

Objective analysis of voice in normal young adults

KC Toran¹, Lal BK²

¹Associate Professor, ²Speech Language Pathologist, Department of ENT-HNS, Kathmandu Medical College, Sinamangal, Nepal

Abstract

Background: Acoustic vocal parameters measure frequency, intensity (amplitude), perturbation (jitter and shimmer) and dynamic range of the voicing vocal folds. Studies have established that a normal standard data is necessary for acoustic analysis.

Objective: The aim of the present study is to standardise Jitter, shimmer, harmonic to noise ratio (HNR) and fundamental frequency (F_0) for young adults with normal voice.

Materials and methods: Values for acoustic voice measurements were obtained from 50 normal individuals with equal number of sexes, without sign and symptoms of voice problems. The vocal data measurement was performed with Doctor Speech (DRS) Tiger Electronics, USA.

Results: Voice analyses were performed with a sustained vowel //i//. The jitter and HNR values were same [1.6% ($\pm 0.47/\pm 0.43$) and 25.8dB ($\pm 2.62/\pm 2.72$)] for both the genders. For the males, the jitter was 0.14% (± 0.02) and 0.16% (± 0.04) for female gender. There was a significant difference in the HNR ($P=0.000$) with 170.05HZ (± 32.78) and 246.45HZ (± 39.73) respectively for male and female genders.

Conclusion: Our results differ from the various literatures; therefore it is important to standardise the program that we use before applying the values for tests designed for a different kind of population.

Key words: Acoustic analysis, Perturbation, harmonics to noise ratio, fundamental frequency, standardised voice.

Voice disorders are generally assessed subjectively by a speech language pathologist or by a laryngologist. This form of perceptual analysis of voice disorders has significant limitations that can lead to confusion between the observers while treating a dysphonic patient. In search of more objective and consistent measures, investigators have introduced the computer based acoustic correlates of abnormal voice qualities^{1,2}. Computer software based acoustic analysis has becoming more popular amongst the otolaryngologists who are interested in voice and their disorders. Objective acoustic instruments can also be used to “facilitate the patients’ awareness of vocal characteristics, manners of voice production, self monitoring, visual feedback, reinforcement as well as documentation in the form of statistical analysis and visual displays as evidence of progress”³. They state that “instrumental measures of the vocal function are an integral component of the clinical process, rather than a supplement to assessment and treatment”. More frequently instrumental measures are incorporated into the clinical practice, the more experienced the clinician becomes in selecting and implementing the measures and in analysing and applying the voice data³.

Acoustic parameters measure fundamental frequency (F_0), Intensity (amplitude), perturbation (jitter and shimmer), harmonic to noise ratio (HNR) and dynamic vocal range⁴. F_0 is the vibratory rate of the vocal folds. It can be measured in hertz (Hz). This measurement reflects the physiological limits of the patients voice⁵. During sustained

vibration, the vocal fold will exhibit slight variation of F_0 and amplitude from cycle to cycle; these phenomena are called frequency perturbation (jitter) and amplitude perturbation (shimmer)⁴. These changes reflect the slight differences in mass, tension and biochemical characteristics of the vocal folds, as well as slight variation in their neural control. Perturbation correlates with the perceived roughness and hoarseness in the voice⁵. HNR is an average ratio of energy of the harmonic components in the range ratio 70-4500Hz to

Correspondence

Dr. Toran KC
Associate Professor, Department of ENT-HNS,
Kathmandu Medical College,
Sinamangal, Kathmandu, Nepal
E-mail: torankc@yahoo.com

the inharmonic components energy in the range 1500-4500Hz.⁶ Harmonics are those frequencies that are integer multiples of the F_0 . If a sound has a steady F_0 and if all the component frequencies are harmonics, then each cycle will be identical. Frequencies that are not integer multiples of F_0 are regarded as noise. The amount of energy conveyed in the F_0 and its harmonics, divided by the energy in noise frequencies, is the harmonic-to-noise ratio. This parameter correlates with the perception of vocal roughness⁷. As the degree of hoarseness increases the noise component increases and replaces the harmonic structure in the spectrogram. Normal voices have low level of noise⁵.

