

Investigation of Pitch and Duration Range in Speech of Sindhi Adults for Prosody Generation Module

Shahid Ali Mahar¹, Mumtaz Hussain Mahar², Shahid Hussain Danwar³, Javed Ahmed Mahar⁴
Department of Computer Science, Shah Abdul Latif University, Khairpur, Pakistan

Abstract—Prosody refers to structure of sound and rhythm and both are essential parts of speech processing applications. It comprises of tone, stress, intonation and rhythm. Pitch and duration are the core elements of acoustic and that information can make easy to design and development for application module. Through these two peculiarities, the prosody module can be validated. These two factors have been investigated using the sounds of Sindhi adults and presented in this paper. For the experiment and analysis, 245 male and female undergraduate students were selected as speakers belonging from five different districts of upper Sindh and categorized into groups according to their age. Particular sentences were given and recorded individually from the speakers. Afterward, these sentences segmented into words and stored in a database consisting of 1960 sounds. Thus, distance of the frequency in pitch was measured via Standard Deviation (SD). The lowest Mean SD accompanied 0.25Hz and 0.28Hz received from male and female group of district Sukkur. The highest Mean SD has measured with male and female group of district Ghotki along 0.42Hz and 0.49Hz. Generally, the pitch of female's speakers was found high in contrast to male's speaker by 0.072Hz variation.

Keywords—Prosody generation; speech analysis; pitch; duration; Sindhi sounds

I. INTRODUCTION

Sindhi Language is being spoken with various accents across the Sindh. Linguists generally divided this language in six dialects [1]. People speak the language in different accent in the same region because either they have migrated from other region or they are living in the districts adjacent to Punjab and Balochistan provinces.

Phonologically, Sindhi language is rich and has sufficient sound inventory [2]. The complex variation in accent is major cause for the less accuracy in Sindhi speech processing software applications specifically prosody generation module. To reach the maximum accuracy in software applications it is mandatory to measure the fundamental frequency of Sindhi sounds.

While speaking the Sindhi language, the variations in sound duration and pitch are normally observed with routine sounds of words but surprisingly these variations are also observed when the homographic words are spoken with the different diacritic symbols even spoken by the same adult. In Sindhi, huge number of homographic words is available which are commonly used and spoken. For instance سُر and سَر are the homographic words used as singular and plural respectively having difference in pitch frequency of 146.89Hz

and 150.61Hz and in sound duration of 0.721ms and 0.566ms seconds respectively can be seen in “Fig. 1” and “Fig. 2”.

During the literature review it is observed that the pitch and duration for Sindhi sounds have not been digitally analyzed at acceptable level and statistically measurement also have not been done for Sindhi sounds parameters like pitch and duration whereas deep analysis is mandatory for various speech processing application specifically prosody generation. Therefore pitch and duration is statistically measured and presented in this paper.

The development of Prosody Generation Module is the main objective of this research for which the pitch and duration ranges are the prerequisites parameters. In solitude condition of prosody, the unusual effects of prosody are complicated to reproduce and also the analysing of prosody is difficult due to the function multiplicity [3]. The prosody generation module is the mandatory components of various speech processing software applications specifically business-related Text-to-Speech systems today make use of rather unsophisticated systems, characteristically conveying a defaulting sentence accent based on the function word distinction [4].

In this study, various male and female adult inhabitants of five districts: Khairpur (K), Sukkur (S), Ghotki (G), Shikarpur (Sh) and Larkana (L) are chosen for recording the sounds to evaluate the fundamental frequencies particularly pitch and duration through which prosodic information pertaining to recorded sounds can be depicted and analyzed for further processing and development of speech based software applications [5] [6].

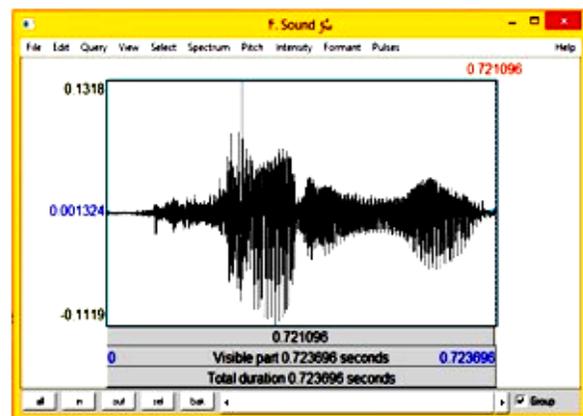


Fig. 1. Sound Information of Word سُر using PRAAT.

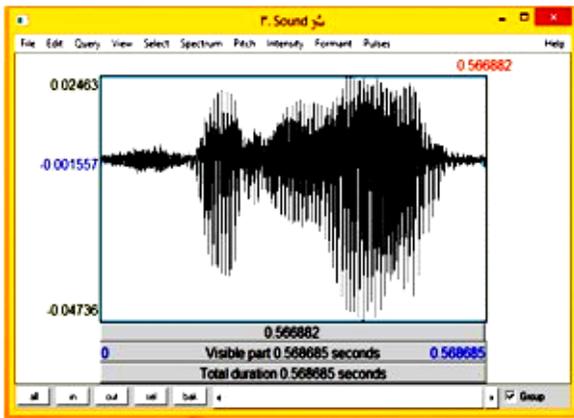


Fig. 2. Sound Information of Word سَئِر using PRAAT.

