VOCAL FOLD BEHAVIOR IN CARNATIC SINGERS

Selvi (T)
Register No- 06SLP016

A Dissertation Submitted in Part Fulfillment of
Final year M.Sc (Speech - Language Pathology)
University of Mysore, Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING
NAIMISHAM CAMPUS, MANASAGANGOTHRI
MYSORE-570006

Dedicated to the Almighty,

My dear Appa & Amma

CERTIFICATE

This is to certify that this dissertation entitled ***"Vocal fold behavior in Carnatic singers"*** is the bonafide work submitted in part fulfillment for the degree of Master of Science (Speech-Language Pathology) of the student (Registration No. 06SLP016). This has been carried out under the guidance of a faculty of this institute and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

V. Basavaraj

Dr. Vijayalakshmi Basavaraj
Director
All India Institute of Speech and Hearing
Manasagangothri
Mysore-570 006.

April, 2008

CERTIFICATE

This is to certify that the dissertation entitled ***"Vocal fold behavior in Carnatic singers"*** is a bonafide work in part fulfillment for the degree of master of (Speech-Language pathology) of the student (Registration N0.06SLP016). This has been prepared carried out under my supervision and guidance. It is also certified that this has not been submitted earlier in any other University for the award of any Diploma or Degree.

Guide

Mysore

April, 2008

(Dr. S.R.Savithri)
Professor of Speech Sciences
Department of Speech-Language Sciences
All India Institute of Speech and Hearing
Mysore-570006.

DECLARATION

This is to certify that this dissertation entitled ***"Vocal fold behavior in Carnatic singers"*** is the result of my own study under the guidance of Dr. S.R.Savithri, Professor of Speech Sciences, All India Institute of Speech and Hearing, Mysore, and has not been submitted in any other university for the award of any diploma or degree.

Mysore

May, 2007 Register No. 06SLP016

ACKNOWLEDGEMENTS

My dearest loving **Almighty god,** thank you for blessing me immensely and holding me in each and every moment of life.

My heartfelt gratitude and thanks to my guide **Prof. S.R.Savithri** without whom this dissertation wouldn't have been possible. Ma'am I thank you with all my heart for all your support, guidance and perfection in making my dissertation complete.

I thank **Dr. Vijayalakshmi Basavaraj,** Director, All India Institute of Speech and Hearing, for permitting me to undertake this study.

I thank **Jayakumar sir** for his timely help at various stages of the study from the beginning. Thank you very much sir, for all the help you have done to complete my study.

Dear **Appa and Amma,** words are not enough to express my love and affection for you. I'm very much blessed to have parents like you and I'm very thankful to you for all the love and encouragement you have given me right from my childhood.

Dearest **Jenu and Ani,** you are the best sisters one can ever have. You were always there during my times of needs and trouble and I thank you both for all the love, care and affection that you have given me to reach my stars.

Dearest **Athaan,** I'm very thankful to you for all the care and affection, and specially helping me out during my data collection.

Dear **Aunty & Uncle** thanks a lot for being there, a constant source of joy for all our days.

Dearest **Paulineakka, Nimmyakka, Gendral, kuttyma, Appu, Gemie, Priya, Joshua n Suji,** have cousins like you is a blessing so special and rare.

Archana. . . . .You have made me rich by entering my life, being my light and warmth through tears and smiles, laughter and heartaches. Thanks for bringing happiness into my life, giving me valuable advice and suggestions whenever I needed them. I found a true friend when I found you.

Anu, Rubi & Ramia ... You guys mean a lot to me, being a part of my life. Life at AIISH would have been an incomplete jigsaw puzzle without u. Thanks for being there during thick n thin times. Rubi my study partner, you are a wonderful friend. The word thanks is too less for all the help and support u gave me when i was in need of it.

Balaji & Poorna ....your friendship is very special indeed. Thanks for always being there.

Hema thanks a lot for being a helping hand and baring my mood swings during our study times.

Dear **Maggie, Ceci, Kavitha n Rajani**.... Great to have friends like you and times & fun we had together will never be forgotten.

