A Study of the Effect of Emotions and Software on Prosodic Features on Spoken Utterances in Urdu Language

Syed Abbas Ali
Computer & Information Systems Engineering, N.E.D University of Engg. & Tech., Karachi, Pakistan.
Email: saaj.research@gmail.com

Maria Andleeb
Department of Computer Engineering, UIT, Karachi, Pakistan.
Email: maria_13augi@hotmail.com

Danish ur Rehman
Computer & Information Systems Engineering, N.E.D University of Engg. & Tech., Karachi, Pakistan.
Email: dani_khan@hotmail.com

Abstract—Speech emotions have potential to provide valuable source of information which can lead us toward human perception and decision making process. This paper analyzes the variation and effect on prosodic features (Formant and Pitch) of female and male speakers in two different emotions (angry and neutral) and softwares (PRAAT and MATLAB) in Urdu language using two ways ANOVA testing. The objective of this paper is to determine the significant effect of emotions and softwares on prosodic features (Pitch and Formant) using recorded speech emotion of both male and female voices of same age group in Urdu language. Experimental results of two-way ANOVA testing considerably show that emotions have effect on pitch and formant both in male and female voice unlike software.

Index Terms—Prosodic Features, Speech Emotion, Urdu Language, Two-way ANOVA Testing.

I. INTRODUCTION AND RELATED WORK

One of the essential tasks of the speech emotion signal analysis is to estimate the prosodic features to convey the psychological state of speaker to others by spoken utterances. Prosodic features are used in daily life conversation as a significant vocal correlates of emotion for recognizing and discriminating emotion from utterances of speakers [1]. Emotions like angry, neutral, sadness, happiness are commonly used to recognize emotions from utterances spoken by speaker [2]. Two prosodic features (Pitch and Format) are considered in this experiment to observe the effect of emotions and software on spoken utterances. Autocorrelation method based comparative analysis of pitch using MATLAB was presented in [3]. Another approach of time domain based algorithm for formant estimation in classifying voice and unvoiced speech signal were presented in [4]. In this

approach, a spoken utterances database is used to test the performance of proposed time domain based algorithm in comparison with PRAAT modified autocorrelation method. Another approach for frequency analysis of Urdu number was proposed in [5] comprises of FFT algorithm in MATLAB to analyze the data. They found the high correlation among frequency contents of the same word, spoken by many different speakers. To alter the attractiveness rating the manipulation of voice pitch has been shown in [6]. This paper concluded that men prefer higher pitched women's voices at very high frequencies, whereas women do not prefer low pitched men's voices at very low frequencies. To estimate the continuous fundamental frequency of speech signals, a different approach was proposed in [7]. The algorithm used in this approach was based on the IFS (frequency estimation) technique. Another idea presented in [8] were based on two MATLAB function namely Generate spectrogram.m and Matrix to Sound.m. It is resolved that the spectra of the sound corresponding to time can be computed using the generate time versus frequency.m matlab files. Likewise an improved method to emotion recognition by pitch is proposed in [9]. The concept of Bayesian classifier is applied and it precisely explained the concept of non-zero pitch. In [10], effect of speech act and tone on rhythm was investigated to compare the mean difference between the PVI (Pair wise Variability Index) values for intensity of different speech acts using One-way ANOVA testing and results concluded that the comparison did not show any significant differences between the two pairs, tone1-tone2, and tone2-tone3. A novel data-driven approach was proposed in [11] to analyze syllable-sized tone-pitch contour similarly in a corpus of Beijing opera (381 arias). The ANOVA statistic modelling and machine learning methods are applied to 30-point pitch contour vectors. Similarly JRPD with a stable stimuli pattern was measured to determine the psychometric functions in [12]. A large standard