II. RELATED WORK

The research efforts have been taken for the investigation of Sindhi phonemes. The phonological problem of Sindhi is addressed by Shaikh [7] and the observation reveals that the melody and intonation of Sindhi language in six dialects are different from each other. The accent of peoples of different dialects are compared based on the waveform visualization of image for resolving the detected problems. The Sindhi phonology is also investigated by Mahar [2] [8] along with the letter to sound conversions. The research is specifically focused to demonstrate the f_0 peak of different classes of syllables in which short and long vowels are used at different positions in the words.

The research contributions pertaining to the analysis of fundamental frequency of Sindhi language are published by Abbasi [9] [10]. The role of pitch between intonation and stress is investigated. The pitch accent is analysed on the recorded sounds of 69 words and the digital experiments are performed on the recorded words having different number of syllables. The obtained results proved that stress is entirely orthogonal to F_0 contours.

The acoustic analysis of Sindhi language is presented by Keerio [11] focusing the consonant sounds. The experiments were performed on the collected sounds having the VCV formats. The research is based on the liquid class of consonants and emphasis is given on the difference in the trill and lateral consonants. Moreover, the vocalic variation in vowels is analysed by Mahwish [12] considering differences among the languages spoken in Pakistan. In their research, the phonology of Sindhi is also discussed and variation is found in vowels with reference to the spoken languages of Pakistan using the PRAAT speech analyzer. Furthermore, the basic idea of digital analysis and statistical measurements of the Sindhi sounds is taken from the available literature mentioned above which will help to develop prosody generation module of Sindhi language.

III. RESEARCH METHODOLOGY

The research methodology of this work is mainly based on seven different phases. The first step is the assortment of the speaker because they are needed for recording the sounds. Therefore, 245 undergraduate students of Department of

Computer Science were selected. The students are belongs to five Districts of Upper Sindh. After the selection of appropriate speakers, the sounds were recorded at the radio station, Khairpur.

The next step is the development of speech database so that 65 descriptive sentences were composed and 8 sentences were randomly given to the male and female speakers. The speech corpus is made to collect the prosodic information available in the recorded sounds. For digital investigation of the recorded sound, PRAAT speech analyzer is used to measure the pitch and duration of recorded Sindhi sounds.

The obtained results are classified into male and female duration in ms and pitch in Hz of recorded sounds. The Mean and SD of durations and pitches are also calculated to quantify the amount of variation. After that obtained results are compared and presented according to age groups and districts.

IV. SELECTED SPEAKERS

Generally in experimental research, the speakers are always required for recording of voices to analyze the pitch and duration [13]. In this research, several undergraduate students from the Department of Computer Science Shah Abdul Latif University Khairpur were randomly selected for sounds recording. Our research is particularly focused on the analysis of the sounds of the people living in five districts of upper Sindh and mostly belongs to dialect Siroli. The speakers are classified in terms of gender and age groups.

Table I presents the number of speakers of a particular District with age groups. 61 speakers from district Khairpur, 53 from district Sukkur, 52 from Ghotki, 45 from Shikarpur, and 34 from Larkana are selected respectively. The selection process was entirely based on the availability of the students and the willingness of the speakers.

TABLE. I. SELECTED SPEAKERS WITH AGE GROUPS

Districts	Age Group (Years)	No. of Speakers
Khairpur	19	09
	20	14
	21	22
	22	16
Sukkur	19	11
	20	12
	21	15
	22	15
Ghotki	19	06
	20	09
	21	17
	22	20
Shikarpur	19	06
	20	14
	21	15
	22	10
Larkana	19	13
	20	03
	21	08
	22	10

V. RECORDING PROCEDURE

The sound files are recorded at the radio station, Khairpur with a well-connected system of four components out of which three are hardware and only one is software product. The four components include a microphone, an audio console, a computer system and adobe auditions. The first component; microphone is a high frequency maintaining equipment that has the capacity of receiving sounds up to 100 decibels properly, the second component; audio console owns the capacity of maintaining, modifying and delivering clear audio to the computer system within the range of 1 to 100 decibels, the third component; computer system is a normal computer with the minimum configuration of Core2duo and the last and most important component is adobe auditions, that plays, records and saves the audio files on the computer hard disk.

VI. SPEECH CORPUS

Actually, the speech corpus is made to gather the prosodic information residing in recorded Sindhi speech with a 16-Khz sampling rate and 16-bit encoding. For this, we have composed 65 descriptive sentences among them 8 sentences were randomly chosen and given to the male and female speakers with some prosodic boundaries for spoken at different timings. The total number of speakers are 245 among them 197 are male and 48 are females speakers. The words spoken by male speakers are $197 \times 8 = 1576$ and the words spoken by females are $48 \times 8 = 384$ so that the total number of spoken words is 1960.