Prasi, Neha, Bindu, Dhanya, Simsom, Swati, Leah, Anjana, Karthi & **Vicky...**you guys have given me a lot of memories to treasure, all the fun we had would be cherished forever.

I thank my seniors **Supraja, Anjali** & **Savitha** . . . . . .Supraja, I thank you specially for helping me a lot during my study and special memories of our dance practice would be remembered and cherished always.

I would like to thank the **Symbionts** and my other classmates for giving me good times and memories to ponder.

I extend my gratitude to the Carnatic singers of the Chennai Music Academy, who have contributed for the study.

I would like to thank **Mr. Shivappa** & **Co.,** for their kind cooperation throughout.

Table of contents

List of tables

CHAPTER I

INTRODUCTION

Human beings experience a wide spectrum of emotions in our life, which can be expressed through music. Music is not only a set of sounds; it evokes emotions, subjectively perceived by listeners. Extracting information on emotions from music is difficult for many reasons. First of all, music itself is a subjective trait, related to culture. Singing is considered as sensory motor phenomenon that requires particular balanced skills (Bunch, 1982). A sensitive performer achieves singing as an art when these skills are developed. Singing is such a human and moving act both for those who do it and for those who listen to it. Hence singers are considered artists in the truest sense as they combine concept, melody and text and stage movement making it all seem effortless and get capable of winning audience appreciation.

Music can be defined in various ways, for instance, as an artistic form of auditory communication incorporating instrumental or vocal tones in a structured and continuous manner, or as the art of combining sounds of voices or instruments to achieve beauty of form and expression of emotion. Therefore, music is inseparably related to emotions. Music, which is composed of seven distinct tones/notes, is actually an expression of human emotions. Music can express most of human emotions like joy, ecstasy, anger and also sorrows, pathos, and agony. A variation in melody and rhythm lead to expression of varieties of emotions.

Indian music is distinguished from other musical systems in terms of its origin and its implication. Music is considered as a gift of god and part of fourfold goals of life. This sociocentric view has its widespread effect. Till today classical music in all its forms is used to worship, praise/hail the glory of formless, so much so that even ragas and notes have their own god. This basic conceptual material is from Vedas. No other musical system has such a definite and strong background as Indian music (Satyanarayana, 1983) Carnatic music is one of the two styles of Indian classical music, the other being Hindustani music.

The Carnatic system is the art form of Southern India and Hindustani is the art form of Northern India. Carnatic music is a very complex system of music that requires much thought, both artistically and technically. The basis of Carnatic music is the system of ra:gas (melodic scales) and ta:las (rhythmic cycles). There are seven rhythmic cycles and 72 fundamental ra:gas. All other ra:gas are considered to have stemmed from these. An elaborate scheme exists for identifying these scales, known as the 72 Melakarta ra:gas. The notes of Carnatic music are "sa-ri-ga-ma-pa-da-ni". These are abbreviations of the real names which are s.adja, r.sabha, ga:ndha:ra, madhyama, pancama,dhaivata and nis a:da. Combination of tones/notes is known as ra:ga.. The melodies contained in each of these ra:ga reflect various moods according to different times during the day. Classical singing is an art that a person develops after rigorous training and practice. Apart from lyrics used, vocal characteristics can express differences in emotions of the song rendered by trained singers. Any discussion of the emotional response to music is faced at the very outset with the fact that very little is known about this response and its relation to stimulus. Evidence that it exists at all is largely based upon the introspective reports of the listeners and the testimony of the composers, performers and critics. Other evidence of the existence of the emotional responses to music is based upon the behavior of performers and audiences and upon the physiological changes that accompany musical perception.

Much confusion has resulted from the failure to distinguish between emotion felt or affect and mood. For several reasons the verbalizations of emotions, particularly those evoked by music are usually deceptive and misleading. Few psychologists dealing with music have been as accurate on this point as Weld, who notes that: 'Emotion is temporary and evanescent; the mood is relatively permanent and stable. A clear distinction should be made between the emotions felt by the composer, listener or critic - the emotional response itself - and the emotional state denoted by different aspects of the musical stimulus. The depiction of musical moods in conjunction with conventional melodic or harmonic formulas, perhaps specified by the presence of a text, can become signs which designate human emotional states. Motives of grief or joy, anger or sorrow found in the works of baroque composers or the affective and moral qualities attributed to special modes or ragas in Arabian or Indian music are examples of such conventional denotative signs.

