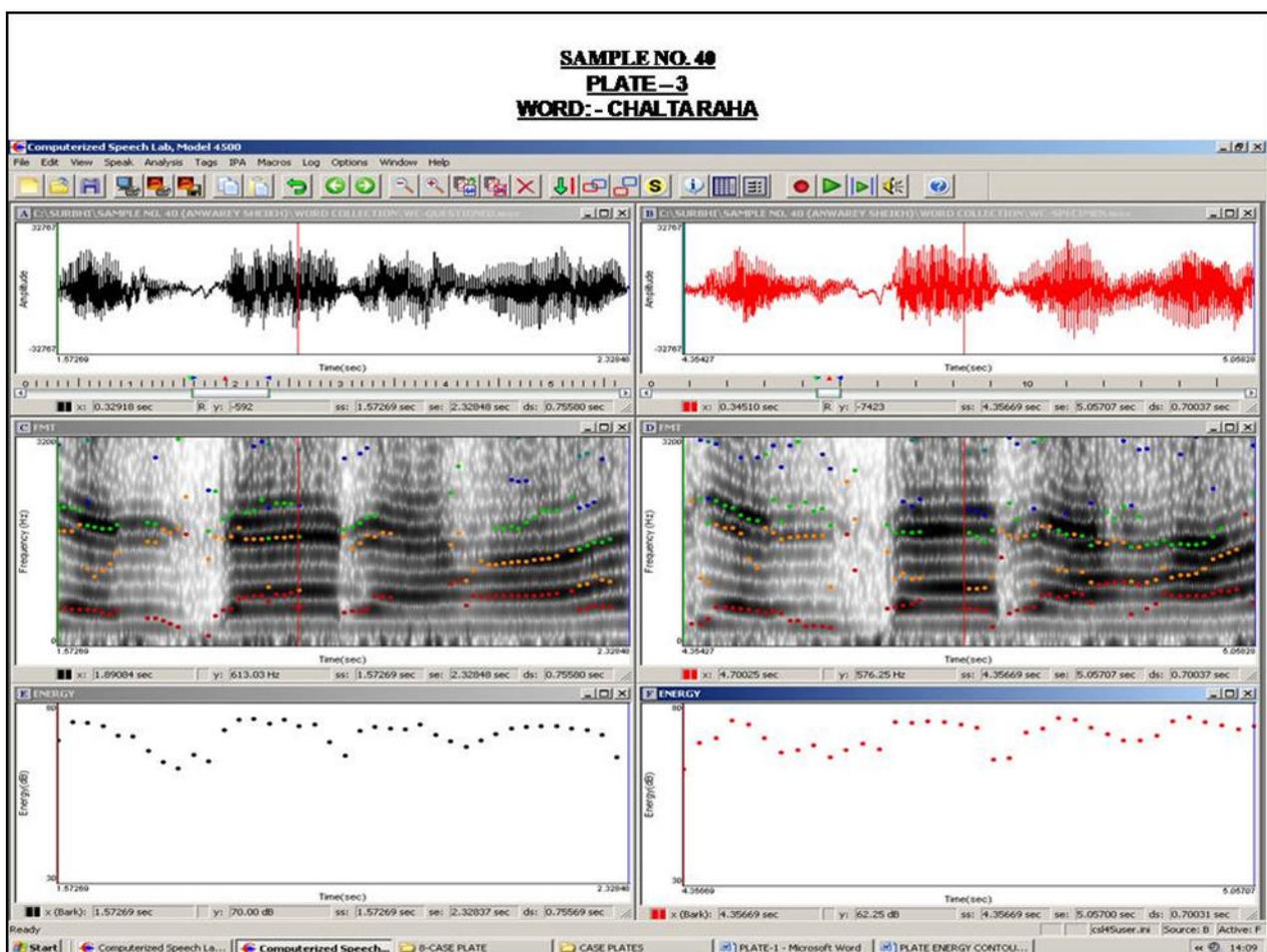


**Figure 5:** Comparison of energy contour of disguise and control voice samples in terms of energy versus time graph using CSL software (Courtesy CSL software for spectrographic analysis of voice, DFS, Gandhinagar).