In our composed sentences, the minimum length of words is one letter and the maximum length is six letters hence, for experiment and results analysis, the recorded sounds of words are classified according to number of letters used to compose the word. The sample of speech sounds in wave form is depicted in “Fig. 3”. The recorded speech was segmented and labeled with prominence values using speech analysis software tools. The segmented sounds along with letters based word classification were separately stored in a speech database.

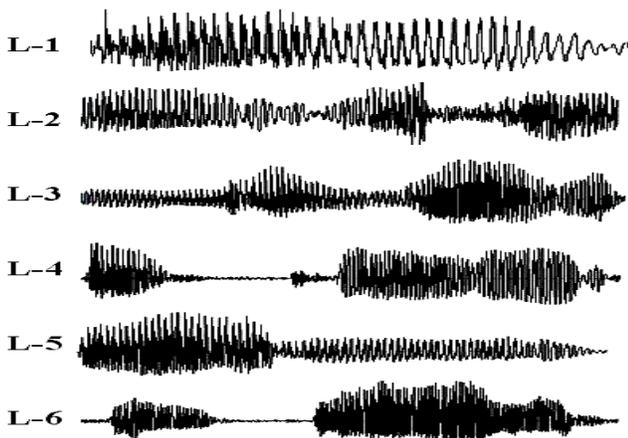


Fig. 3. Sample Wave Forms of Recorded Sounds.

VII. SPEECH ANALYZER

Various speech analysis tools are available for investigating, analysing and restoration the audio speech signals. The PRAAT speech analyzer has great features like spectrographic, intensity measurement and formant analysis and it is freeware software. Due to these characteristics most of the researchers used this software [14]. Hence, PRAAT speech analyzer is selected for experiments and sound analysis of Sindhi adults.

The recorded sound for testing purpose is initially saved in wav format which is later loaded for evaluation. The pitch of the sound file, spectrogram and signal waves of recorded data when uploaded in the PRAAT software it has some default setting for pitch and signal waves. Duration of the word is shown in seconds in a duration bar which appeared after selecting a particular word from a sound file. In the same way to find out the pitch measurement in term of Mean pitch in Hz of any specific word.

VIII. RESULTS

The obtained results using PRAAT are further synthesized and classified into male and female duration and pitch of recorded sounds. Due to the large number of calculations, the calculated sound durations in ms and pitches in Hz are summarized by calculating Mean value. Furthermore, set of received values were measured by the researchers with Standard Deviation (SD) to quantify the amount of variation specifically in pitch using the formula given below [15]. The same process is adopted for presenting the obtained pitch results.

$$\sqrt{\frac{1}{n} \sum_{i=1}^n (\bar{X} - x_i)^2}$$

The recorded sounds of sentences are separated as 1 to 6 letter(s) words to evaluate the duration and pitch of the speakers selected from the five Districts of Upper Sindh. The results are separately presented in four columns: Mean Duration Male (MDM), Mean Duration Female (MDF), Mean Pitch Male (MPM) with SD and Mean Pitch Female (MPF) with calculated SD.

Table II presents the calculated results with 1 letter words. The inferior MDM of 0.1581ms and 0.1426ms and maximum MDM of 0.1924ms and 0.1799ms are received from the male and female speakers. The lowest MPM of 156.43Hz and 158.07Hz are received. The highest MPM of 159.78Hz and 161.14Hz are obtained from the recorded sounds of male and female speakers. We have also calculated Mean SD of each District to measure the variation. The Mean SD of male and female pitches with 1 letter words is depicted in “Fig. 4”. The lowest Mean SD is received with the speakers belongs to district Sukkur and the highest Mean SD is calculated with the speakers of Ghotki.

Table III presents the calculated results of words having 2 letters. The minimum MDM of 0.2160ms and the maximum MDM of 0.2403ms are calculated. The minimum MDF of 0.1715ms and maximum MDF of 0.1906ms are received from the male and female sounds. The lowest MPM of 136.51Hz and highest MPM of 137.87Hz are measured with male sounds. The low MPF of 254.01Hz and the maximum MPF of 255.39Hz are obtained with female sounds. The Mean SD is graphically represented in “Fig. 5”. The lowest SD is obtained from the speakers of Shikarpur District and the highest SD is calculated with the speakers of Ghotki.

The pitch and duration is also calculated with the recorded sounds of male and female speakers. The calculated results of spoken words having 3 letters are presented in Table IV. The lowest MDM of 0.2974ms and 0.2309ms are calculated and greatest MDM of 0.2994ms and 0.2392ms are received.

The lowest MPM of 142.54Hz and the highest MPM of 144.83Hz are obtained from male sounds and the lowest MPF of 268.29Hz and the maximum MPF of 269.45Hz are received from female sounds. For the analysis, Mean SD of male and female pitch is calculated and presented in “Fig. 6”. The surprising results are received from this analysis because the lowest SD is calculated with the speakers of Larkana and the highest SD is calculated with the speakers of Khairpur. However, male and female speakers of district Ghotki have high pitch.

Table V presents the calculated results of 4 letter words. The lowest MDM of 0.3828ms and the highest MDM of 0.4006ms are calculated. From the female speakers, the minimum MDF of 0.3902ms and the maximum MDF of 0.4129ms are calculated from the recorded sounds. The lowest MPM of 147.46Hz and the highest MPM of 147.98Hz can be seen in the table.