deviation 0, 1,2,3,4 of temporal and Gaussian distributed jitter was applied to electric pulse pattern. It is concluded that the pulse rate jitter affects the JRPD and therefore should be reflected in current coding strategies. Similarly Hsin-Yi Lin investigated Pitch and durational cues of emotional speech in Taiwan Mandarin. It involves samples of five males and five females' speakers of Taiwan Mandarin and five emotional types in total i.e. Anger, joy, sorrow, fear and neutral were targeted [13]. It is concluded that there is a correlation between F0 height and speech rate with high arousal dimensions i.e. angry and joy. Negative emotions i.e. anger and sorrow had longer lengthening than positive ones. Experimental results in [14], focuses on speech formant transition. In this investigation speech category boundary was shifted by the presence and direction of the inducer glide. It is concluded that the glide direction of a shorter inducer, of only 100ms, could influence phonetic judgments along a /ba/-/da/ continuum, unlike the longer precursor, 500ms, had no significant influence on the phonetic judgments. Investigation in [15] found that males and females differ in terms of all formants scores that correlate to accents in great details using two-way ANOVA and one-way ANOVA and the plots of the normal fit of individual formants. This study explored the usefulness up to the fifth formant. Synthesis method was evaluated in [16] to investigate the effect of prosodic factors on spectral balance. Spectral balance was measured by using the energy in four broad frequency bands that corresponds to the formant frequency. Speech rate at high arousal dimension was correlated with the height of F0 [17]. Singing voice was correlated with the role of acoustic features using one way ANOVA model [18]. One way ANOVA was performed by considering acoustic cues with affect ratings. Two hypothesis was examined in [19] regarding the modulation of anger and aggressive behavior. Anger and aggressive behavior was counteracted by sadness whereas they are promoted by fear. This paper determines the effect of emotions (Angry and Neutral) and software's (PRAAT [20] and MATLAB) on prosodic features (Formant [22] and Pitch [21]) using recorded speech emotion of both male and female voices in Urdu language. Rest of the paper is organized as follow. The consequent section defines recording specifications and corpus collection of emotion utterances. Research methodology including Software's for prosodic features estimation and Two-way ANOVA testing are discussed in section 3. The experimental results and discussions are presented in section 4. Finally, conclusions are drawn in section 5.

II. Recording Specifications and Corpus Collection

The data evaluated in this study were collected from the university students of age group 18-19 years, and consist of one Urdu language sentence that is suitable to be uttered with two emotions (Angry and Neutral). In this initial work, the samples of emotional speech were randomly taken from daily life personnel and non-actors

with an objective of estimating the recorded speech emotion utterances for real-world implementation. The corpus recording has been performed in Urdu language based on ITU recommendations with following specifications: SNR≥45dB, bit rate of 24120 bps. Microsoft Windows 7 built-in sound recorders has been used to record the entire speech emotion utterances from speakers in Urdu language with following recording format: 48 KHz sampling rate with microphone impendence and sensitivity of 2.4ohm and 56dB ±2.5dB respectively, cable length is 1.8m and stereo type of 3.5mm. Carrier sentence in this experiment was based on well-known desiderata, according to following considerations 1) should be easy to analyze, 2) semantically neutral, 3) dependability on presented situation and 4) should have analogous meaning for each languages. Based on the earlier results [23], the selected sentence was: "**It is not your problem**". The sentence is spoken in two different emotions (Neutral and Angry) in Urdu Language: یہ آپ کامسئلہ نہیں ہے

III. Research Methodology

The presented experimental framework for analyze the effect of emotions and software's on prosodic features is divided into three stages:

1. Analysis methods for Pitch and Formant
2. Software's for parameters Estimation
3. Two-way Analysis of Variance

A. Analysis of Pitch and Formant

The fundamental frequency of excitation source is defined as the pitch. Pitch is loosely related to the log of the frequency. As the frequency gets double, the perceived pitch increases about an octave. The pitch can be analyzed in frequency and time domain. Time domain methods of pitch estimation are ZCR (zero crossing rate), autocorrelation method etc. Frequency domain methods include cepstral analysis, maximum likelihood estimator etc. We use cepstral analysis to estimate the mean pitch value. Formant models the time evolution of speech signal and is considered as the frequency characteristic and also the prosodic feature. Several studies have been attempted to use spectral analysis to extract accent features in accent recognition such as filter bank analysis[24], cepstral analysis such as Mel-frequency Cepstral coefficient[25]etc. In this experiments, linear predictive coding coefficient method and cepstral analysis were used for prosodic parameters estimation.

B. Software's for prosodic features Estimation

Two software's were considered to analyze the effect of emotions and software's (PRAAT and MATLAB) on prosodic features (Pitch and Formant) using recorded speech emotion of both male and female voices in Urdu language. Speech emotion corpus consists of utterances of one Urdu sentence uttered in angry and neutral emotions by male and female speakers. The reason of

using PRAAT is to overcome the instability of speech signal. PRAAT assumes that the speech signal stable by taking into account only the small fragments of it. PRAAT is used to analyze the mean pitch of the recorded samples. In MATLAB, the pitch value is estimated by using the cepstral analysis using following steps:

- The wav files are first read in MATLAB using [x, fs]= wavread ('filename.wav);
- The audio speech signal is plotted
- Windowing of the speech signal is performed
- Fourier transform of the windowed signal is taken using:

$$Y=fft(x);$$

- The spectrogram is plotted
- Cepstrum coefficient is formed which is the DFT spectrum.
- The cepstrum signal is plotted
- Conversion of cepstrum value from s-domain to Hertz.

