www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 4 Issue 7 July 2015, Page No. 13165-13168

Pitch, Formant Frequencies, Intensities as Speech Cue to preceived age

Tarunam Chaudhary[1], Sawati Bhasin[2], Saloni[3]

[1] M.Tech Scholar, Geeta Institute of Management & Technology, Kurukshetra (Haryana)
tanuchoudhary61@gmail.com

[2] Assistant proffessor, Geeta Institute of Management & Technology, Kurukshetra (Haryana)
missbhain@gmail.com

[3] Assistant proffessor, Haryana Engineering college Jagadhri (Haryana)
er.saloni83@gmail.com

Abstract: ***Voice of a person gives an idea regarding the age, gender and health of a person. In this paper we tried to have an idea about the age of speaker. Our study includes 30 speakers of 14-30, 31-49, 50-80 age groups. we have considered different voice parameters fundamental frequency F0 , formant frequencies F1, F2, F3, F4 and intensities. In this different voiced vowels are used to study the changes between the three age groups.***
Keywords: Voice parameters, Fundamental frequency (F0), Formant Frequencies (F1, F2, F3, F4) and Intensities.

1. Introduction

The voice is generated with the help of biological patterns. Voice is produced by the airflow through the lungs which is converted by the vocal folds. Spectral shaping can be done by the converting airflow through vocal tract. Cavities are used to forward the airflow to the lips to produce the sound. Voice is produced with the help of voice organ. The voice organ is classified into main three categories; lungs, larynx, and vocal tract. Lungs are used to provide the air pressure to the larynx. Larynx converted the airflow as an input given by the lungs. It contains two vocal folds which are made by masses of flesh, ligament, and muscles. The vocal folds are used to produce the stretching between the front and back parts of larynx. Glottis is as a space between two folds. the focal folds are open during breathing process. The folds can be vibrate depending upon the speaking state. Voice source like vowels then the vocal folds are in a vibrating state . The vocal folds are open and closed quickly when the voice source is vowels. For the other voice sources than vowels the vocal folds are not open and closed quickly. After the process of larynx the signal forward to the vocal tract. Vocal tract contains three cavities ; pharynx cavity, oral cavity, and nasal cavity. These are used for giving shape to the changed airflow and then adjusting the quality of the speech. When the vibration of the vocal folds then the voice can be got in the form of pulse and this pulse is known as glottal pulse. The glottal pulse presents in the form of open and closed. When the folds in a closed position means airflow is on a zero level and when the vocal folds in an opened position means airflow shows increasing values. When the persons age increase not only face but also voice will change so voice is an important parameter which is used to detect the age, emotions, beauty of the persons, happiness, sadness, stress. Here we identifying the age based on the voice parameters like Pitch, Intensity and Formants (F1, F2, F3,F4) by the use of vowels 'a', 'e', 'i', 'o', 'u'[1]. When the speech is produced then we can get many types of information about the speaker like vocal organs, gender, speaker's age range and speaker's emotional state[2]. Larynx changes after when the person reached its maximum size in puberty. This changes in larynx with age changes the affect mainly to the fundamental frequency and the voice quality. Persons listeners having the quality to detect the persons age at the different levels of the age with the help of frequency because fundamental frequency of the females rises when age between (10-20) years after that the frequency decreases in the mid age from (30-50) year after that the frequency remains constant at the age level of (above 50) years [3]. Because speech information can be represented in the forms of fundamental frequencies, Formants frequencies (F1, F2, F3, F4), Intensities, Jitter, Shimmer, Autocorrelation, Harmonics to Noise ratio (HNR) and Noise to Harmonic ratio (NHR). By the estimation of these information we can easily detect the age at the different levels[4].

2. Data Collection and Methodology

In this paper we have recorded the voices of 30 healthy female persons of different age groups . 10 aged between 14 - 30 years , other 10 aged between 31 - 49 years and other 10 aged between above 50 years. Voices are analyzed and following voice parameters like Pitch, Intensity and the Formant frequencies(F1, F2, F3, F4) are extracted. All the voice samples were recorded with no background noise in the home environment with the help of microphone M27 which have frequency response: 100Hz- 16kHz, sensitivity: -58dB ±3dB, and S/N ration: More than 60dB. Then this voice is analyzed and various features are extracted based on the samples of 5 vowels (a, e, i, o, u) .we analyzed the three samples of each vowel. Formant frequencies (F1, F2,F3, F4),

Pitch and Intensity are computed these characteristics are calculated for adult, mid and old age. After extraction the formant frequencies of each vowels the average was calculated for F1,F2,F3,F4, Pitch and Intensity. Frequencies of each vowels has been computed to check age related changes among the speakers.