The low MPF of 265.33Hz and the maximum MPF of 266.22Hz are received. “Fig. 7” shows the calculated Mean SD of the pitch of male and female speakers. The lowest SD is calculated with the speakers of Shikarpur and Larkana and the highest SD is calculated with the speakers of Khairpur. The variability in the male and female speakers is observed at low level with the speakers of Shikarpur and Larkana but similar cumulative SD is received from the speakers of both districts.

Table VI presents the calculated results of 5 letter words. The lowest MDM is 0.3952ms and the highest MDM is 0.3989ms. The minimum MDF is 0.3636ms and the maximum MDF is 0.3695ms. The lowest MPM is 127.76Hz and the highest MPM is 128.17Hz. The low MPF is 235.53Hz and the maximum MPF is 235.89Hz. The Mean SD of all calculated pitch of male and female sounds is depicted in “Fig. 8”. The lowest Mean SD is calculated with the speakers of Larkana. The Mean SD of male speakers of Khairpur is high but averagely Mean SD of male and female speakers of Sukkur are same and high. It is natural because there is small distance between both cities and peoples are usually shifted and interconnected with same business.

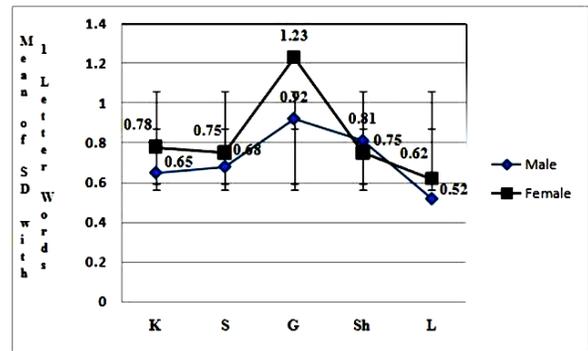


Fig. 4. Mean SD of Male and Female Pitch with 1 Letter Words.

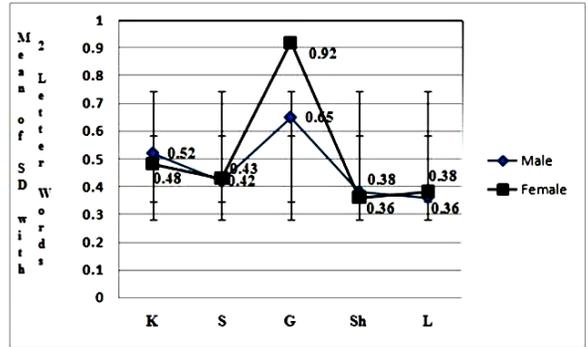


Fig. 5. Mean SD of Male and Female Pitch with 2 Letter Words.

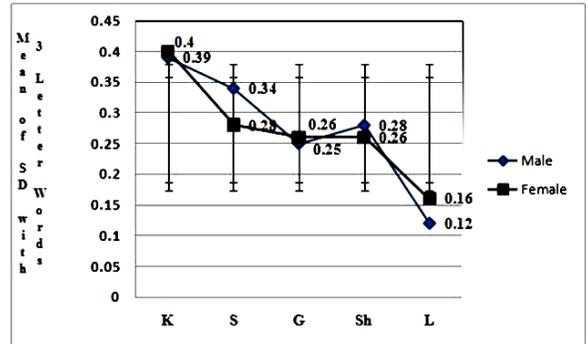


Fig. 6. Mean SD of Male and Female Pitch with 3 Letter Words.

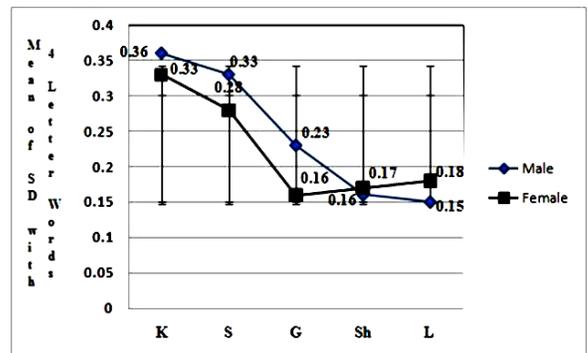


Fig. 7. Mean SD of Male and Female Pitch with 4 Letter Words.

