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 (F0) 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 % (±0.47/±0.43) and 25.8dB (±2.62/±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. 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”. 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 data.

Acoustic parameters measure fundamental frequency (F0), Intensity (amplitude), perturbation (jitter and shimmer), harmonic to noise ratio (HNR) and dynamic vocal range. F0 is the vibratory rate of the vocal folds. It can be measured in hertz (Hz). This measurement reflects the physiological limits of the patients voice. During sustained vibration, the vocal fold will exhibit slight variation of F0 and amplitude from cycle to cycle; these phenomena are called frequency perturbation (jitter) and amplitude perturbation (shimmer). 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 voice. HNR is an average ratio of energy of the harmonic components in the range ratio 70-4500Hz to
the inharmonic components energy in the range 1500-
4500Hz. Harmonics are those frequencies that are
integer multiples of the F0. If a sound has a steady F0,
and if all the component frequencies are harmonics,
then each cycle will be identical. Frequencies that are
not integer multiples of F0 are regarded as noise. The
amount of energy conveyed in the F0 and its harmonics,
divided by the energy in noise frequencies, is the
harmonic-to-noise ratio. This parameter correlates with
the perception of vocal roughness. As the degree of
hoarseness increases the noise component increases
and replaces the harmonic structure in the spectrogram.
Normal voices have low level of noise.

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 certification. 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. 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 scale.
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 speech.

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 (F0), 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 (F0). 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 F0 (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 the all
the four parameters revealed no statistical significance
between these two age groups of both the genders either
(P<0.05).

<table>
<thead>
<tr>
<th>Item</th>
<th>Male Average values(N=25)</th>
<th>Mean(SD)</th>
<th>Female Average values(N=25)</th>
<th>Mean(SD)</th>
<th>p-value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jitter</td>
<td>0.14%</td>
<td>0.14(0.02)</td>
<td>0.16%</td>
<td>0.14(0.04)</td>
<td>0.568</td>
</tr>
<tr>
<td>Shimmer</td>
<td>1.6%</td>
<td>1.66(0.47)</td>
<td>1.6%</td>
<td>1.57(0.43)</td>
<td>0.494</td>
</tr>
<tr>
<td>HNR</td>
<td>25.81dB</td>
<td>25.81(2.62)</td>
<td>25.8dB</td>
<td>25.88(2.75)</td>
<td>0.930</td>
</tr>
<tr>
<td>F0</td>
<td>170.05Hz</td>
<td>1.75(32.78)</td>
<td>246.45Hz</td>
<td>2.55(39.73)</td>
<td>0.000</td>
</tr>
</tbody>
</table>

* 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 respectively. Relative jitter values in our study were much lesser than other studies. Wuysts 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 al14(0.85%), Higgins17(0.38%), wuysts13(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 studies15,16. But, our values for relative shimmer were much lower than those observed by smits15, 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% respectively. As compared to other studies we observed relatively lower shimmer values in both the gender groups. 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. Wuysts et al observed a value of 17.5dB and 18.3dB for male and female gender respectively.

Average F0 was included because of its possible effect on the perception of voice quality21. F0 is an acoustic measure that directly reflects the vibrating rate of vocal folds and is expressed in Hertz. The average F0 was 170.05Hz and 246.45 Hz for males and females respectively. The value of F0 was higher in women than the men is due to the vocal folds in men being longer and thicker17. In another acoustic study with the DRS, the average value of F0 in Belgian population was 115.5Hz and 197.0Hz respectively.6 On reviewing the several studies the F0 for males has ranged from 115.0 to 129.0 Hz20,12,13, while that of female ranged from 198.0 to 224 Hz22,23,24,25. In our study, there was no significant difference between the age groups of either sex. This demonstrates aging doesn’t influence the F0 before 39 years of age. But the study needs further evaluation of older age groups also. With respect to age, Wang26 observed a trend of higher F0 in males from 20 to 79 years old, but a decreasing trend in F0 for females from 20 to 79. Howe ever, 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 F0. However, measures of perturbation in the various analysis packages use different algorithms, provide results in different units, and often yield values for voices that violet 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 ± 32.72) than the males (170.05 Hz ±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