Linear Prediction coding coefficient method [26] is used to estimate the formant frequency values by taking following steps:

- The .wav files are first read in MATLAB using [x,fs]= wavread ('filename.wav);
- Resampling of audio signal using:

$$x=resample(x, 10000, fs);   (1)$$

- The signal is then passed to the linear predictive filter (LPC) with coefficients:

$$ncoeef=2+fs/1000;   (2)$$

- Copying of coefficient values in the matrix i.e. A:

$$A=Lpc(x, ncoeff);   (3)$$

- Plotting of frequency response:

$$[h, f]=freqz(1,A,512,fs);   (4)$$

- Formant frequencies are then found by root solving:

$$r=roots(A);$$

C. Two-way ANOVA Testing

Two-way Analysis of Variance (ANOVA) testing is used to determine the significant effect of emotions and software on the pitch and formant values. The influence of two different independent variables on one dependent variable is examined by using Two way Analysis of Variance (ANOVA). It not only examined the main effect of each independent variable but also the interaction

between them if exist. The general model of equation of Two way ANOVA is represented in (1):

$$y = \mu + \alpha i + \beta j + (\alpha\beta)ij   (5)$$

Where,

i=1, 2.................n, j=1, 2,.............m
y= Independent factor, consider as (pitch and formant)
μ= Constant value
αi= Main effect of first factor, consider as (emotion_angry, emotion_neutral)
βj = Main effect of second factor, consider as (software_PRAAT, software_MATLAB)
$(\alpha\beta)ij$= Two factor interaction effect.

The following preliminaries are defined in two way ANOVA testing for pitch and formant value:

- Degree of freedom (DF): the number of values in the final calculation of ANOVA that are free to vary is called degree of freedom. It is the number by which a dynamic system can move in independent way without violating any constraint imposed on it.
- Adjusted Sum of Squares (Adj SS): It is the measure of variation from the mean. In ANOVA, the total variation that can be attributed to various factors is expressed by sum of squares.
- Adjusted Mean Squares (Adj MS): dividing the sum of squares value by corresponding degree of freedom yields the mean squares value. It is used to determine whether factors have significant effect or not.
- F-value: The ratio of the two mean squares is called F-value. This value is close to 1 most of the time, if the null hypothesis is true. It determines whether the particular factor is significant or not.
- P-value: Each F-value corresponds to a particular P-value. The cut off point for the alpha level of significance is usually represented by the critical value i.e. 0.05 and it is the p-value associated with certain F-statistics. If P-value is less than 0.05 then factor has significant effect.

The ANOVA table for pitch and formant (P-value and F-value) of both female and male speakers are shown in Table 1., Table 2., Table 3. and Table 4. respectively.

Table 1. ANOVA Table for Female Speakers Pitch.

SOURCE	DF	Adj SS	Adj MS	F- Value	P-Value
Emotions	1	19357	19357	7.87	0.021
Softwares	1	4507	4507	1.83	0.209
ERROR	9	22137	2460		
Lack-of-fit	1	5066	5066	2.37	0.162
Pure Error	8	17071	2134		
TOTAL	11	46001			

Table 2. ANOVA Table for Male Speakers Pitch

SOURCE	DF	Adj SS	Adj MS	F- Value	P-Value
Emotions	1	7475.8	7475.8	5.17	0.049
Softwares	1	1444.0	1444.0	1.00	0.344
ERROR	9	13017.2	1446.4		
Lack-of-fit	1	103.9	103.9	0.06	0.806
Pure Error	8	12913.3	1614.2		
TOTAL	11	21937.0			

Analysis of Variance

Table 3. ANOVA Table for Female Speakers' Formant

SOURCE	DF	Adj SS	Adj MS	F- Value	P-Value
Emotions	1	61456	61456	5.99	0.037
Softwares	1	51418	51418	5.01	0.052
ERROR	9	92413	10268		
Lack-of-fit	1	17898	17898	1.92	0.203
Pure Error	8	74514	9314		
TOTAL	11	205286			

Table 4. ANOVA Table for Male Speakers' Formant

SOURCE	DF	Adj SS	Adj MS	F- Value	P-Value
Emotions	1	38306.2	38306.2	6.16	0.035
Softwares	1	910.8	910.8	0.15	0.711
ERROR	9	55968.2	6218.7		
Lack-of-fit	1	3014.4	3014.4	0.46	0.519
Pure Error	8	52953.7	6619.2		
TOTAL	11	95185.2			

IV. EXPERIMENTAL RESULTS & DISCUSSION

The pitch and formant values of male and female speakers as extracted from PRAAT and MATLAB in neutral and angry emotions are shown in Table.5, Table.6, Table.7 and Table.8 respectively. The regression equations of pitch value for both female and male speakers in neutral and angry emotions are represented in (6) and (7) respectively:

- Pitch = 317.4 + 40.2 Emotions_Angry - 40.2 Emotions Neutral + 19.4 Softwares_Matlab – 19.4 Softwares_Praat (6)
- Pitch = 171.1 + 25.0 Emotions Angry - 25.0 Emotions Neutral + 11.0 Softwares_Matlab – 11.0 Softwares_Praat (7)

Table 5. Mean Pitch Values of Female Speakers in Neutral and Angry Emotions.