Sundberg (1990) stated that the singers must gain control over all perceptually relevant voice parameters so that they do not change by accidents and do not signal an unintended boundary. Technical control of voice components is a pre-requisite for artistic expressions. A professional singer must be able to produce the optimal vocal product at the same time preserve the mechanism producing it.

It is a well known fact that vocal gestures relate to the physiology of larynx and voice quality differences produced by laryngeal features are one of the fundamental factors differentiating vocal emotional expression. The ground work for studying voice quality changes in the vocal expression of emotion has been laid in previous research, including emotion induction, description and classification, as well as robust inverse filtering and glottal flow parameterization methods. However, previous analyses have either used a very limited number of test subjects or they have analyzed voce qualities or stress styles instead of a well defined emotion. Till date, no study has been specifically devoted to a detailed exploration of study of variations in vocal fold behavior during expression of emotions in classical singing. Therefore, this study analyzed the vocal fold behavior during expression of various emotions in Carnatic singers. Specifically, The study analyzed the various measures depicting vocal fold behavior of Carnatic singers at high frequency namely Pitch period [TO], open quotient [OQ], speed quotient (SQ), leakage quotient [LQ], excitation ratio [EI/EE], dynamic leakage [AR], harmonics [HO], harmonic ratio (H1-H0) and spectral roll-off during the expression of emotions joy, neutral and sorrow and investigated variation in these glottal parameters vary across gender in Carnatic singers.

CHAPTER II

REVIEW OF LITERATURE

According to Scherer (2000) emotions are processes of events affecting several psychological components of an organism, namely physiological arousal, motor expression and subjective feeling. The observable changes in physiological arousal and motor expressions of emotions may be considered to possess an inherently communicative role. Within the last few years, research on vocal expression of emotion has advanced rapidly, due to psychological as well as phonetic and speech processing research. However, despite the progress in the study of vocal expression of emotion, knowledge of the topic is still far from conclusive.

In speech, emotion is communicated by a combination of features at all three principal levels of speech abstraction: suprasegmental, segmental and intrasegmental (Murray & Arnott, 1983). All these three levels can be considered to consist of two components namely verbal (words) and vocal (intonation, voice quality and intensity). At suprasegmental and segmental levels, this information includes fundamental frequency (FO), sound pressure level (SPL) and duration patterns as well as variations in formant configurations. While suprasegmental patterns are undoubtedly important for vocal expression of emotion, Pollack (1980) noted that emotion can be recognized in segments of speech as short as 60 ms. On this intrasegmental level, vocal expression of emotion is performed with various voice quality adjustments generated by manipulation of voice source.

Multiple studies on acoustic parameters of emotional speech were in both suprasegmental, segmental as well as intrasegmental levels. The earliest being done in 1930s, after the invention of phonophotographic oscillography. Skinner (1935) investigated F0 and intensity patterns in the expression of induced happiness and sadness.

Fairbanks & Pronovost (1938, 1939), Fairbanks (1940) and Fairbanks & Hoaglin (1941) explored F0, intensity and durational patterns of a speech passage elicited in five different dramatized expressions. They found differences in the acoustic parameters across the emotions. Later research generally confirms the findings of their studies.

In 1962, Kaiser studied how single vowels could express affects. Six acted emotions were analyzed using oscillograms, spectrograms and F0 and intensity patterns. The F0 and timbre were found to have the strongest effect on the perception of emotions. Suprasegmental features mainly F0 contours, were analyzed for real emotional recordings acquired from radio traffic between a civilian pilot and a control tower operator (Williams & Stevens, 1969). The same features were later analyzed for simulated emotions yielding results that correlated with the previous one. The FO ranges and contours were found to vary according to expressed emotions, with a definite increase of F0 range in distress.

