Objective Identification of Prepubertal Female Singers and Non-singers by Singing Power Ratio Using *Matlab*

***M. Usha, *Y. V. Geetha, and †Y. S. Darshan,** **Mysuru, India, and †Long Beach, California*

Summary: Background. The field of music is increasingly gaining scope and attracting researchers from varied fields in terms of improvising the art of voice modulation in singing. There has been a lot of competition, and young budding singers are emerging with more talent. This study is aimed to develop software to differentiate a prepubertal voice as that of a singer or a non-singer using an objective tool—singing power ratio (SPR)—as an objective measure to quantify the resonant voice quality.

Method. Recordings of singing and phonation were obtained from 30 singers and 30 non-singer girls (8–10 years). Three professional singers perceptually evaluated all samples using a rating scale and categorized them as singers or non-singers. Using *Matlab*, a program was developed to automatically calculate the SPR of a particular sample and classify it into either of two groups based on the normative values of SPR developed manually.

Results. Positive correlation for SPR of phonation or singing was found between perceptual and manual ratings, and objective values of SPR. Software could automatically give the SPR values for samples that are fed and could further differentiate them as singer or non-singer.

Conclusion. Researchers need not depend on professional singers or musicians for the judgment of voice for research purposes. This software uses an objective tool, which serves as an instrument to judge singing talent using singing and phonation samples of children. Also, it can be used as a first line of judgment in any singing audition process, which could ease the work of professionals.

Key Words: non-singers–objective evaluation–prepubertal singers–singing power ratio–voice quality.

INTRODUCTION

Expressing the words with a set of tune and making the musical sounds with the voice constitutes singing. The art of singing needs fine coordination of all the speech subsystems in an orderly way. It also changes for every different styles and musical tones in a wide variety of range. There is a variety of singing styles, and in India, the two major categories of classical singing are Hindustani and Carnatic, both of which have a great previous history of their own. Each tune constitutes different expressions and shades of emotions in its composition. Good singers possess special breathing habits where the pitch and loudness are coupled together even for their normal speech.[1] In singing, each and every note constitutes particular intended pressure, which is dependent on both pitch and loudness for every particular song. A good singing should have the overall best vocal performance to maintain the rhythm, to follow the music in tune with its respective melody formulae, brilliance or ring in their voice, good breath management, less strain in the voice, among others. All these parameters make singers more evident from inexperienced or non-singers. All these aspects of good singing can be evaluated perceptually by an experienced trained singer. Objective parameters to serve the same purpose are few and often cannot rely on just a single parameter to come to the final conclusion.

In various western studies, the presence of singer's formant has been documented in talented singers.[2] A study reports that the singer's formant lies between 2 and 4 kHz in the power spectrum. Studies with respect to singer's formant vary across the different styles of singing. Few Indian studies report the absence of singer's formant in singers, because in western singing the singers have to project their voice against loud accompaniment or musical instruments.[3] The Carnatic classical singing gives equal importance to both aspects, leading to the absence of singer's formant. However, in spite of many research, no uniform agreement exists for the definition of the singer's formant.

One objective tool that evolved from the concept of singer's formant to classify the voice quality in singers is the singing power ratio (SPR). It is an objective measure that can quantify the resonant quality of singers' voice. The SPR indicates the acoustic characteristics and quality of resonant tuning in the vocal tract of a singing voice.[4] It is the ratio of the highest peak intensity taken between 2–4 kHz and 0–2 kHz frequency bands. This can be analyzed in the sustained phonation or a singing sample.

Omori et al[4] reported that SPR is greater in singers than in non-singers in both men and women, but there was no significant difference between professional and nonprofessional male and female singers. Also, the SPR was greater for sung vowel /a/ than for spoken vowel. This study used formant ratio method for calculating the SPR.[5] Similarly, the SPR calculated using the formant difference method reported significant difference of SPR in talented and non-talented singers, where the SPR was 8 decibels (dB) lesser in the talented group than in the non-talented group.[3] An Indian study showed no statistical difference between SPR values of trained Carnatic and untrained singers, and also the absence of the singer's formant.[6] SPR values could differentiate three levels (junior, senior, and vidwath) of Carnatic singers. All these studies are on adult singers.

These days, in majority of the cases, singing is taken up as a serious hobby right from childhood, and mastering is the same as in a professional singer. There are many young budding singers

Accepted for publication June 28, 2016.

From the *All India Institute of Speech and Hearing, Manasagangothri, Mysuru, Karnataka, India; and the †California State University, Long Beach, California.