		PRAAT (Hz)	MATLAB (Hz)
Speaker	Neutral	329	328.61
	Angry	334	407.5
Speaker	Neutral	272.9	229.44
	Angry	326.4	448.91
Speaker	Neutral	233.37	270.22
	Angry	292.57	336.12

Table 6. Mean Pitch Values of Male Speakers in Neutral and Angry Emotions.

		PRAAT (Hz)	MATLAB (Hz)
Speaker	Neutral	146.4	145.93
	Angry	175.32	188.385
Speaker	Neutral	147.41	119.10
	Angry	184.13	155.65
Speaker	Neutral	120.64	197.58
	Angry	187.10	285.99

Table.5 and Table.6 provide the comparative analysis of pitch with two different emotions: Angry and Neutral when extracted from two different softwares. The mean values of pitch and the variation among them in term of speech emotion samples is presented in above tables. The comparison of the speech emotion samples has been made on the basis of the mean values. Experimental results show following observations:

- It is quite clear that in neutral emotions, the pitch values of both male and females are less than as compared to the pitch values in angry emotions.
- The male speakers pitch values are less than the female speakers pitch values in both emotions.
- The mean pitch values of male speakers have average range of 115Hz-195Hz in neutral emotion and 125Hz-290Hz in angry emotion.
- The mean pitch values of female speakers have average range of 225Hz-350Hz in neutral emotion and 290Hz-490Hz in angry emotion.

Similarly, the regression equations of Formant value for both female and male speakers in neutral and angry emotions are represented in (8) and (9) respectively:

- Formant = 879.9 + 71.6 Emotions Angry - 71.6 Emotions Neutral + 65.5 Softwares_Matlab-65.5 Softwares_Praat (8)
- Formant = 879.9 + 71.6 Emotions Angry - 71.6 Emotions Neutral + 65.5 Softwares_Matlab-65.5 Softwares_Praat (9)

Table 7. Mean Formant Values of Female Speakers in Neutral and Angry Emotions

		PRAAT (Hz)	MATLAB (Hz)
Speaker	**Neutral**	768.24	884.6
	Angry	890.78	1019.5
Speaker	**Neutral**	853.94	883
	Angry	931.05	1184.2
Speaker	**Neutral**	722.39	738.0
	Angry	720.40	963.0

Table 8. Mean Formant Values of Male Speakers in Neutral and Angry Emotions

		PRAAT (Hz)	MATLAB (Hz)
Speaker	**Neutral**	728.23	872.8
	Angry	847.37	930
Speaker	**Neutral**	759.17	638.7
	Angry	850.253	799.4
Speaker	**Neutral**	619.53	742.8
	Angry	843.40	768.8

Table.7 and Table. 8 provide the comparative analysis of first formant with two different emotions: Angry and Neutral when extracted from two different softwares. The mean values of formant and the variation among them in term of speech emotion samples is presented in above tables. The mean values of speech emotion samples are used to compare the samples of speech emotions. The following observation has been obtained from experiments:

- It is quite clear that in neutral emotions, the formant values of both male and females are less than as compared to the formant values in angry emotions.
- The first formant values of female speakers' lies range 800Hz-1200Hz.
- The first formant values of male speakers' lies range 700Hz-1000Hz.

A. Pictorial view of Pitch in PRAAT and MATLAB

To evaluate the experimental results in neutral emotion, Fig.1 and Fig.2 provide the pictorial view of the prosodic feature (Pitch) of the sentence "It is not your problem" in Urdu language using PRAAT. Figures shown that the peak of pitch speech signal of both female and male speakers is in the range defined above. The PRAAT and MATLAB software was used to observe the mean value of pitch in Urdu languages with two different emotions.
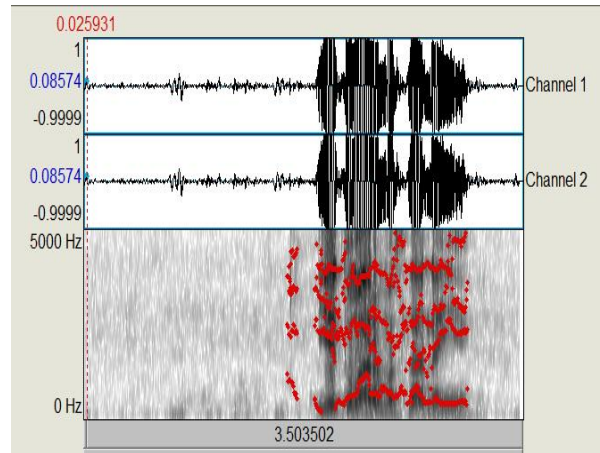