Banse & Scherer (1996) studied emotion portrayals by professional actors in 14 emotions varying in intensity and valence. Multiple data including different F0, energy, speech rate, and spectral parameters were complied and analyzed. Rather good performance in statistical emotion classification was reported. But in this study, although some information on the voice quality was encoded in the spectral parameters, an explicit analysis of the voice quality or voice source parameters was not attempted.

Laukkanen, Vilkman & Alku (1997) conducted a study coupling variations of glottal flow signal and emotions. Nonsense utterances were produced in five simulated emotional states by three subjects. In addition to regular F0 and SPL measures, glottal flow was estimated and parameterized using the speed quotient and open quotient. The emotions were found to differ from one to another in the F0 and SPL results.

The variation of open quotient with fundamental frequency was also been explored (Childers, Naik, Larar, Krishnamurthy & Moore, 1983; Hanson, Gerratt& Berke, 1990). Healthy normal male and female adults age ranging from 21-38 years served as subjects in these studies. In case of male speakers, research did not show any relationship between open quotient and fundamental frequency. In the case of female speakers, an increase of open quotient with an increase of fundamental frequency was found by Kitzin & Sonneson (1974) using photoglottography.

The role of voice quality in communication of emotions and moods and attitudes was explored by Christer & Chaisade (2003). Listeners' reactions to an utterance synthesized with seven different voice qualities were elicited in terms of pairs of opposing affective attributes. The voice qualities included harsh voice, tense voice, modal voice, breathy voice, whispery voice, creaky voice and lax-creaky voice. These were inverse filtered and synthesized using a formant synthesizer, and the voice source parameter settings were guided by prior analytic studies as well as auditory judgements. Results offer support for some past observations on the association of voice quality and affect, and suggested a number of refinements in some cases. Listeners' ratings further suggested that these qualities are considerably more effective in signaling milder affective states than the strong emotions. It is clear that there is no one-to-one mapping between voice quality and affect: rather a given quality tends to be associated with a cluster of affective attributes.

Airas & Alku (2006) analyzed the glottal flow of vowel segments in continuous speech during various emotions using inverse filtering procedure. 5 males and 4 females, who were professional stage actors, native speakers of Finnish in the age range of 26-45 years, served as subjects for the study. The subjects' task was to recite a text passage containing 83 words of Finnish prose in a give emotion. The emotions taken were neutral, sadness, joy, anger and tenderness. Two glottal parameters *amplitude quotient* - voice source parameter that represents the ratio of the glottal flow cycle amplitude to maximum negative peak of its first derivative and CLQ, closing quotient which is defined as the ratio between the duration of the glottis closing phase and the fundamental period. The results showed that there was significant difference in the gender and emotions for the glottal parameters analyzed.

Singers use their voices in special ways that may be different from nonsingers and untrained singers. In the past, some efforts have been made to understand the vocal fold behaviour in singers. Sundberg (1989) researched on singer's breathing, phonation and articulatory patterns during singing and comparisons were made with typical speech patterns. He reported that breathing aspects of singers are special. In normal speech, loudness and pitch are coupled together, both being dependent on subglottal pressure. In singing, each note requires its own particular pressure and this pressure depends on both the intended pitch and intended loudness of the note. A failure to produce the required pressure results in singing out of tune. Therefore singers are clearly special with regard to breathing. It was also found that singers' phonatory habits are special in at least three different aspects. First, mean pitch and mean loudness are interrelated in speech so that a change in loudness is normally associated with a change in pitch. In singing, on the other hand, pitch and loudness are independent. Second, changes in pitch and loudness are typically associated with change of mode of phonation in speech. In singing, neither pitch nor loudness changes should automatically cause audible, automatic changes of mode of phonation; such changes should be reserved for musical expression. Finally, Singers must tailor the activity of pitch raising cricothyroid muscles with regard not only to loudness but also to lung volume. Thus singers are special with regard to phonation. Singers are also special with respect to formant frequencies. In normal speech, formant frequencies are used merely to pronounce different vowels and consonants. Also, when pitch varies, larynx height normally varies in speech and a change in larynx position results in changes of the air cavity volumes in the low pharynx.

