ISSN Print: 2394-7500
ISSN Online: 2394-5869
Impact Factor: 5.2
IJAR 2015; 1(9): 996-1000
www.allresearchjournal.com
Received: 29-06-2015
Accepted: 31-07-2015

Jasdeep kaur
Astt. Prof. in Department of physics, A.S College, Khanna, Punjab, India

Correspondence
Jasdeep kaur
Astt. Prof. in Department of physics, A.S College, Khanna, Punjab, India

Comparison of normal and disguised voice on basis of pitch and fundamental frequency

Jasdeep kaur

Abstract
People generally changed their voices to hide their identities for many purposes, for example to make emergency police calls, for extortion and also for kidnapping. Disguise voice may be defined as an intentional action of a speaker who wants to change his voice for falsifying reason. There is change in values of fundamental frequency and pitch pattern before and after disguise voice. A comparison of each disguise voice with normal voice was taken to put check on speaker identification and verification. The results show that there exists a linear relationship between fundamental frequency and pitch pattern for both normal and disguised voice.

Keywords: Speaker identification, Fundamental frequency, Pitch pattern and disguised voice.

1. Introduction
It would be a great boon to society if speech based interface could be used to communicate with machines in native languages specifically. Majority of population will get the benefits of information technology if contribution of this sector cares for their native languages [1]. For instance, Speech enabled applications helps a lot in public areas such as railway stations, Airports or tourist information centers by assist the customers with answers to their spoken query. Presently Automatic speech recognition (ASR) is widely used for data entry with voice or commands while employee's hands are otherwise occupied. SR technology recommends greater liberty to physically challenged persons in many ways. For example it makes available a natural mechanism to support them such as voice enabled chair.
Speech recognition software is used for deaf people by automatically generated conversations such as in conferences, class room lectures etc. This technology is very useful for people having disabilities or injuries in their hands.
Through a number of experimental studies, [2, 3, 4] it can be said that articulatory and acoustic bases of various consonants are now fairly well understood, though the details of some specific aspects of individual language remain to be clarified. "English" remain the first choice for most of the researchers for the purpose of study and easily available databases. Instead of English, there are also other languages for which speech recognition is under process by different feature extraction.
Speech recognition research work can be found for various Indian languages like Punjabi [5], Kangri [6], Kannada [7], Tamil [8], Malayalam [9], Telugu [10], Bengali [11], Gujarati [12], Marathi [13] and Maithili [14]. For Punjabi, not so much work was done so far. Few researchers [15, 16], [17] have worked on segmentation of speech into smallest units (syllable like). In another paper [18] on Punjabi language, authors studied speech synthesis architecture by make use of Hidden Markov Model (HMM).
What is important for us in our study of speech is that acoustic signal is completely recognizable; we can confine everything the listener perceive sounds in the form of a recording and then measure whichever aspect of the signal that we want to know about. In addition to recognition, even understanding of voice is possible when a recognition system is combined with NLP (Natural language processing systems [19]. Keyboard, mouse and display are main crossing point to exchange a few words with devices that are activated with the assist of software. To convert an acoustic waveform into content equivalent to the information being revealed by the speaker, speech recognition system is used [20].

Many of behaviors felt by human beings can be detected with analysis of person's voice and speech. Even including

(i) Types of emotions (anger, happiness, stress, fear, excitement, depression etc.)
(ii) Effect of external conditions (drug intoxication)
(iii) Intentional behaviors (lie, disguise etc.)
(iv) Health states (cold/flu and fatigue).

Principle of disguise consists in modifying the voice of one person to sound differently or like another person [21]
Following parameters help us to study voice disguise:

- Pitch,
- Formant frequency,
- Voice quality,
- Duration of segment,
- Dialect,
- Articulation,
- Prosody
- Others (Masking on mouth, objects in mouth etc.)

It is the purposeful change of perceived age, gender, identity, personality of any person. This is also possible with help of electronic as well as non-electronic voice changers. Our main focus is to detect type of disguise and as well as original voice if possible. It also involves voice transformation, conversion and even alteration by machines.

2. Hindi Pitch Pattern

In general, Shape of tone and pitch of tone were used to describe Hindi tone. When there are vibrations in vocal cords, it will create reflection in fundamental frequency so from here one can easily delineate tone shape [22]. In order to determine various personality characteristics of individual tones, it is necessary to get detailed decomposition of pitch change.