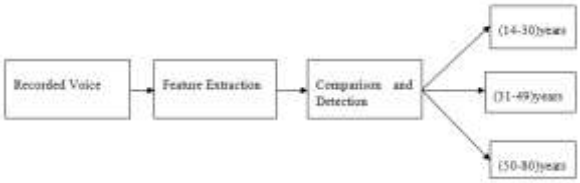

Figure 1. Block Diagram

(i) Fundamental frequency: Fundamental frequency is defined as the lowest frequency of a periodic waveform.

$$F0=\frac{1}{T}$$

(ii) Intensity or Vocal Intensity: Intensity or vocal intensity of the speech signal refers to the loudness effect of speech signal. Vocal intensity is related to the sub glottis pressure of the airflow, which depends on the tension and the vibrations of the vocal folds.

(iii) **Formant frequency:** The formant frequencies can be estimated by taking the frequency response of the vocal tract filter. The peaks of the response are the formant frequencies. which have, in the spectrum, a high degree of energy.

3. RESULTS and DISSCUSION

We detected different age groups based on the voice parameters as F0, F1, F2, F3, F4 and Intensity. We concluded that the frequency levels decreased with increase in age for females by . The following bar graphs proved the age detection process. In figure 2 compare the F1 at the three age groups for all Vowels ('a', 'e', 'i', 'o', 'u') we concluded that F1 decreased when the increased in the age groups. Similarly in Figure 3 compare the F2 at three age groups for vowel ('a' ,'e', 'i', 'o', 'u') there is also F2 decreased with increased in age groups. In figure 4. compare the F3 at three age groups for vowel ('a' ,'e', 'i', 'o', 'u') this is also decreased with increased with age. In figure 5 we compare the F4 at the three age groups for all Vowels ('a', 'e', 'i', 'o', 'u') in this F4 increased with increased with age for the vowels ('a', 'e', 'i', 'o') ,but the F4 decreased with increased in age only for vowel 'u'. In figure 6 we compare the F0 for all vowels ('a', 'e', 'i', 'o', 'u') in this F0decreased for increased in age. In figure 7 we compare the intensities for the all vowels ('a', 'e', 'i', 'o', 'u') it is also decreased with increased with age.

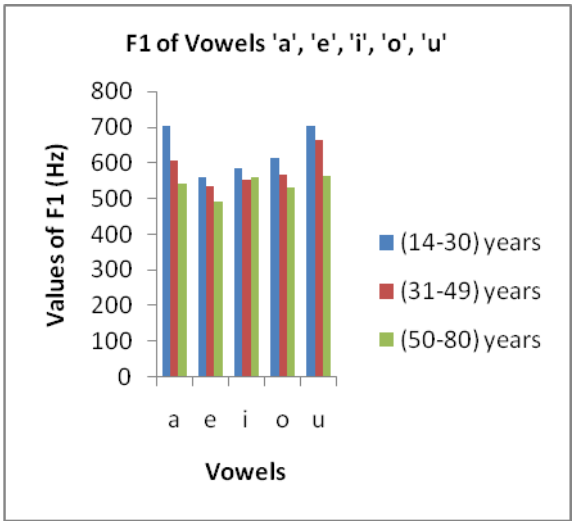

Figure 2: Compare the F1 at the three age groups for all Vowels ('a', 'e', 'i', 'o', 'u')

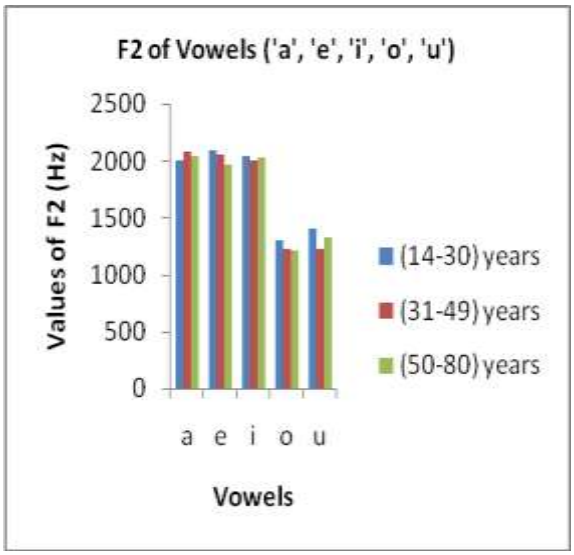

Figure 3: Compare the F2 at three age groups for vowel ('a' ,'e', 'i', 'o', 'u')