TABLE. II. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 1 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.1552	0.1427	156.43 (0.67345)	158.07 (0.96406)
20 years	14	0.1554	0.1428	156.49 (0.50542)	158.33 (0.92923)
21 years	22	0.1554	0.1429	156.68 (0.89033)	158.35 (0.62593)
22 years	16	0.1556	0.1429	156.73 (0.53236)	158.46 (0.59609)
Sukkur					
19 years	11	0.1551	0.1426	156.65 (0.46367)	158.09 (0.91513)
20 years	12	0.1551	0.1427	156.68 (0.84448)	158.11 (0.76362)
21 years	15	0.1552	0.1427	156.69 (0.87684)	158.12 (0.60557)
22 years	15	0.1552	0.1428	156.78 (0.51511)	158.15 (0.70495)
Ghotki					
19 years	06	0.1921	0.1797	159.34 (0.65965)	160.88 (1.38108)
20 years	09	0.1923	0.1799	159.45 (1.00263)	160.90 (1.52768)
21 years	17	0.1923	0.1799	159.63 (0.85926)	160.93 (0.97471)
22 years	20	0.1924	0.1801	159.78 (1.15661)	161.14 (1.01935)
Shikarpur					
19 years	06	0.1581	0.1577	158.07 (1.00097)	159.77 (0.53488)
20 years	14	0.1582	0.1581	158.14 (0.98074)	159.79 (0.80353)
21 years	15	0.1583	0.1582	158.29 (0.58041)	159.85 (0.81503)
22 years	10	0.1585	0.1582	158.46 (0.68147)	159.86 (0.82335)
Larkana					
19 years	13	0.1708	0.1701	158.22 (1.10646)	158.94 (0.64442)
20 years	03	0.1710	0.1702	158.25 (0.04321)	158.96 (0.35693)
21 years	08	0.1710	0.1704	158.27 (0.47053)	158.96 (0.75082)
22 years	10	0.1711	0.1704	158.31 (0.44960)	158.99 (0.70429)

TABLE. III. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 2 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.2161	0.1716	136.58 (0.48095)	254.05 (0.46912)
20 years	14	0.2162	0.1716	136.56 (0.53630)	254.01 (0.64991)
21 years	22	0.2164	0.1717	136.59 (0.57440)	254.12 (0.49295)
22 years	16	0.2165	0.1719	136.61 (0.47028)	254.17 (0.30033)
Sukkur					
19 years	11	0.2160	0.1715	136.51 (0.42895)	254.21 (0.31493)
20 years	12	0.2162	0.1716	136.55 (0.45345)	254.11 (0.39143)
21 years	15	0.2162	0.1717	136.56 (0.38056)	254.28 (0.53760)
22 years	15	0.2163	0.1717	136.61 (0.41055)	254.31 (0.45835)
Ghotki					
19 years	06	0.2398	0.1903	137.78 (0.69318)	255.27 (1.64697)
20 years	09	0.2399	0.1905	137.82 (0.45722)	255.28 (0.54939)
21 years	17	0.2402	0.1905	137.83 (0.56371)	255.35 (0.90806)
22 years	20	0.2403	0.1906	137.87 (0.89304)	255.39 (0.56721)
Shikarpur					
19 years	06	0.2211	0.1832	137.01 (0.34761)	254.49 (0.34137)
20 years	14	0.2211	0.1833	137.17 (0.37842)	254.33 (0.58161)
21 years	15	0.2215	0.1834	137.23 (0.41825)	254.42 (0.22334)
22 years	10	0.2216	0.1834	137.24 (0.35847)	254.52 (0.29957)
Larkana					
19 years	13	0.2189	0.1814	136.69 (0.40352)	254.17 (0.52989)
20 years	03	0.2189	0.1814	136.60 (0.17664)	254.36 (0.01415)
21 years	08	0.2191	0.1815	136.71 (0.36486)	254.40 (0.43064)
22 years	10	0.2192	0.1815	136.76 (0.49667)	254.47 (0.52972)

TABLE. IV. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 3 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.2975	0.2309	142.55 (0.52502)	268.50 (0.41492)
20 years	14	0.2977	0.2309	142.54 (0.41127)	268.49 (0.42730)
21 years	22	0.2977	0.2313	142.59 (0.24799)	268.52 (0.38579)
22 years	16	0.2978	0.2312	142.64 (0.36012)	268.51 (0.35541)
Sukkur					
19 years	11	0.2974	0.2309	143.21 (0.51085)	268.29 (0.23413)
20 years	12	0.2976	0.2311	143.22 (0.36499)	268.31 (0.37265)
21 years	15	0.2977	0.2316	143.26 (0.25099)	268.36 (0.25007)
22 years	15	0.2979	0.2321	143.29 (0.22663)	268.37 (0.26895)
Ghotki					
19 years	06	0.2991	0.2378	144.75 (0.23296)	269.41 (0.32935)
20 years	09	0.2993	0.2374	144.78 (0.21889)	269.40 (0.42042)
21 years	17	0.2993	0.2380	144.83 (0.31262)	269.42 (0.17314)
22 years	20	0.2994	0.2383	144.80 (0.24741)	269.45 (0.13532)
Shikarpur					
19 years	06	0.2983	0.2392	143.63 (0.31172)	268.85 (0.29603)
20 years	14	0.2981	0.2389	143.59 (0.25574)	268.86 (0.32054)
21 years	15	0.2984	0.2389	143.59 (0.28339)	268.86 (0.24416)
22 years	10	0.2988	0.2390	143.61 (0.25999)	268.88 (0.17135)
Larkana					
19 years	13	0.2977	0.2317	143.45 (0.14444)	268.59 (0.30207)
20 years	03	0.2975	0.2318	143.48 (0.00817)	268.62 (0.01415)
21 years	08	0.2980	0.2322	143.47 (0.10747)	268.63 (0.11554)
22 years	10	0.2983	0.2323	143.45 (0.22199)	268.60 (0.19979)