2.1 Attainment of fundamental frequency and its processing

While complex method used algorithm to extract fundamental frequency and also considering the noise resistance performance and other factors [23]. However by applying either algorithm to get fundamental frequency or by any other method, it is necessary to make use of software (Praat) that can model, adjust the curve data at same time to ensure smooth of curve and label fundamental frequency.
By selecting 10 sampling points to measure fundamental frequency values from extraction.
Fig. 1 depicts fundamental frequency value at a particular time with help of praat software. Some frequency value error may occur near starting and ending position of fundamental frequency band [24]. In absence of sudden flip, value of frequency from spectrogram is correct. But this is not applicable if point is free so one need to pay special attention to this kind of signal.

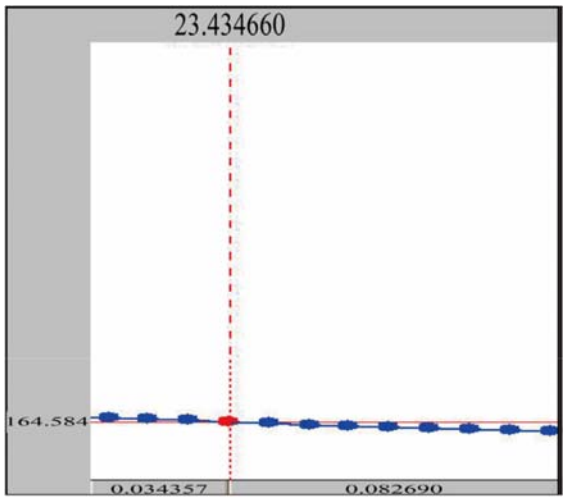

Fig 1: Analysis of fundamental frequency value of PRAAT.

2.2 Pitch pattern curve fitting

From smooth pitch pattern observation and comparison one can measure pitch of speaker and its characteristics [25]. By considering starting and end point of pronunciation stability problem, one need to do suitable correction.

3. Experiment and Analysis

Different corpus from all speakers in their normal and disguise voices were selected. We did analysis of change in frequency value before and after disguise. Speaker identification from different pitch pattern was done with help of comparative analysis of all different pitch patterns.

3.1 Corpus and design

4 speakers with disguised and normal pronunciation of 100 syllables with four tones were taken. There were 8 corpuses that included 400 syllables, numbered as I, II, III, IV for normal and ID, IID, IIID, IVD for disguise voice of each speaker respectively. Recording was done in a professionally recording studio.

3.2 Analysis of results

Pitch pattern for normal voices of four different speakers is shown in fig. 2. These pitch pattern have different curve pattern for instance, (a) and (c) are curved in end as compared to (b) and (d). Furthermore, there is a huge difference in blank area, which is formed by positive tone, rising and falling- rising tone of each speaker (Table 2).