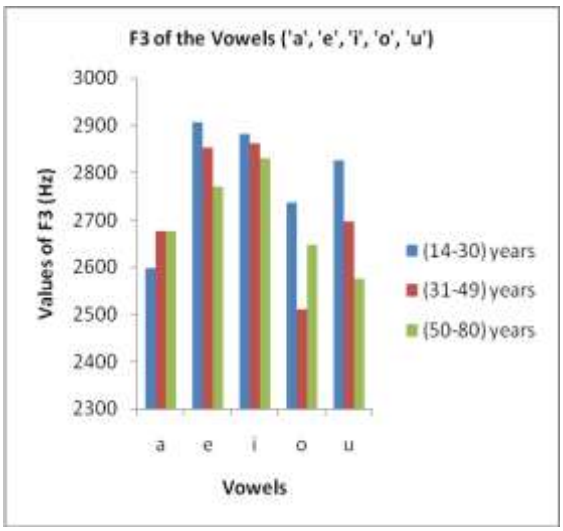

Figure 4: Compare the F3 at three age groups for vowel ('a' ,'e', 'i', 'o', 'u')

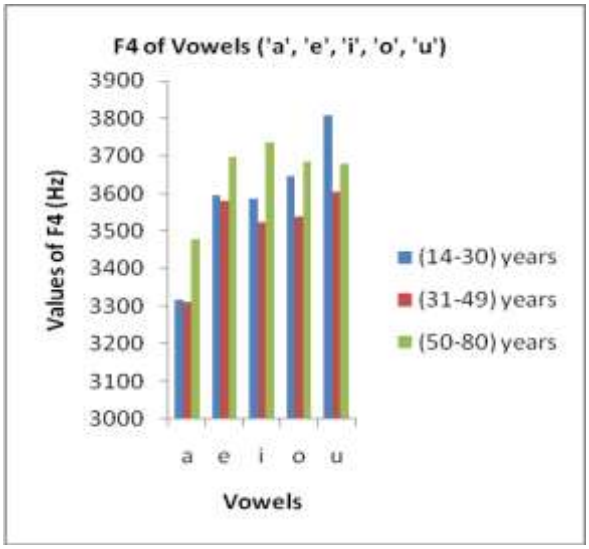

Figure 5: Compare the F4 at the three age groups for all Vowels ('a', 'e', 'i', 'o', 'u')

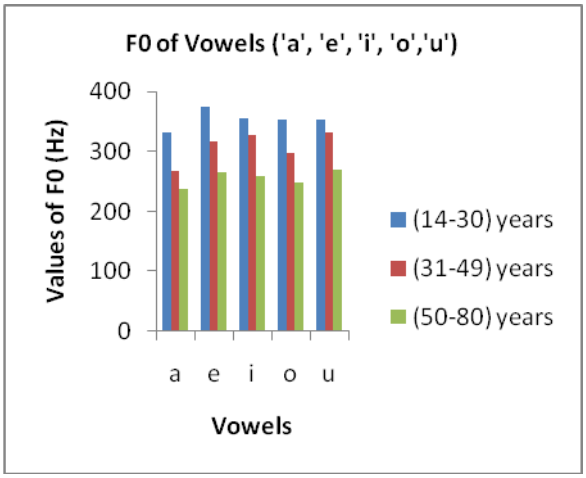

Figure 6: Compare the F0 at the three age groups for all Vowels ('a', 'e', 'i', 'o', 'u')