TABLE. V. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 4 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.3828	0.3904	147.46 (0.39525)	265.33 (0.31539)
20 years	14	0.3831	0.3902	147.53 (0.39305)	265.38 (0.32641)
21 years	22	0.3832	0.3907	147.54 (0.34606)	265.39 (0.39097)
22 years	16	0.3834	0.3906	147.57 (0.29944)	265.43 (0.29307)
Sukkur					
19 years	11	0.3832	0.3906	147.48 (0.26711)	265.36 (0.32485)
20 years	12	0.3832	0.3909	147.49 (0.35555)	265.36 (0.31182)
21 years	15	0.3837	0.3911	147.56 (0.30175)	265.38 (0.19963)
22 years	15	0.3835	0.3913	147.58 (0.39915)	265.39 (0.30115)
Ghotki					
19 years	06	0.4003	0.4126	147.95 (0.30904)	266.17 (0.16351)
20 years	09	0.3999	0.4128	147.93 (0.18997)	266.16 (0.15427)
21 years	17	0.4006	0.4128	147.97 (0.19232)	266.19 (0.14414)
22 years	20	0.4004	0.4129	147.98 (0.23106)	266.22 (0.19558)
Shikarpur					
19 years	06	0.3946	0.3952	147.59 (0.11591)	265.56 (0.17244)
20 years	14	0.3943	0.3951	147.61 (0.16562)	265.57 (0.21075)
21 years	15	0.3945	0.3954	147.63 (0.13525)	265.59 (0.17424)
22 years	10	0.3946	0.3951	147.64 (0.21977)	265.60 (0.12845)
Larkana					
19 years	13	0.3839	0.3912	147.54 (0.23933)	265.41 (0.19807)
20 years	03	0.3843	0.3909	147.56 (0.05354)	265.46 (0.19442)
21 years	08	0.3841	0.3915	147.54 (0.16477)	265.43 (0.18507)
22 years	10	0.3842	0.3918	147.53 (0.13131)	265.45 (0.12633)

TABLE. VI. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 5 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.3958	0.3637	127.80 (0.28794)	235.56 (0.29829)
20 years	14	0.3959	0.3639	127.78 (0.21504)	235.56 (0.06095)
21 years	22	0.3957	0.3640	127.76 (0.24179)	235.53 (0.17719)
22 years	16	0.3961	0.3641	127.78 (0.37423)	235.55 (0.10069)
Sukkur					
19 years	11	0.3952	0.3636	127.77 (0.10296)	235.58 (0.36873)
20 years	12	0.3955	0.3637	127.78 (0.21985)	235.61 (0.13479)
21 years	15	0.3956	0.3636	127.81 (0.20695)	235.62 (0.09041)
22 years	15	0.3956	0.3637	127.82 (0.30479)	235.60 (0.25177)
Ghotki					
19 years	06	0.3986	0.3689	128.17 (0.11972)	235.89 (0.07528)
20 years	09	0.3986	0.3687	128.14 (0.09877)	235.85 (0.08179)
21 years	17	0.3987	0.3688	128.16 (0.12916)	235.87 (0.14492)
22 years	20	0.3989	0.3689	128.15 (0.12045)	235.86 (0.16146)
Shikarpur					
19 years	06	0.3962	0.3695	127.87 (0.22657)	235.67 (0.13881)
20 years	14	0.3962	0.3695	127.88 (0.17217)	235.67 (0.13989)
21 years	15	0.3965	0.3694	127.89 (0.12628)	235.68 (0.08189)
22 years	10	0.3966	0.3693	127.88 (0.07362)	235.69 (0.06588)
Larkana					
19 years	13	0.3957	0.3641	127.81 (0.11046)	235.59 (0.17436)
20 years	03	0.3959	0.3645	127.80 (0.04321)	235.62 (0.07257)
21 years	08	0.3961	0.3643	127.82 (0.05196)	235.61 (0.14169)
22 years	10	0.3962	0.3644	127.79 (0.07655)	235.58 (0.13409)

TABLE. VII. INVESTIGATED MEAN PITCH AND DURATION OF RECORDED SOUNDS OF 6 LETTER WORDS

District and Age Group	n	μ Duration in ms (Male)	μ Duration in ms (Female)	μ Pitch in Hz and σ (Male)	μ Pitch in Hz and σ (Female)
Khairpur					
19 years	09	0.5422	0.4791	150.94 (0.17349)	240.79 (0.15727)
20 years	14	0.5423	0.4794	150.94 (0.15469)	240.81 (0.15933)
21 years	22	0.5426	0.4792	150.97 (0.20631)	240.78 (0.20991)
22 years	16	0.5425	0.4792	150.95 (0.21795)	240.80 (0.21755)
Sukkur					
19 years	11	0.5426	0.4723	150.97 (0.23584)	240.82 (0.21329)
20 years	12	0.5426	0.4724	150.98 (0.22628)	240.83 (0.22329)
21 years	15	0.5429	0.4727	150.96 (0.20865)	240.83 (0.20465)
22 years	15	0.5431	0.4726	150.97 (0.22993)	240.85 (0.21357)
Ghotki					
19 years	06	0.5576	0.4853	151.62 (0.29456)	241.37 (0.21347)
20 years	09	0.5574	0.4851	151.63 (0.24536)	241.36 (0.20554)
21 years	17	0.5575	0.4854	151.66 (0.22054)	241.37 (0.22534)
22 years	20	0.5573	0.4854	151.64 (0.20722)	241.38 (0.23852)
Shikarpur					
19 years	06	0.5486	0.4792	151.13 (0.22138)	240.77 (0.21618)
20 years	14	0.5486	0.4788	151.11 (0.20167)	240.75 (0.21122)
21 years	15	0.5487	0.4793	151.14 (0.21753)	240.74 (0.21711)
22 years	10	0.5487	0.4793	151.11 (0.24956)	240.78 (0.24376)
Larkana					
19 years	13	0.5438	0.4728	150.96 (0.22847)	240.29 (0.19608)
20 years	03	0.5440	0.4728	150.95 (0.23424)	240.29 (0.16083)
21 years	08	0.5438	0.4729	150.95 (0.24546)	240.30 (0.16553)
22 years	10	0.5439	0.4727	150.97 (0.22628)	240.33 (0.23825)