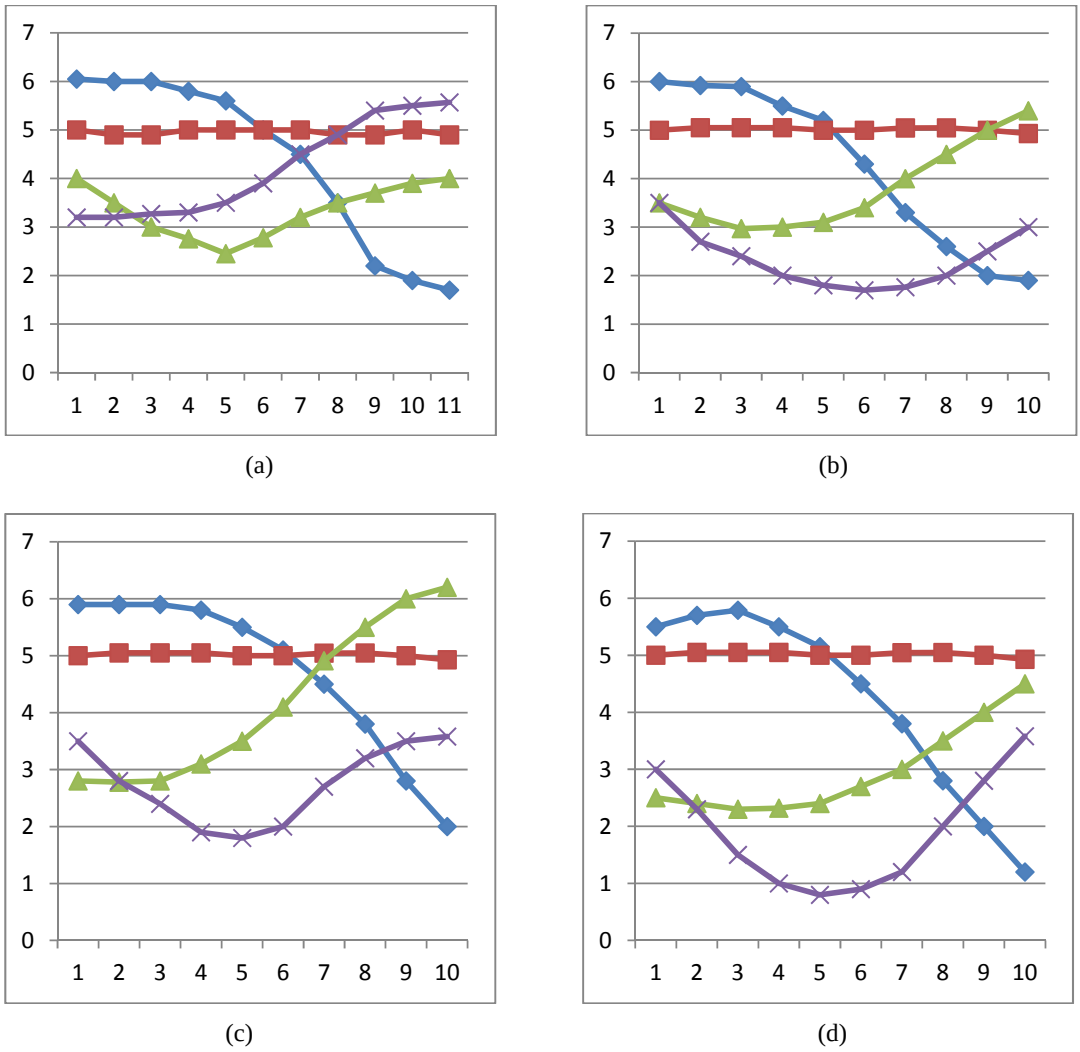

Fig 2: Normal pronunciation representation based on pitch pattern.

3.3. Analysis of fundamental frequency value with normal and disguised voice

From statistical analysis, measured statistical standard deviation is less than 0.05 for change in tone of four speakers before and after voice disguise. The results shows that fundamental frequency of normal and disguised voice has strong correlation also have strong linear ratio between change in fundamental frequency for both normal and disguise voice.

Table 1: Average value and upper tone limit of corpus

Corpus	Level tone with average value	Positive tone of upper limit
(a)	144	5.38
(b)	162.5	5.29
(c)	158.9	6.11
(d)	236.7	4.36

Table 2: Upper and lower limit of falling rising tone

Corpus	Upper limit of falling-rising tone	Lower limit of falling-rising tone
(a)	6.11	3.37
(b)	6.01	2.29
(c)	5.89	1.69
(d)	5.61	0.76

3.4. Normal and disguised pitch pattern of same speaker:

It can be revealed from fig. 3 and 4 that extractions from II and IID attains same values. No doubt, there are very small differences between these two patterns but this can be neglected on basis of overall shape and comparative analysis of each pitch pattern. Pitch pattern remains same before and after camouflage, there may be chances of small differences due to measurement errors. So we can say that speaker identification method is applicable to disguise voice analysis, too.

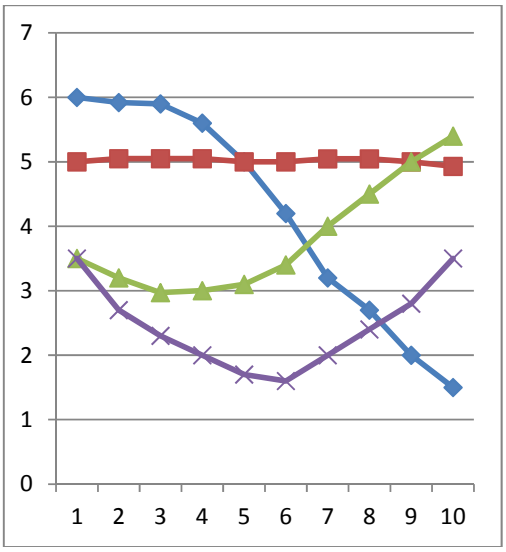

Fig 3: Normal pitch pattern

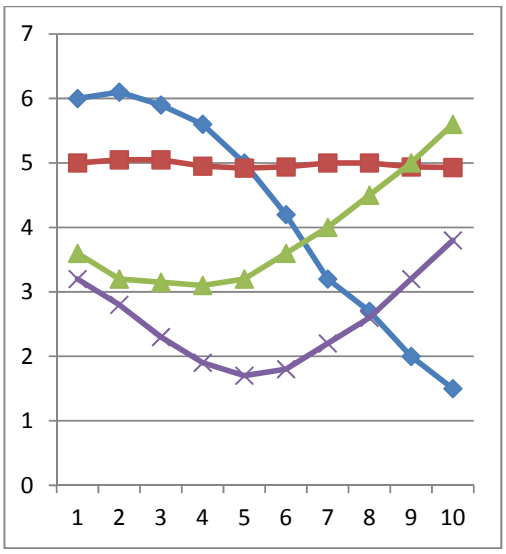

Fig 4: Disguised pitch pattern

4. Conclusions

This paper combines with tone pattern theory by make use of praat and excel to realization of one's pitch pattern. It can be concluded now; Fundamental frequency values for normal and disguise voice have a strong linear relation. Furthermore, Pitch pattern of same speaker in curve shape before and after voice change has a relation. Experiments results proclaimed that, there is a clear-cut difference in pitch pattern between different speakers.