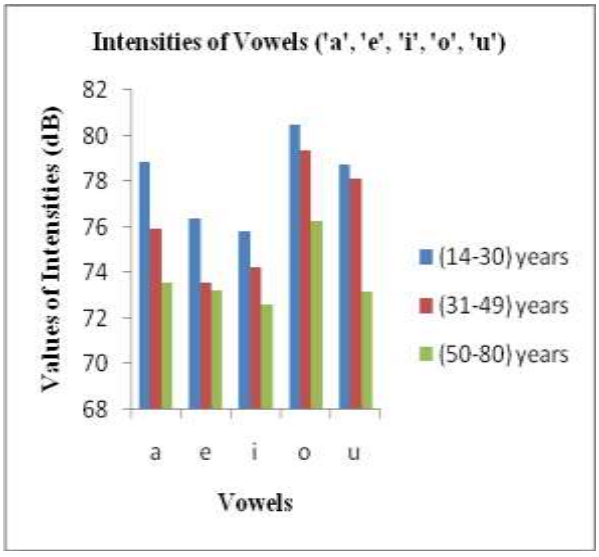

Figure 7: Compare the Intensities at three age groups for vowel ('a' ,'e', 'i', 'o', 'u')

4. CONCLUSION

In this paper, we give statement about the practical studies on changes in different vocalizations parameters with aging for vowels. Fundamental frequency changes in stable ways for the speakers. we have concluded that formant frequencies also changes with age changes. On average, Fundamental frequency (F0), Formants frequencies (F1, F2, F3, F4) and Intensities decreased with aging. The fundamental and formant frequencies and intensities are high for the ages between (14-30) years. These parameters are decreased with the increase in ages between (31-50) and (50 -80) years

References

[1] Mongia, Puneet Kumar, and R. K. Sharma. "Estimation and statistical analysis of human voice parameters to investigate the influence of psychological stress and to determine the vocal tract transfer function of an individual." Journal of Computer Networks and Communications 2014 (2014).

[2] Mengdi Yue, Ling Chen, Jie Zhang, and Hong Liu, "Speaker Age Recognition Based on Isolated Words By Using SVM",2014 IEEE.

[3] Schötz, Susanne. "Acoustic analysis of adult speaker age." Speaker Classification I. Springer Berlin Heidelberg, 2007. 88-107.

[4] Das, Biswajit, et al. "Effect of aging on speech features and phoneme recognition: a study on Bengali voicing vowels." International Journal of Speech Technology 16.1 (2013): 19-31.

[5] Reubold, Ulrich, Jonathan Harrington, and Felicitas Kleber. "Vocal aging effects on F 0 and the first formant: a longitudinal analysis in adult speakers." Speech Communication 52.7 (2010): 638-651.

[6] Płonkowski, Marcin. "Using bands of frequencies for vowel recognition for Polish language." International Journal of Speech Technology 18.2 (2015): 187-193.

[7] Tani Rajput, Vikas Mittal and Tarun Gulati," Voice Parameters Analysis for Detection of Effect of Stress on Voice: A Review", IJRIT 2015

[8] El Haddad, Kevin, et al. "Speech-laughs: An HMM-based approach for amused speech synthesis." Internation Conference on Acoustics, Speech and Signal Processing (ICASSP 2015).

[9] Brückl, Markus, and Walter Sendlmeier. "Aging female voices: An acoustic and perceptive analysis." ISCA Tutorial and Research Workshop on Voice Quality: Functions, Analysis and Synthesis. 2003.

[10]Vojnović, Milan, Ivana Bogavac, and Ljiljana Jeličić Dobrijević. "Maximal Vowel Space Metod In Analysis of Vowels in Prelingual Speech Phase." Specijalna edukacija i rehabilitacija 13.2 (2014).

[11] Moerman, Mieke, et al. "Objective evaluation of the quality of substitution voices." European Archives of Oto-

Rhino-Laryngology and Head & Neck 261.10 (2004): 541-547.

[12] Shivaji Chaudhari and Ramesh Kagalkar, "A Review of Automatic Speaker Age Classification, Recognition and Identifying Speaker Emotion Using Voice Signal", International Journal of Science and Research (IJSR) 2012.

[13] Pati, Debadatta, and SR Mahadeva Prasanna. "Speaker recognition from excitation source perspective." IETE Technical Review 27.2 (2010): 138-157.

[14] Fourakis, Marios S., et al. "Effect of frequency boundary assignment on vowel recognition with the Nucleus 24 ACE speech coding strategy." Journal of the American Academy of Audiology 15.4 (2004): 281-299.

[15] Rabiei, Mohammad, and Alessandro Gasparetto. "A Methodology for Recognition of Emotions Based on Speech Analysis, for Applications to Human-Robot Interaction. An Exploratory Study." Paladyn, Journal of Behavioral Robotics 5.1 (2014).