The calculated results of words having 6 letters are given in Table VII. The least MDM of 0.5422ms and the highest MDM of 0.5576ms are measured. The minimum MDF is 0.4723ms and the maximum MDF is 0.4854ms. The lowest MPM is 150.94Hz and the highest MPM is 151.66Hz. The lowest MPF of 240.29Hz and largest calculated MPF of 241.38Hz are received from the recorded male and female recorded sounds. The calculated results of Mean SD are shown in “Fig. 9”. The mean SD of Khairpur shows that male and female pitch with large words are same and the peoples speak slightly. The high SD results are received with the speakers of Ghotki.

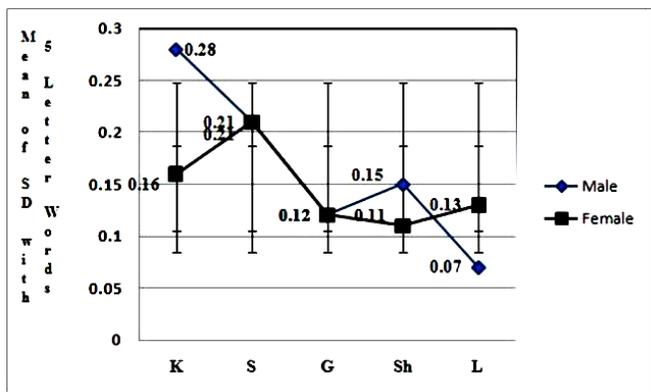


Fig. 8. Mean SD of Male and Female Pitch with 5 Letter Words.

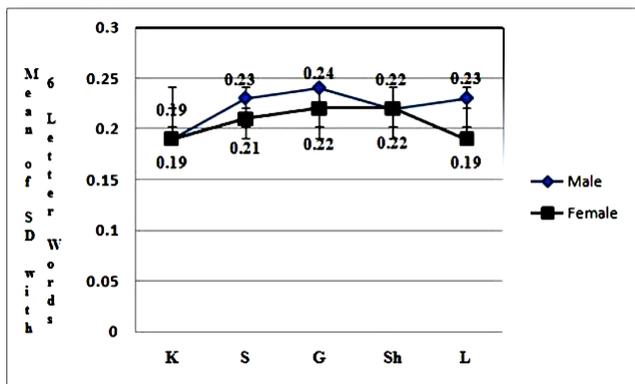


Fig. 9. Mean SD of Male and Female Pitch with 6 Letter Words.

IX. DISCUSSION AND CONCLUSION

According to the phonology, prosody is conceptual phenomena acquiring from the recorded speech. The deep phonetic understanding of cognitive concepts like rhythm, intonation and accentuation are the notations of pitch and duration as these both are core transmitters of the acoustic information and mandatory for designing, developing and validation of prosody module generation for Sindhi language. Hence, we have evaluated and investigated the two core parameters i.e. pitch and duration of the recorded sounds of Sindhi male and female participants and presented in this paper. The outcome of this research study will be helpful to develop efficient and effective Sindhi prosody generation module.

The sentences were given to speakers for investigation of pitch and duration from the recorded sounds. For this, 245 undergraduate students are selected from the five Districts; Khairpur, Sukkur, Ghotki, Shikarpur and Larkana of Upper Sindh having different ages. Eight sentences were given individually which are randomly selected from the prepared 65 sentence. The total words were spoken by male and female speakers comprises on 1960.

Furthermore, all the recordings are recorded at the radio station, Khairpur on different timings and due to short time only limited time was given by the authority. The recorded sounds are segmented into words and then stored into the computer. The PRAAT speech analyzer is used for segmentation of sounds of words and analysis of the recorded sounds.

The duration and pitch of the recorded sounds are separately calculated and presented according to the words based on the number of letters. The lowest MPM 127.76Hz is calculated with 5 letter words and the highest MPM of 159.78Hz is recorded with 1 letter word. Experimented results proved that the pitch in Sindhi sounds is entirely based on the syllabification. Almost, 1 letter word has one syllable while it is possible the number of syllables increases when word is based on more than 1 letters. It is also observed that the pitch of Sindhi people is high at the start of the word and at the end of the syllable particularly when syllable ends with ‘Jazm’. The cumulative Mean SD of all words spoken by Male and Female speakers is depicted in “Fig. 10”.