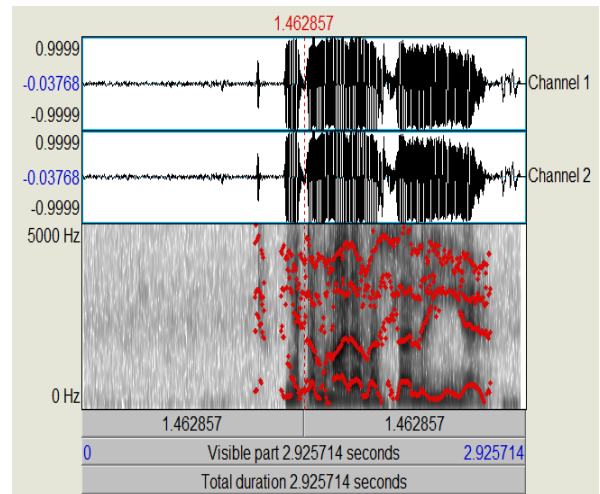Fig.1. PRAAT Speech Signal Waveform of Male Speaker in Neutral Emotion.

Fig.2. PRAAT Speech Signal Waveform of Female Speaker in Neutral Emotion.

Similarly, Fig.3 and Fig.4 shows the MATLAB results of pitch analysis of both male and female speakers' speech samples in the same sentence. Pitch was estimated using the cepstral method.
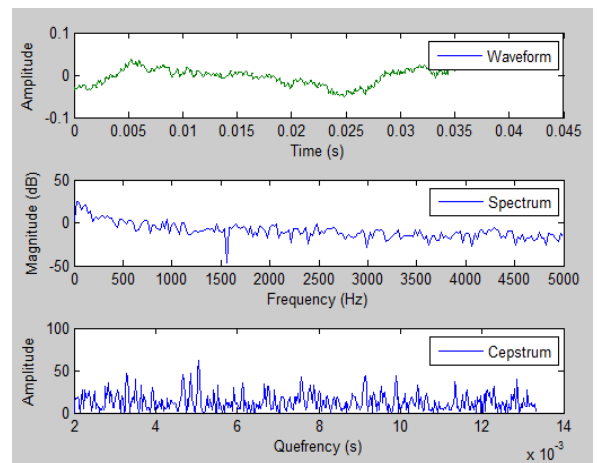

Fig.3. MATLAB Result of Pitch Analysis Waveform Using Cepstral Analysis of Male Speaker in Neutral Tone.
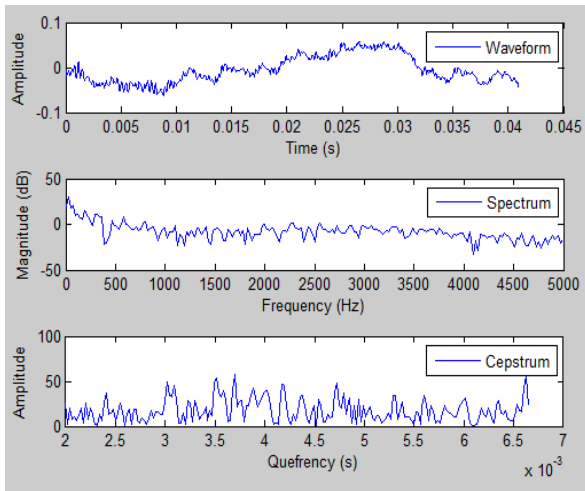
Fig.4. MATLAB Result of Pitch Analysis Waveform using Cepstral Analysis of Female Speaker in Neutral Emotion.

Fig.5 and Fig.6 show the MATLAB results of formant analysis of both male and female speakers. Formant was estimated using the Linear Prediction Coding Coefficient Method.
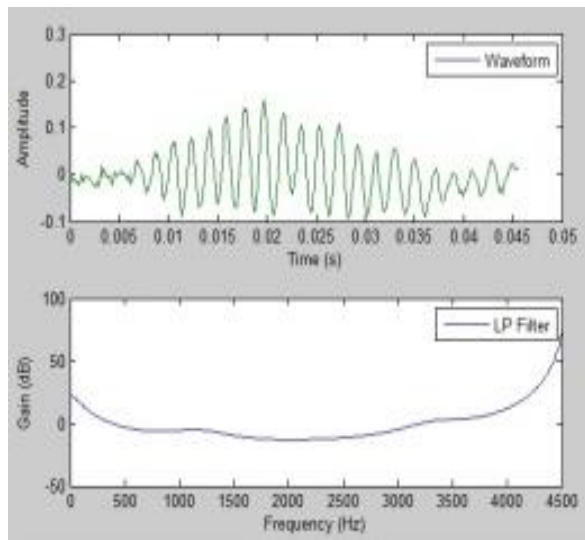