Henrich, d'Alessandro, Castelengo & Doval (2004) studied glottal open quotient measurements in case of singing voice production. They explored the relationship between open quotient and laryngeal mechanisms, vocal intensity and fundamental frequency. Eighteen professionally trained singers served as subjects for the study. Since male and female phonation differs in respect to vocal fold vibratory properties, a distinction was made between two different glottal configurations, i.e.

mechanism 1, related to chest, modal and male head register and mechanism 2, related to falsetto for male and head register for female. It was found that open quotient depended on laryngeal mechanisms, which ranged between 0.3 to 0.8 in mechanism 1 and between 0.5 to 0.95 in mechanism 2. Open phase was longer in mechanism 1 shorter in mechanism 2.

Henrich, d'Alessandro, Castelengo & Doval (2005) explored the perceptual relevance of the variations of glottal flow parameters and to what extent a small variation can be detected. Just Noticeable Differences (JNDs) was measured for three values of open quotient (0.4, 0.6, and 0.8) and two values of asymmetry coefficient (2/3 and 0.8), and the effect of changes of vowel, pitch, vibrato, and amplitude parameters were tested. Two main groups of subjects were analyzed: a group of 20 untrained subjects and a group of 10 trained subjects. The results showed that the JND for open quotient is highly dependent on the target value: an increase of the JND is noticed when the open quotient target value is increased. The relative JND is constant: Oq/Oq = 14% for the untrained and 10% for the trained. In the same way, the JND for asymmetry coefficient is also slightly dependent on the target value—an increase of the asymmetry coefficient value leads to a decrease of the JND. The results showed that there was no effect from the selected vowel or frequency (two values have been tested), but that the addition of a vibrato had a small effect on the JND of open quotient. The choice of an amplitude parameter also had a great effect on the JND of open quotient.

Petterson & Westgaard (2006) studied the activity patterns of neck muscles during classical singing, particularly muscle usage during inhalation and phonation and the relationship to changes in pitch and vocal loudness. Five professional opera singers (3 males and 2 females) participated in the study. Singing tasks were performed with variation in vocal loudness and pitch. Surface electromyographic activity was recorded from the muscles of anterior and posterior neck regions. It was found that the neck muscles showed correlated activity during inhalation and phonation in classical singers and activity of neck muscles were markedly elevated when singing in the highest pitch in classical singers.

Judith & Rieger (2008) examined the physiology of pitch change in terms of laryngeal and respiratory mechanisms in professional singers. Nine female professional singers were recruited to participate in the study. Videoendoscopic recordings of the participants producing one-octave ascending and descending scales were used as a basis to apply a ratio measurement of vocal fold length for each note produced on the scale. Simultaneous respiratory data using Respitrace were also collected. Questionnaires (Voice Handicap Index and Vocal Questionnaire) were used to obtain information about psychosocial aspects related to voice use. Two vocal fold lengthening patterns *(static* and *dynamic)* were observed with pitch change. Participants exhibiting a static pattern of vocal fold lengthening had fewer years of vocal training, exhibited a more variable use of vital capacity, and also began the singing task at a higher position in their vital capacity. The reverse was true for participants exhibiting a dynamic vocal fold pattern. These preliminary data indicate that the pattern of vocal fold lengthening exhibited by singers may be related to the number of years of training possessed. Furthermore, the data indicates that stability in one subsystem may result in variability in another, as shown by the interaction between the vocal fold and respiratory patterns.

The review indicates that some glottal parameters have been studied in speech and singing under different emotions. However, till date no study has explored vocal fold behaviour in Carnatic singing at high pitches. Therefore, the present study analyzed the vocal fold behavour during expression of various emotions in Carnatic classical singing at high pitches.

CHAPTER III

METHOD

Subjects: Sixteen Carnatic singers, eight males and eight females, in the age range of 18-28 years participated in the study. The subjects had passed the junior music examination and had no history of any laryngeal pathology, vocal abuse or vocal misuse behavior and the subjects had been practicing regularly.

