

Effect of Sample Duration and Temporal Resolution on Perturbation Measurements

Syed Absaruddin Hasan

Student, A.I.I.S.H, Mysore

Voice is the laryngeal modulation of pulmonary air stream which is further modified by the configuration of vocal tract, (Michael and Wendhal, 1971).

Voice plays an important role in speech communication. The production of voice requires a complex and precise control by the central nervous system on a series of synchronous events in respiratory, phonatory and resonatory systems. Any anatomical and physiological deviation in these systems would lead to voice disorder. The evaluation of voice includes acoustic analysis of voice. The acoustic analysis of voice may be one of the most powerful and attractive method for assessing the phonatory function as it is non-invasive, objective and quantitative.

One of the important and powerful acoustic measure of voice is pitch and amplitude perturbation. These methods have been extensively used in voice clinics [Lieberman (1961,1963); Koike (1973); Kitazima and Gould (1976); Horii (1979); Murry and Doharthy (1980); Askenfelt and Hammerberg (1986); Higgins and Saxman (1989); Venkatesh, Satya and Jenny (1992); Gokul Krishna (1992); and Pathak (1994)].

The measurement of pitch and amplitude perturbation is influenced by several factors. They are :-

1. Sample duration.
2. Temporal resolution (Sampling frequency).
3. Pitch extraction method.
4. Type of microphone.

The various sample duration which were taken by different investigators was:- Horri, (1980 and 198F); Fritz and Frank (1985) had taken 3 secs duration; Imaizuma (1985) had take, 0.5 sec duration.

The earlier studies whidi measured pitch and amplitude perturbation used sampling frequencies from 8 KHz to 50 KHz. Horii (1980 and 1985) had taken 40 KHz; Fritz and Frank (1985) had taken 25 KHz; Imaizuma (1986) had taken 20 kHz; Milenkovic P. (1987) had taken 8.3 KHz; Venkatesh, Sathya and Jenny (1992) had taken 16 KHz with interpolation; Raghunath, Baldva and Venkatesh (1992) had taken 16 KHz with interpolation.

- The present study is aimed at studying the effect of duration of voice sample and temporal resoJution of perturbation measurements. The sample durations of 0.5 secs, 1 sec, 2 sec and 3 sees and temporal resolution of 8,000 Hz and 16,000 Hz are compared for their efficacy in the measurement of frequency and amplitude perturbations.

Thirty normal adult males, thirty normal adult females and 10 dysphonics served as subjects. The age range of normal subjects varied from 19 to 35 years., whereas the age range of dysphonics group varied from 30 to 55 years, three trials fo EGG recordings for 5 seconds, for each of the vowels /a/, /i/ and /u/ were obtained using Kay-laryngograph. The EGG recordings were digitized at 16 KHz sampling frequency using 12 bit Analog-to-Digital Converters. The

digitized data was stored on the hard disk of PC-AT 386. The digitized EGG waveforms were smoothed, differentiated and peak-picking method was used to extract the fundamental frequency and intensity. The obtained cycle-to-cycle fundamental frequency and intensity data were subjected to further computations using PC-AT 386 to obtain Jitter Ratio (JR), Directional Perturbation Quotient (DPQ), Shimmer in (dB), Directional Perturbation Quotient for Amplitude (DPQ-A).

Five normal adult males and five normal adult females with the age range of 18 to 25 years served as subjects for the second experiment. Three trials of EGG recordings for 5 seconds, for each of the vowels /a/, /i/ and /u/ were obtained using Kay-laryngograph. The EGG recordings were digitized at 16 KHz sampling frequency using 12 bit analog-to-Digital converter.

The digitized data was stored on the hard disk of PC-AT 386. The digitized data was further digitally filtered for 3.5 KHz and then down sampled to 8 KHz and this data was also stored in a separate file. By this, it was possible to obtain a digitized data for the same voice at two different sampling frequencies i.e., 16KHz and 8 KHz. The digitized EGG wave forms were smoothed, differentiated, interpolated and peak-picking method was used to extract the fundamental frequency and intensity. The obtained cycle-to-cycle fundamental frequency and intensity data were subjected to further computations using PC-AT 386 to obtain Jitter Ratio, Directional Perturbation Quotient, Shimmer in, Directional Perturbation Quotient for Amplitude.

The measured pitch and amplitude perturbation values were subjected to descriptive statistical procedures, t-test and to ANOVA. Based upon the results obtained the following conclusions were drawn:

1. There is no significant difference between the three trials for all the four measures of pitch and amplitude perturbations. Hence, it is not really necessary to take all the three trials for the measure of pitch and amplitude perturbation. Even a single trial can give same amount of information.
2. There was no significant difference between adult normal males and adult normal females for four pitch and amplitude perturbation measures.
3. There was a significant difference between the vowel /a/ and /i/, /u/ whereas these was no significant

difference between the vowel /i/ and /u/ for the four measures of pitch and amplitude perturbations. The difference observed between the low vowel i.e., /a/ and the high vowels i.e., /i/ & /u/ may be due to the different amount of laryngeal tension applied during the production of these vowels. Hence it is not necessary to take all the three vowels for the purpose of voice evaluation. The voice recording of one high vowel /i/ & one low vowel; /a/ is sufficient to obtain valid and reliable data for pitch and amplitude perturbation

4. There was no significant difference between different sample durations (i.e., 0.5, 1, 2 and 3 seconds) for DPQ, shimmer (dB) and DPQ-A, whereas Jitter ratio (JR) showed that its values for 0.5 second sample duration is significantly different from other sample duration. (1, 2, 3 sec) J.R. did not show any significant mean difference across other three sample durations (i.e., 1, 2 & 3 seconds).

From the above result it may be inferred that larger sample duration is not necessary to obtain a reliable and valid pitch and amplitude perturbation measures. Hence the recording of the voice sample for only one second is sufficient for the measurement of pitch and amplitude perturbation.

5. There is no significant difference between the two temporal resolutions. (16 KHz & 8KHz) for the measurement of DPQ, shimmer, -- dB & DPQ-A, however, in the measurement of Jitter Ratio there was significant difference between two temporal resolutions (i.e., 16 KHz & 8KHz). Due to the low experimental sample and varied result, it is not possible for the investigator to draw any concrete reliable inference, however, it may be suggested to use the sampling rate higher than 16 KHz with interpolation for the purpose of measurement of perturbation measures.

Limitations:

1. The sample size of the dysphonic population taken for checking reliability and validity of the data was very small.
2. The sample size taken for the assessment of significant difference across different sampling frequencies (i.e., 16 KHz & 8 KHz) was very small to draw valid conclusions.