Fig.5. MATLAB Results of Formant Analysis Waveform of Female Speaker in Neutral Emotion using Linear Prediction Coding Method.
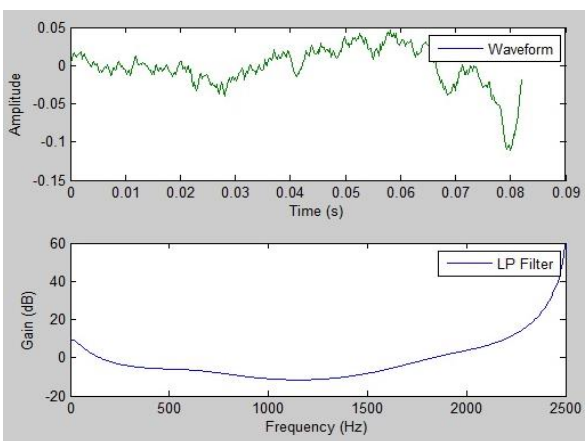

Fig.6. MATLAB Results of Formant Analysis Waveform of Male Speaker in Neutral Emotion using Linear Prediction Coding Method.

ANOVA was carried out with 5% level of significance. Both the male and female pitch shows following observations:

1) Emotions has a significant effect on pitch of human voice ($P<0.05$).
2) Software didn't has any significant effect on pitch of human voice ($P>0.05$).
3) The interaction of emotion and software also has no significant effect on pitch of human voice ($P>0.05$).

B. Main Effect Plot of Pitch and Formant using Two Way Anova:

The main effect plots of emotions and software of both female and male speakers pitch are presented in Fig.7 and Fig.8 respectively:

1) For female speakers the mean pitch value in neutral emotion is 276 Hz and in angry tone is 359 Hz.
2) For male speakers the mean pitch value in neutral tone is 149 Hz and in angry tone is 165 Hz.
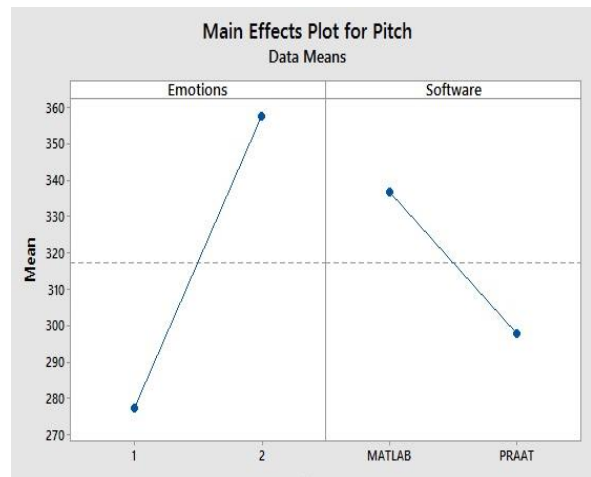

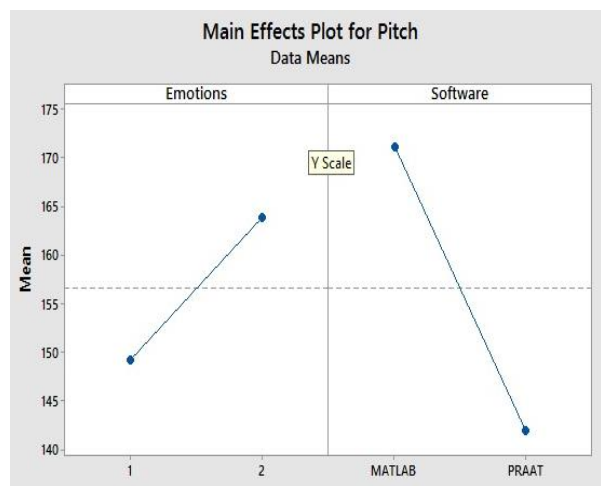Fig.7. Main Effect Plot for Female Pitch Values of Two Way ANOVA.

Fig.8. Main Effect Plot for Male Pitch Values of Two Way ANOVA.