With above mentioned objective criteria one can now easily record and analyse the phonatory data of vibrating vocal folds. And such measurements are only possible with these kinds of acoustic instruments. None of the English journal has published any voice data of Nepalese population so far. Acoustic analysis standardisation also educates, simplifies, saves time, money and effort and assures certification⁸. We evaluated the normal voice data of young Nepali volunteers of both the sexes for quick standard reference and to understand their normal vocal parameters.

Materials and methods

Fifty informed volunteers from age 20 to 39 years were included in the study. The age range thus selected was to avoid the vocal changes of the growing age and the old age.⁹ The data were collected and analysed with Dr. Speech (DRS) Tiger Electronics, USA in between January and March 2008. Both the sexes were of equal in number. The persons with smoking habit, vocal abusers or recent history of cough and cold were excluded from the study. All of the volunteers were non singers. All of them were perceptually analysed with GRBAS scale¹⁰. Those with rating more than 0 was excluded from the

study. All the data were recorded in a sound treated room, during the morning hours before 12 noon to avoid error due to vocal fatigue or overuse. Voice recording was done directly in their habitual loudness with a microphone placed at 30cm from the volunteer's upper lip, and the person sitting in a comfortable position. For the acoustic assessment sustained vowel //i// was preferred instead of continuous speech¹¹.

All the recordings were repeated for three times. To avoid voice onset effects, first 500ms of the voice data were not included for the study. Data were analysed from mid 3 seconds of the vowel //i//. Analysis was performed in terms of perturbation (jitter and shimmer), fundamental frequency (F_0) and harmonic to noise ratio. The statistical analysis was performed in statistics software SPSS 16.0 for windows.

Results

A total of fifty volunteers of equal number of sexes were evaluated for the acoustic analysis. The evaluations were done in terms of jitter, shimmer, HNR and fundamental frequency (F_0). Both the gender groups were also analysed separately. The average values of all the four parameters are as depicted in the table below.

It was observed that there were no significant difference in jitter, shimmer and HNR ($P > 0.05$) of acoustic parameters between the two groups of both the sexes, except the F_0 ($P = 0.000$). Only the female averages for jitter and fundamental frequency were higher than the male groups. The data were also analysed in terms of two age groups of 20-29 years and 30 to 39 years of age for both the sexes separately. The analysis of all the four parameters revealed no statistical significance between these two age groups of both the genders either ($P < 0.05$).

Item	Male		Female		p-value*
	Average values(N=25)	Mean(SD)	Average values (N=25)	Mean(SD)	
Jitter	0.14%	0.14(0.02)	0.16%	0.14(0.04)	0.568
Shimmer	1.6%	1.66(0.47)	1.6%	1.57(0.43)	0.494
HNR	25.81dB	25.81(2.62)	25-8dB	25.88(2.75)	0.930
F_0	170.05Hz	1.75(32.78)	246.45Hz	2.55(39.73)	0.000

* t-test was applied.

Discussions

The average of jitter for sustained vowels /i/ observed in males was 0.14% while in females it was 0.16%. The normative range for the DRS data is 0.5%. In another study with DRS revealed 0.40% and 0.38% for normal males and females respectively⁶. Relative jitter values in our study were much lesser than other studies. Wuyts¹² et al has average relative jitter for males was 0.81% while lesser value observed was 0.037%^{13, 14}. Amongst the females the average jitter value was 0.16% which is slightly higher than the males. As to the average jitter for vowel /i/ our result was lower than the one found by Aroujo et al¹⁵ (0.85%), Higgins¹⁴ (0.38%), wuyts¹² (1.04). Even though the jitter value was different between the two gender groups, they were not significantly different ($P=0.568$). The lack of difference in the jitter value also don't support the hypothesis about which would be the reasons for the better sound control shown by females¹⁶. The average value for shimmer was 1.6% which was same for both the gender groups while the DRS database is 3.0%. Similar data was also observed in other studies^{15,16}. But, our values for relative shimmer were much lower than those observed by smits⁶, which were 3.98% and 4.5% for males and females respectively. In other studies relative shimmer for males was 3.6% and females were 3.4% respectively¹². As compared to other studies we observed relatively lower shimmer values in both the gender groups⁶. Increased Jitter or shimmer values have been associated with phonatory instability due to aging, neurological disorder like amyotrophic lateral sclerosis¹⁷ and various laryngeal pathologies¹⁸. Voice perturbation is a probable indicator for a physiological disorder. However, several factors, including frequency, intensity and vowel selection, would affect various phonatory stability measures and that the highest perturbation values almost always occur in low frequency- low intensity situations¹⁹.