5. References

1. Dua M, Kaur P, Saini P. Hindi automatic speech recognition using Hidden Tool Kit (HTK), International journal of engineering trends and technology (IJETT). 2013; 4: 2233.
2. Diehl RL. The role of phonetics within the study of language, Phonetics, 1991; 48:120-134.
3. Honda M. Human Speech Production Mechanisms, Selected Papers, 2003; 1(2):24-29.
4. Keating PA. Phonetic representations in a generative grammer, Journal of phonetics, 1990; 18:321-324.
5. Choudhary A, Chauhan RS, Gupta G. Automatic speech recognition system for isolated & connected words of Hindi language by HTK, Int J speech comm. 2004; 2(3):35-39.
6. Eaton R. Kangri in context: An aerial perspective, Thesis submitted to University of Texas at Arlington.
7. Hegde S, Achary KK, Shetty S. Statistical analysis of features and classification of alphasyllabary sounds in Kannada language, Int J speech technol. 2014, 67-74.
8. Thangarajan R, Selyam M. Syllable modeling in continuous speech recognition for tamil language, International Journal of speech technology, 2009; 12(1):47-57.
9. Soavithri SR, Jayaram MMM, Rajasudhakar RR, Venugopal MB. A comparative study of base of articulation in Dravidian and indo- Aryan languages, J Acous Soc Ind. 2005, 33.
10. Bhaskar PV, Mohan Rao R, Gopi A. HTK Based Telugu Speech recognition, International journal of advanced research in computer science and software engineering (IJARCSSE). 2012, 2(12).
11. Das B, Mandal S, Mitra P, Basu A. Effect of aging on speech teatures and phonemic recognition: A study of Bengali voicing vowels, International journal of speech technology. 2013; 16:19-31.
12. Vuppala AK, Bhaskararao P. Automatic detection of breathy voiced vowels in Gujarati speech, Automatic detection of breathy voiced vowels in Guajarati speech, International journal of speech technology. 2014; 17:7-82.
13. Nathoosing KC. Isolated word recognition for Marathi language Isolated Word recognition for Marathi language by VQ and HMM techniques, science research reporter, 2012; 2(2):161-165.
14. Yadav R. The influence of aspiration on Vowel duration in Maithili, CNAS, 1980, 7 (1 and 2).
15. Dua M, Aggarwal RK, Kadyan V, Dua S. Punjabi automatic speech recognition using HTK, IJCSI, 2012, 9(4).
16. Kaur A, Singh T. Segmentation of continuous Punjabi speech signal into syllables, in proceedings of the world congress proceedings on engineering and computer science (WCECS 10), 2010, 1.
17. Kumar R. Comparison of HMM and DTW for isolated word recognition of Punjabi language, in proceedings of progress in pattern recognition, Computer vision and applications, Springer, 2010, 6419, 244-252.
18. Bansal D, Goel A, Jindal K. Punjabi speech synthesis system using HTK, International Journal of information sciences and techniques (IJIST), 2012, 2(4).
19. Ladefoged P. Vowel and consonants – An introduction to sounds of languages, Wiley-Blackwell, ISBN 0-631-21412-7, 2001, 1-2.
20. Kumar K, Aggarwal RK. A Hindi speech recognition system for connected words using HTK, Int J Computational systems engineering. 2012, 1.
21. Kuttruff H. Acoustics- an Introduction, Taylor and Francis group, London and New York, ISBN 0-203-97089-6, 2007, 233-236.
22. Cho T, Ladefoged P. Variations and universals in VOT: Evidence from 18 languages, Journal of phonetics. 1999;

27:207-219.

23. Lisker L. Closure duration and the intervocalic voiced – voiceless distinction in English. 1957; 33(1):42-49.
24. Lisker L, Abramson SA. Voice onset time in production and perception of English stops, Linguistic society of America, 1964, 3.1-3.15.
25. Lisker L. Stop duration and voicing in English, in papers in linguistic and phonetics to the memory of Pierre delatrre, Paris, 339-343.