Both the male and female first formant results show that:

1) Emotions has a significant effect on formant of both male and females (P<0.05)
2) Software didn't has any significant effect on formant of male voices (P>0.05) but for the female speech samples software nearly effect the formant as P-value corresponds to 0.052 which is exceeding the significance level only by 0.2%.
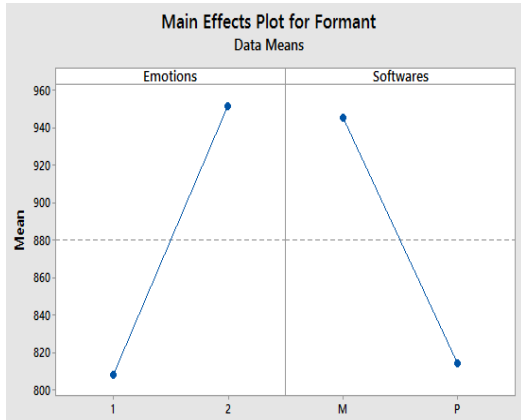

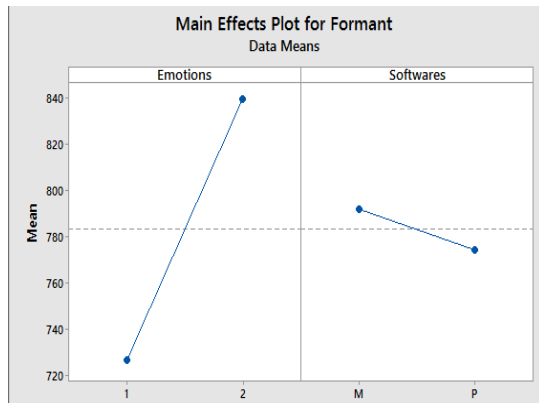Fig.9. Main Effect Plot for Female Formant Values of Two Way ANOVA.

Fig.10. Main Effect Plot for Male Formant Values of Two-Way ANOVA.

The main effect plots of emotions and software of both female and male speakers' first formant are shown in Fig.9 and Fig.10 respectively:

1) For female speakers the mean formant value in neutral emotion is 808.36 Hz and in angry tone is 951.48 Hz.
2) For male speakers the mean formant value in neutral tone is 726.87 Hz and in angry tone is 839.87 Hz.

V. CONCLUSION

This paper presented the two way ANOVA testing to analyze the effect of the two factors i.e. emotions and software's on prosodic features (Pitch and Formant) of male and female speech emotion samples recorded in Urdu language. Demonstrative experiments have been conducted using two different emotions (angry and neutral) and softwares (PRAAT and MATLAB) to analyze the influence of software's and emotions on prosodic features (Formant and Pitch) using recorded speech emotion of both male and female voices in Urdu language. Experimental results provide evidence from Two-way ANOVA table that emotions have significant effect on prosodic features (Pitch and Formant) unlike the software and the interaction of the two factors, whether it is extracted from PRAAT or MATLAB the angry and neutral pitch values have approximately ±50 Hz variation. From future research prospective, authors are focusing that Two-way ANOVA testing can be observed by pursuing more emotions and analyze their impact on other prosodic features.

REFERENCES

[1]   S.R. Karathapalli, S.G. Koolagudi, "Emotion recognition using speech features", Springer Science+ Business Media New York (2013).
[2]   P. Ekman, "An argument for basic emotions", Cognition and Emotion, Vol. 6, pp. 169-200, 1992.
[3]   B.L. Jena, B. P. Panigrahi, "Gender Classification by Pitch Analysis", International Journal on Advanced Computer Theory and Engineering (IJACTE), vol 1, 2012.
[4]   H. Boril, P. Pollak, "Direct Time Domain Fundamental Frequency estimation of Speech in noisy condition" in Proceedings of European Signal Processing conference, pp. 1003-1006, 2004.
[5]   S.K Husnain, Azam Beg, Muhammad Samiullah Awan, "Frequency Analysis of Spoken Urdu Numbers using MATLAB and Simulink", PAF KIET Journal of Engineering & Sciences, vol. 1, p. 5, December 2007.
[6]   D. E. Re, J. J. M. O'Connor, P.J. Bennett, D. R. Feinberg, "Preferences for Very Low and Very High Voice Pitch in Humans", PLOS one, vol.7, March 5, 2012.
[7]   L. Qin, H. Yang and S. N. Koh, "Estimation of Continuous Fundamental Frequency Of Speech Signals" in proceedings of Speech Science and Technology Conference, Perth, 1994.
[8]   V. Mohan, "Analysis and Synthesis of Speech using MATLAB", International Journal of advancements in Research and Technology, vol.2, pp. 373-382, May 2013.
[9]   K.H. Hyun, "Improvement of emotion recognition by Bayesian classifier using non-zero-pitch concept", in proceedings of IEEE International Workshop on Robot and Human Interactive Communication, pp 312-316, 2005.
[10]   C-L-Chaung, "Phonetics of Speech Arts: A pilot Study", in proceedings of 24th conference on computational linguistics and speech processing, September 21-22, 2012, Chung-Li, Taiwan.
[11]   S. Zhang, R. C. Repetto, X. Serra, "Study of the Similarity between linguistic tones and melodic pitch contours in Beijing Opera Singing", in proceedings of 15th International Society for Music Information Retrieval Conference , pp 343-348, 2014.
[12]   A. Bahmer, U. Baumann, "Psychometric function of jittered rate pitch discrimination", Elsevier, pp 47-54, 10th May 2014.
[13]   H. Y. Lin and J. Fon, "Prosodic and acoustic features of emotional speech in Taiwan Mandarin in Proceedings of 6th International Conference on Speech Prosody, pp.