## Results and Discussions

The subjects were asked to give one of the voice samples by doing some modifications in their original voice. Various disguise techniques were adopted by different subjects including constriction of tract (6%), lowering of pitch (6%), pinching nostrils (9%), pulling cheeks (3%), raising pitch (10%), changing tone/accent (1%), covering mouth (34%), simulating anger (5%), state of cold (2%), mimicry (3%), with some obstacle in mouth (9%), protruding lips (3%), throat infection (3%) and whispering (6%) (Table 1 and Graph 1).

| Method of Disguise | No. of Subject |
|---|---|
| By constricting tract | 12 (6%) |
| By lowering pitch | 12 (6%) |
| By pinching nostrils | 19 (9%) |
| By pulling cheeks | 7 (3%) |
| By raising pitch | 20 (10%) |
| Changing the accent or tone | 2 (1%) |
| Covering the mouth | 68 (34%) |
| Simulating anger | 10 (5%) |
| In state of cold | 5 (2%) |
| Mimicry | 7 (3%) |
| Obstacle in mouth | 17 (9%) |
| Protruding Lips | 5 (3%) |
| State of Throat infection | 5 (3%) |
| Whispering | 11 (6%) |
| Total | 200 |

**Table 1:** Distribution of disguise techniques chosen by the subjects (N=200).
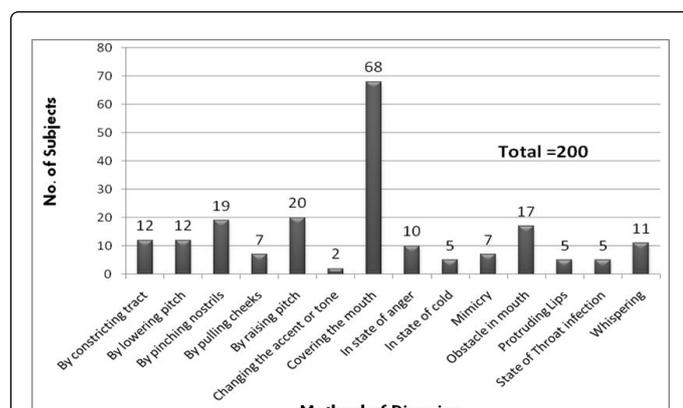


**Graph 1:** Bar chart showing the disguise techniques preferred by the different speakers (Total, N=200).

Majority of subjects have chosen to disguise by covering the mouth externally by using either handkerchief or hand followed by varying

pitch levels, either by increasing or decreasing the normal values and so on. Among the least popular were changing of normal accent/tone, protruding lips, bad throat and condition of cold.

A Positive correlation (+0.514) was observed in the values of F0 in all disguised and control samples indicating 51% chances (i.e. moderate chances) of similarity between disguised and control samples. Z-value of F0 was calculated as -2.38 (p=0.007), indicating significant difference in the F0 values of all disguised and control samples. The mean of F0 values in disguised samples was found to be less than the mean of F0 in control samples.

In case of male subjects, a positive correlation (+0.621) was observed in the values of F0 in disguised and control samples, indicating 62.1% chances of similarity between the two samples [11-22].

Z-value of F0 was calculated as -2.19 (p=0.02), indicating moderate difference in the F0 values of disguised and control samples of male subjects. The mean of F0 values of male subjects, in disguised samples was found to be less than the mean of F0 in their control samples.

In case of female subjects, a positive correlation (+0.421) was observed in the values of F0 in disguised and control samples, indicating 42.1% chances of similarity between the two samples. Z-value of F0 was calculated as -3.03 (p=0.0011), indicating significant difference in the F0 values of disguised and control samples of female subjects. The mean of F0 values of female subjects, in disguised samples was found to be less than the mean of F0 in their control samples.