The average of harmonic-to noise ratio was similar in both males and females groups, i.e., 25.81dB and 25.80dB respectively. Similar HNR values for both the gender subgroups with respect to age were also observed in other studies⁶. But the HNR data observed with the DRS in a Belgian population has value of 16.31dB and 21.09dB for male and female gender groups⁶. Our values for both the gender groups were similar and above the ones observed by Rodrigues et al. which were 8.63dB and 10.17 dB for male and female groups respectively. The significantly higher HNR values in women in comparison to men may be related to the fact that men use fluid voice more frequently. It leads to lesser glottal closure and this favours voice production with less harmonics and or greater amount of glottal noise²⁰. Wuyts et al observed a value of 17.5dB and 18.3dB for male and female gender respectively¹².

Average F_0 was included because of its possible effect on the perception of voice quality²¹. F_0 is an acoustic measure that directly reflects the vibrating rate of vocal folds and is expressed in Hertz. The average F_0 was 170.05Hz and 246.45 Hz for males and females respectively. The value of F_0 was higher in women than the men is due to the vocal folds in men being longer and thicker⁴. In another acoustic study with the DRS, the average value of F_0 in Belgian population was 115.5Hz and 197.0Hz respectively.⁶ On reviewing the several studies the F_0 for males has ranged from 115.0 to 129.0 Hz^{22,12,13} while that of female ranged from 198.0 to 224 Hz^{12,23,24,25}. In our study, there was no significant difference between the age groups of either sex. This demonstrates aging doesn't influence the F_0 before 39 years of age. But the study needs further evaluation of older age groups also. With respect to age, Wang⁴ observed a trend of higher F_0 in males from 20 to 79 years old, but a decreasing trend in F_0 for females from 20 to 79. However, the difference between the age groups was not significant for subjects before 50 years old in both genders. According to Beilamowicz et al.,²⁶ commercially available acoustic analysis programs agreed well, but not perfectly, in their measures of F_0 . However, measures of perturbation in the various analysis packages use different algorithms, provide results in different units, and often yield values for voices that violate the assumption of quasi-periodicity. As a result, poor rank order correlations between programs using similar measures of perturbation were noted. Therefore, if our data are used as a base line for normal adults, we advocate that comparisons be made only with data gained from using the same procedure and analytical software systems.

Conclusion

The average normal perturbation values (jitter and shimmer) observed while producing the vowel /i/ were 0.14% for both the genders. The value for jitter was 0.14% (± 0.02 & 0.04) and 1.6% (0.47 & 0.43) for the shimmer, for respectively for both the sexes. Harmonic to noise ratio also did not reveal any significant difference. Conversely, the value of fundamental frequency was higher in females (246.45Hz \pm 32.72) than the males (170.05 Hz \pm 39.73).

As the aforementioned results the discussions demonstrate, we have developed a reliable normal acoustic data for various parameters of acoustic analysis in different gender and age groups in normal voicing Nepalese adults. Importantly, this standardised data can be used as a convenient reference to evaluate the voices.

Acknowledgement

We would like to thank Mr. Umesh Aryal, department of community medicine in helping to analyse the statistical data.