Address correspondence and reprint requests to M. Usha, All India Institute of Speech and Hearing, Manasagangothri, Mysuru, Karnataka, India. E-mail: usha_yeshu@yahoo.com

Journal of Voice, Vol. 31, No. 2, pp. 157–160

0892-1997

http://dx.doi.org/10.1016/j.jvoice.2016.06.016

who possess good voice quality and can be trained well as good singers later. There are many studies that report that there is a difference in the singing voice quality between children and adults. The description of children's voice has been described perceptually as being clearer, non-fuzzy, echoey, and purer.[7] Also, listeners who are not trained in music have commented on the characteristics of children's voice as having clean white tone, lightness, and clarity based on perceptual judgments. There are few Indian studies focusing on prepubertal singers and non-singers.[8] Using formant power difference method, a study on prepubertal age range of 8–12 years found the SPR differences among three groups, that is, trained classical singers, untrained talented singers, and non-talented singers, using a limited number of participants in each group: 14 children under trained singers category, 12 children under untrained talented singers category, and 18 under non-singer category.[9] The same study, continued in prepubertal female singers with 30 participants in each of the three groups, found the normative for classifying singers as trained, talented, and non-talented singers using the SPR values that correlated with the perceptual values for the same samples of phonation and singing.

NEED FOR THE STUDY

The normative values for classifying prepubertal children's voice as trained, untrained talented, and non-singers have been developed from a study based on the values of SPR.[9] But in that study, the calculation of SPR was manual and it requires a lot of time and effort for the same. It is not always possible to rely on experienced professional singers to evaluate the samples perceptually. It would be helpful if software is developed for the same purpose, which will ease the work of calculating SPR manually and reduce the calculation errors.

The main aim of the present study was to develop software to automatically classify the prepubertal female voice and singing samples as that of singers or non-singers based on the norms obtained in the previous study.

Objective

- To develop software in *Matlab* using fast Fourier transform (FFT), which can calculate SPR using the program algorithm and classify the phonation and singing samples of children's voice using the normative
- To check the validation of the software by running recorded samples of phonation and singing

METHOD

Samples of phonation and singing were collected from two groups of prepubertal children (age range of 8–10 years) where group 1 consisted of 30 children who were trained in Carnatic music for a minimum of 2 years and group 2 consisted of 30 non-singers without any formal training. For singing sample "Lambodara lakumikara" song from Carnatic music, which is in Malahari raga, Roopaka thala was used from all the participants. Non-singers were given sufficient time to practice the song along with audiotaped sample and lyrics.

Matlab is a useful tool that is used to analyze speech signals that are read in dot wav format. The frequency of any unknown audio signal can be calculated using FFT in *Matlab*. The following functions are used to analyze the samples:

- Wavread: reads speech signal
- Windowsize: defines window function of transformation
- Wavplay: produces speech signal after transformation
- Length: defines length of speech to be processed by transforming principle
- Audio recorder: prepares audio recording real time

Using these samples, further software was developed. First *Matlab* software should be run; then, the software will accept a real-time or recorded audio sample in the format of .wav from the microphone, which has been further analyzed to find out maximum peak or amplitude of a speech in a given interval of frequency range between 0–2 kHz and 2–4 kHz by using FFT analysis and the above-listed functions. The program will automatically give the highest amplitude of 0–2 kHz and 2–4 kHz frequency, in a unit of dB; the conversion of dB will be made by the program itself.

$$A = 20 * \log_{10}(y)$$

Further program was done to calculate the SPR automatically by subtracting the values of highest peak at 0–2 kHz from the values of highest peak at 2–4 kHz. Then, to classify the sample as that of a singer or a non-singer, the software used the formula that was obtained after running all the samples and setting the criterion value for differentiating the singing samples as singers' or non-singers'. The formula used for obtaining decision regarding singing sample is:

$$\text{Decision} = ((\text{fspe1} * \text{SPR}) + (\text{fspe2} * \text{SPR}))\,/\,(\text{fspe2} * \text{fspe1} * \text{SPR})$$

fSpe1—first highest singing power energy/peak from 0 to 2 kHz
fSpe 2—second highest singing power energy/peak from 2 to 4 kHz

For phonation samples, the normative values of SPR developed by Usha and Geetha[9] were considered, which were also used to cross-check all the data and validate the results. That is, all the recorded samples of classical prepubertal singers and non-singers were given to three professional singers for the perceptual evaluation using a rating scale with eight parameters to rate and to categorize them as singers or non-singers. The eight parameters considered were an overall rating of the aesthetic and technical quality of singing voice (overall vocal performance), maintenance of raga (melodic formulae), rhythmic scale/thala (rhythm follow), brilliance of tone (ring), singing in tune (pitch accuracy), efficient breath management (breath management), ability to sing freely throughout the pitch and dynamic range without inappropriate change in voice quality (evenness throughout the range), and voice quality that gives impression of excessive vocal effort (strain). All the recorded samples were then run in