450–453, 2012.

[14] N. Wang and A. J. Oxenham, "Spectral motion contrast as a speech context effect", *Journal of Acoustic Society of America*, Vol. 136, pp 1237-1245, September 2014.

[15] M. A Yusnitaa, M.P. Paulrajb, Sazali Yaacobb, M. Nor Fadzilaha, A.B. Shahriman, "Acoustic Analysis of Formants Across Genders and Ethnical Accents in Malaysian English Using ANOVA", *Elsevier, Procedia Engineering*, Vol. 64, pp 385-394, 2013.

[16] Q. Miao, X. Niu, E. Klabbers, J.V Santen, "Effects of Prosodic Factors on Spectral Balance: Analysis and Synthesis", in speech prosody, Dresden, Germany, 2006.

[17] H.Y. Lin and J.Fon, "Prosodic and acoustic features of emotional speech in Taiwan Mandarin" in Proceedings of 6th International Conference on Speech Prosody, pp. 450–453, Taiwan, 2012.

[18] P. Mouawad , M.D Catherine , A.G Petit, C.Semal, "The role of the singing acoustic cues in the perception of broad affect dimensions", in proceedings. of the 10th International Symposium on Computer Music Multidisciplinary Research, Marseille, France, October 2013.

[19] Zhan J, Ren J, Fan J and Luo J, "Distinctive effects of fear and sadness induction on anger and aggressive behavior" Frontiers in Psychology, vol. 6:725, pp 1-12, 15 June 2015.

[20] http://www.fon.hum.uva.nl/praat/.

[21] A.S. Utane and S.L. Nalbalwar, ―Emotion recognition through Speech‖ International Journal of Applied Information Systems (IJAIS), pp.5-8, 2013

[22] E. Bozkurt, E, Erzin, C. E. Erdem, A. Tanju Erdem, ―Formant Position Based Weighted Spectral Features for Emotion Recognition‖, Science Direct Speech Communication, 2011.

[23] S.A. Ali., S Zehra., et.al. ―Development and Analysis of Speech Emotion Corpus Using Prosodic Features for Cross Linguistic‖, International Journal of Scientific & Engineering Research, Vol. 4, Issue 1, January 2013.

[24] P.P. Vaidyanathan, "MultiRate digital filters, Filter bank, Polyphaser network, and application: A tutorial", *in proceedings of IEEE* vol.78, pp.56-93, Jan 1990.

[25] S. Davis, P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," IEEE Trans. on Acoustics, Speech and Signal Processing, vol. 28, pp. 357-366, Aug 1980.

[26] S. Kwong and K. F. Man, "A Speech Coding Algorithm based on Predictive Coding," In Proc. Data Compression Conference, Hong Kong, pp.455, March 1995.

Authors' Profiles

Syed Abbas Ali received his B.E. and M.Engg. degrees in Computer Systems & Electrical Engineering from the N.E.D University of Engineering & Technology, Pakistan, in 1999 and 2002 respectively. Also, he received Ph.D. degree in Computer Science from N.E.D University of Engineering & Technology, Pakistan, in 2015. Since 2005, he has been working as an Assistant Professor in Department of Computer & Information Systems Engineering, Faculty of N.E.D University. His current research interests include Machine Learning & Speech Recognition, Speech Emotion Recognition and Computational Intelligence.

Maria Andleeb Siddiqui received her B.E degree in Telecommunication Engineering in 2010 from NED University of Engineering and Technology, Karachi, Pakistan. She received M.E degree in Computer and Information system engineering in 2013 from NED University of Engineering and Technology. Currently she is enrolled in PhD in the same University. Her research area is speech emotion recognition. She is lecturer in Usman Institute of Technology.

Danish Ur Rehman Khan received his B.E. degree in Computer Systems Engineering in 2000 from NED University, Karachi, Pakistan. He received his first M.E. degree in Electrical Engineering with majors in Telecommunications (by coursework) in 2007 and second M.E. degree in Computer and Information Systems Engineering with majors in Computer Networks and Performance Evaluation (ISP) in 2012. He is also accepted for PhD in NED in 2014. Current research interests include Human Behavior and Speech Recognition.