### Fundamental Frequency (F0) in disguise by constricting tract in males and females

A Positive correlation (+0.811) and (+0.763) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -1.15 (p=0.1056) for male subjects and -0.9 (p=0.1711) for female subjects respectively, indicating strong association in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise by lowering pitch in males and females

A Positive correlation (+0.789) and (+0.561) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -1.36 (p=0.074) for male subjects and -2.98 (p=0.0016) for female subjects respectively, indicating strong association in the F0 values of disguised and control samples of male subjects and strong difference in the F0 values of disguised and control samples of female subjects.

### Fundamental Frequency (F0) in disguise by pinching nostrils in males and females

A Positive correlation (+0.351) and (+0.281) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -3.31 (p=0.0004) for male subjects and -3.01 (p=0.0011) for female subjects respectively, indicating strong difference in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise by pulling cheeks in males and females

A Positive correlation (+0.975) and (+0.727) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -0.61 (p=0.258) for male subjects and -0.108 (p=0.125) for female subjects respectively, indicating strong association in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise by raising pitch in males and females

A Positive correlation (+0.362) and (+0.306) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -2.97 (p=0.0011) for male subjects and -3.12 (p=0.0008) for female subjects respectively, indicating strong difference in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise by changing accent/tone

A Positive correlation (+0.99) was observed in the values of F0 in disguised and control samples of the male subjects. Z-value of F0 was calculated as -0.67 (p=0.227), indicating strong association in the F0 values of disguised and control samples of male subjects.

### Fundamental Frequency (F0) in disguise by covering mouth in males and females

A Positive correlation (+0.792) and (+0.853) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as -1.02 (p=0.147) for male subjects and -1.07 (p=0.146) for female subjects respectively, indicating strong association in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise by simulating anger in males and females

A Positive correlation (+0.908) and (+0.98) was observed in the values of F0 in disguised and control samples of the male and female subjects respectively. Z-value of F0 was calculated as 2.12 (p=0.984) for male subjects and 2.67 (p=0.997) for female subjects respectively, indicating strong association in the F0 values of disguised and control samples of both male and female subjects.

### Fundamental Frequency (F0) in disguise in state of cold

A Positive correlation (+0.789) was observed in the values of F0 in disguised and control samples of the male subjects. Z-value of F0 was calculated as -1.03 (p=0.147), indicating strong association in the F0 values of disguised and control samples of male subjects.

### Fundamental Frequency (F0) in disguise by mimicry

A Positive correlation (+0.521) was observed in the values of F0 in all disguised and control samples. Z-value of F0 was calculated as -2.01 (p=0.020), indicating moderate similarity in the F0 values of disguised and control samples of all males and females.

### Fundamental Frequency (F0) in disguise by obstacle in mouth

A Positive correlation (+0.452) was observed in the values of F0 in all disguised and control samples. Z-value of F0 was calculated as -3.16 (p=0.0006), indicating strong difference in the F0 values of disguised and control samples of all males and females.

### Fundamental Frequency (F0) in disguise in state of throat infection

A Positive correlation (+0.437) was observed in the values of F0 in all disguised and control samples. Z-value of F0 was calculated as -2.87 (p=0.0016), indicating strong difference in the F0 values of disguised and control samples of all males and females.

### Fundamental Frequency (F0) in disguise by whispering

A Positive correlation (+0.563) was observed in the values of F0 in all disguised and control samples. Z-value of F0 was calculated as -3.26 (p=0.0004), indicating strong difference in the F0 values of disguised and control samples of all males and females.

## Conclusions

F0 was found to be more reliable, accurate and consistent parameter for examination and comparison of disguised and normal speech of individuals. The values of F0 in disguise samples including constricting tract, pulling cheeks, changing accent/tone, covering mouth, simulating anger and in state of cold were found to be significantly associated with the values of F0 in their respective control samples, in both males and females.