References

1. Hammarberg B, Fritzell B, Gauffin J et al. Perceptual and acoustic correlates of abnormal voice qualities. *Acta Otolaryngologica*. 1974; 90: 442-51.
2. Muirry T, Singh S, Sargent M, Multidimensional classification of abnormal voice qualities. *Journal of Acoustical Society of America*. 1977; 61: 1630-5.
3. Behram A, Orlikoff R. Instrumentation in voice assessment treatment: What's the Use? *American Journal of Speech- Language Pathology*. 1997; 6(4): 9-16.
4. Wang CC, Huang HT, Voice acoustic analysis of normal Taiwanese adults. *J Chin Med Assoc*. 2004; 67: 179-84.
5. Colton RH, Casper JK, Understanding voice problems: A physiological perspective for diagnosis and treatment. 2nd edition. Baltimore: Lippincott Williams & Wilkins; 2006.
6. Smits I, Ceuppens P, De Bot MS, A comparative study of Acoustic measurements by Means of Dr. Speech and Computerized speech lab. 2005; 19(2): 187-96.
7. Woodson GE, Cannito M. *Voice analysis*. In: Cummings. CW, editors. *Otolaryngology-head and neck surgery*. 3rd edition. St. Louis: Mosby-Year Book;1998.p.1876-90.
8. Titz IR. Towards standards in acoustic analysis of voice. *J Voice*. 1994; 8(1):1-7.
9. Honjo I, Isshiki N, Laryngoscopic and voice characteristics of aged persons. *Arch Otolaryngol*. 1980; 106(3): 149-50.
10. De Bodt MS, Wuyts FL, Heyning PH et al. Test-retest study of the GRBAS scale: influence of experience and profession background on perceptual rating of voice quality. *J of Voice*. 1997; 11(1):74-80.
11. de Felipe, Grillo MHMM, Grechi TH, Standardization of acoustic measure of normal voice patterns. *Rev Bras Otorhinolaringol*. 2006; 72(5): 659-64.
12. Wuyts FL, De Bodt M, Bruckers L et al, Normative data steam. *Acta Otorhinolaryngologica.Belgica*. 1996; 50:331-41.
13. Voster L, Onderzoek van. objectieve parameters ter beoordeling van de normale en de pathologische stem. Unpublished thesis. Katholieke Vlamamse Hogeschool, Antwerpen, o.l.v.S. Peeters & J. Dirckx. 1996.
14. Higgins MB, Saxman JH. A comparison of selected phonatory behaviors of healthy aged and young adults. *J speech Hear Res*. 1991; 34:1000-10.
15. Araujo SA, Grellet M, Periera JC. Normatizacao de medidas acusticas da voz normal. *Rev Bras Otolaryngol*. 2002; 68(4):540-4.
16. Behlau MS, Tosi O. Determinacao da frequencia fundamental e suas variacoes em altura (jitter) e intensidade (shimmer) parafalates do portugueses brasileiro. *Acta AWHO*. 1985; 4(1):5-10.
17. Kent J, Kent R, Rosenbek et al., Quantitative description of dysarthria in women with amyotrophic lateral sclerosis. *J Speech Hear Res*. 1992; 35:723-33.
18. Wolfe V, Cornell R, Palmer C, Acoustic correlates of pathological voice types. *J Speech Hear Res*.1991; 34:509-16.
19. Gelfer MP. Fundamental frequency, Intensity and vowel selection: effects on measures of phonatory stability. *J Speech Hear Res*. 1995; 38:1189-98.
20. Rodrigues S, Behlau M, Pontes P. Proporcão harmônico-ruído: valores para indivíduos adultos brasileiros. *Acta AWHO*. 1994; 13(3): 112-6.
21. Emanuel FW, Smith WF. Pitch effects on vowel roughness and spectral noise. *Journal of Phonetics*. 1974; 2:247-53.
22. Decoster W. Akoestische kenmerken van de ouder wordende stem. Doctoral dissertation. Department Speech Hearing Science, Social Health Science. University of Louven, Leuven, 1998; 29-123.
23. Brown WS Jr, Morris RJ, Michel JF. Vocal jitter in Young adult and aged female voices. *J voice*. 1989; 3:113-9.
24. Lenville SE. Acoustic- perceptual studies of aging voice in women. *J Voice*. 1987;1:44-8.
25. Klingholtz F, Martin F. Quantitative spectral evaluation of shimmer and jitter. *J Speech Hear Res*. 1985;28:169-74.
26. Bielowicz S, Kreimen J, Gerratt BR et al., Comparison of voice analysis systems for perturbation measurement. *J Speech Hear Res*.1996; 39:126-34.