On the basis of received results and the calculate Mean SD of all words spoken by the male speakers it is found that speakers pertaining to district Ghotki have high pitch as compare to others. And the speakers of Larkana district have low pitch. However, little variation in pitch is observed with the male speakers of Khairpur, Sukkur and Shikarpur Districts. Approximately same results of the female speakers are received during the experiments. It is found that the pitch of the female speakers are high then the pitch of male speakers but the variation is only 0.072.

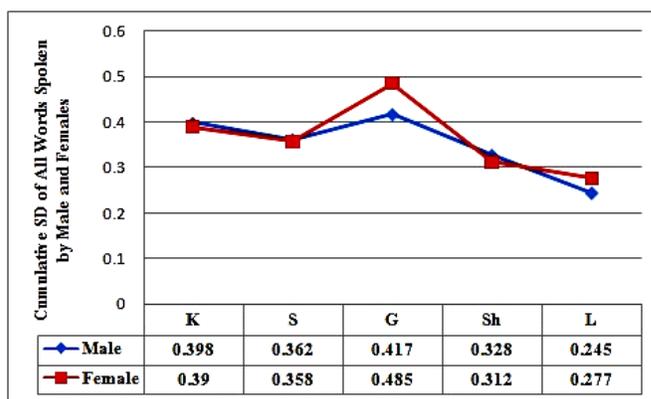


Fig. 10. Cumulative SD of Male and Female Pitch.

REFERENCES

- [1] Hassan, "Assimilation and Incidental Differences in Sindhi Language", Eurasian Journal of Humanities, vol. 2, Issue 1, pp. 1-16, 2016.
- [2] J. A. Mahar, G. Q. Memon, "Phonology for Sindhi Letter-to-Sound Conversion", Journal of Information & Communication Technology, Vol. 3, No. 1, pp. 11-21, Spring 2009.
- [3] S. Hoffmann, B. Pfister, "Employing Sentence Structure: Syntax Trees as Prosody Generators", In 13th Annual Conference of the International Speech Communication Association, pp. 470-473, September 2012.
- [4] P. B. Dasgupta, "Detection and Analysis of Human Emotions Through Voice and Speech Pattern Processing", International Journal of Computer Trends and Technology, Vol. 52 No. 1, pp. 1-3, 2017.
- [5] K. Waghmare, S. Kayte, B. Gawali, "Analysis of Pitch and Duration in Speech Synthesis", Communications on Applied Electronics, Vol. 4, No. 4, pp. 10-18, February 2016.
- [6] X. Sarasola, E. Navas, D. Tavarez, L. Serrano, I. Saratxaga, I. Hernaez, "Application of Pitch Derived Parameters to Speech and Monophonic Singing Classification", Applied Science, Vol. 9, 3140, pp. 1-16, August 2019.
- [7] H. Shaikh, J. A. Mahar, G. A. Malah, "Digital Investigation of Accent Variation in Sindhi Dialects", Indian Journal of Science and Technology, Vol. 6, No.10, pp.5429-33, 2013.
- [8] J. A. Mahar, G. Q. Memon, H. A. Shah, "Perception of Syllables Pitch Contour in Sindhi Language", Proceeding of the IEEE Natural Language Processing and Knowledge Engineering, pp. 593-597, September 2009, China.
- [9] A. M. Abbasi, S. Hussain, "The Role of Pitch between Stress and Intonation in Sindhi", ELF Annual Research Journal, Vol. 17, pp. 41-54, 2015.
- [10] A. M. Abbasi, H. Pathan, M. A. Channa, "Experimental Phonetics and Phonology in Indo-Aryan & European Languages", Journal of Language and Cultural Education, Vol. 6, No. 3, pp. 21-52, 2018.
- [11] A. Keerio, N. Channa, Y. A. Malkani, B. Qureshi, J. A. Chandio, (2014), "Acoustic Analysis of the Liquid Class of Consonant Sounds of Sindhi", Sindh University Research Journal (Science Series.), Vol.46, No. 4, pp. 505-510, 2014.
- [12] M. Farooq, "Acoustic Analysis of Corner Vowels in Six Indigenous Languages of Pakistan", Journal of Research in Social Sciences, Vol. 6 No. 2, pp. 2305- 6533, 2018.
- [13] M. Swain, A. Routray, P. Kabisatpathy, "Databases, Features and Classifiers for Speech Emotion Recognition: A Review", International Journal of Speech Technology, Vol. 21, pp. 93-120, 2018.
- [14] HT Lathadevi, S. P. Guggarigoudar, "Objective Acoustic Analysis and Comparison of Normal and Abnormal Voices", Journal of Clinical and Diagnostic Research, Vol. 12, No. 12, pp. 1-4, December 2018.
- [15] M. P. Gelfera, L. Sara, "Speaking Fundamental Frequency and Individual Variability in Caucasian and African American School-Age Children", American Journal of Speech-Language Pathology, Vol. 23, pp.395-406, 2014.