The values of F0 in disguise by lowering pitch were found to be significantly associated with the values of control samples, only in case of male subjects, while values of F0 in such disguise technique vary more in case of female subjects, when compared to their control samples.

F0 was found to be less reliable in case of disguise including pinching nostrils, raising pitch, obstacle in mouth, throat infection and whispering. The values of F0 in such type of disguise samples were found to be significantly different from the values in their respective control samples, for both male and female subjects. Cases of mimicry shows moderate similarity in values of F0, in comparison to that of their control voice samples.

The values of F0 decreases with constriction of tract, lowering of pitch, pulling cheeks, changing accent/tone, covering mouth, pinching of nostrils, obstacle in mouth, throat infection, whispering and in cold, when compared with their respective control voice samples. Simulation of anger and raising pitch results in increase of F0 with respect to control samples. Mimicry results in more fluctuations in F0 values. The female subjects show more variations in the values of F0 of their disguised samples as compared to male subjects.

## References

1.  Rose P (2002) Forensic speaker identification. London: Taylor and Francis.
2.  Hollien H (2002) Forensic voice identification. San Diego, Calif.: Academic Press.
3.  Kent RD, Read C (1992) The acoustic analysis of speech. San Diego, California.

4. Perrot P, Chollet G (2008) The question of disguised voice. The Journal of the Acoustical Society of America 123: 3878.

5. Zhang C, Tan T (2008) Voice disguise and automatic speaker recognition. Forensic Science international 175: 118-122.

6. Farrus M, Wagner M, Erro D, Hernando J (2010) Automatic Speaker Recognition as a Measurement of Voice Imitation and Conversion. International Journal of Speech language and The Law.

7. Gibbon D, Moore R, Winski R, de Gruyter W (1997) Handbook of standards and resources for spoken language systems.

8. Mathur S, Choudhary S, Vyas J (2014) Speaker Recognition System and its Forensic implications: A review. International Journal of Latest Technology in Engineering, Management and Applied Science.

9. Kersta L (1974) Voiceprint Identification: Police Law Quarterly.

10. Masthoff H (2013) A report on a voice disguise experiment. International Journal of Speech, Language and the Law 3: 160-167.

11. Amin T, German J, Marziliano P (2013) Detecting voice disguise from speech variability: Analysis of three glottal and vocal tract measures. The Journal Of The Acoustical Society Of America 134: 4068.

12. Parul, Dubey RB (2012) Automatic Speaker Recognition System. International Journal of Advanced Computer Research.

13. Kulashreshṭha, Manisha, Mathur, Ramkumar (2012) Dialect accent features for establishing speaker identity.

14. Neustein A, Patil HA (2012) Forensic Speaker Recognition. Law Enforcement and Couter Terrorism.

15. Farrus M, Wagner M, Erro D, Hernando J (2010) Automatic Speaker Recognition as a Measurement of Voice Imitation and Conversion. Internatiopnal Journal of Speech, Language and Law.

16. Kunzel HJ (2000) Effects of voice disguise on speaking fundamental frequency. Forensic Linguistics 7: 149-179.

17. Zetterholm, Elisabeth (2003) Voice Imitation: A Phonetic Study of Perceptual Illusions and acoustic Success (Doctoral Dissertation). Lund University.

18. Beigi, Homayoon (2011) Fundamentals of speaker recognition.

19. Benesty, Jacob, Sondhi, Huang MM, Yiteng (2008) Springer handbook of speech processing.

20. Dellwo V, Ramyead S, Dankovicova J (2009) The influence of voice disguise on temporal characteristics of speech.

21. Campbell JP, Shen W, Campbell WM, Schwartz R, Bonastre JF, Matrouf D (2009) Forensic speaker recognition. Signal Processing Magazine 26:95-103.

22. Masthoff H (2013) A report on a voice disguise experiment. International Journal of Speech, Language and The Law 3: 160-167.