Procedure: Subjects were tested individually. Initially the subjects were instructed to sing sapasa. Carnatic ra:gas depicting three emotions - joy, sorrow and neutral - were selected. Ra:gas taken by the Carnatic singers were /kalya:ni/, /mukhari/, /punnagavara:l.i/, /bilahari/, /a:rabhi/, /ca:ruke:si/, /cittraranjini/, /a"nandabhairavi/, /karaharapriya/, /hindola/, /ka:phi/, /re:vagupti/, /sahana/, /sanmugapriya/, syamala:ngi/, /kama:c/, /raghupriya/, /cakrava:ka/, and /sankara:bharana/.. They were instructed to sing the kriti, till they were satisfied that they have expressed the intended emotion. All the singings were audio-recorded using a Digital mini-disc recorder and were transferred directly onto the computer at a sampling frequency of 16 kHz. Two musicians judged the emotions of these ra:ga and only those that are judged 100% to express intended emotions were considered for further acoustic analysis. VAGHMI software (Voice & Speech Systems, Bangalore) was used for the vocal fold behavior analyses. Each singer's base frequency was extracted from sa of sapasa using F0 program of VAGHMI. Vowel /a/ segment at highest frequency above the singer's base frequency in each emotion was selected and saved as separate files. These were then subjected to inverse filtering.

Speech is the product of source and filter (P = S * T). In order to obtain the source, the filter effect should be nullified. Indirect measurements of the glottal waveform involve inverse filtering. This technique is effectively a reversal of the speech production process. The speech signal is passed through a filter whose transfer function is the inverse of the supraglottal transfer function. In principle, this yields the voice source in its prefiltered form, as the filtering effect of the voice is cancelled.

This method enables the researcher to estimate the glottal flow signal by using the speech signal. The acoustic signal is recorded using a pressure-sensitive microphone, which corresponds only to the acoustic pressure and not flow. Thus inverse filtering gives the derivative of glottal flow called voice source pulses. The relationship between the glottal flow pulses and voice source pulses is illustrated for a typical case in figure 1.

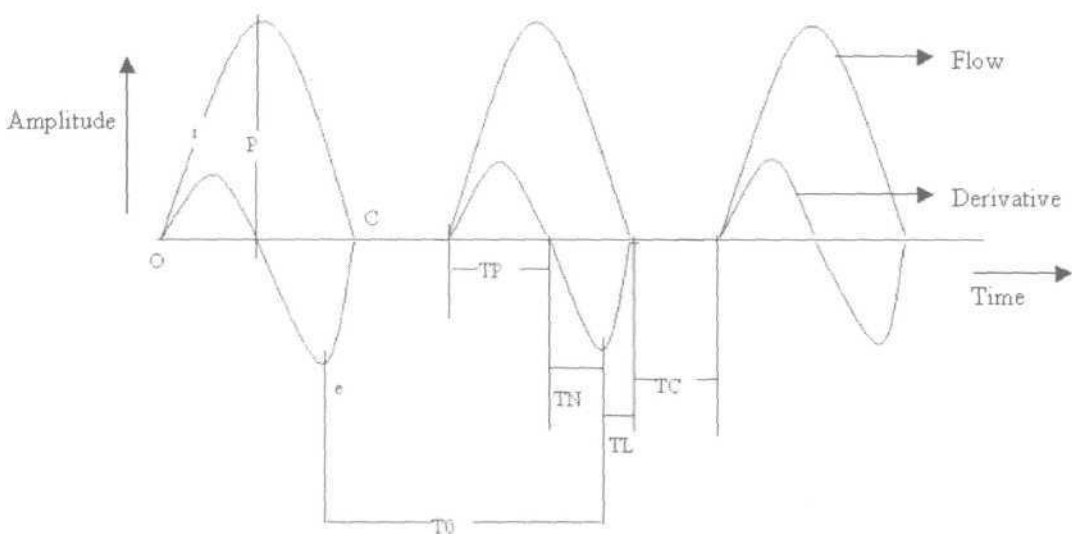

Figure 1: Illustration of inverse filtering.

Instants - O: Glottal onset, i: Inflection, p: Peak flow, C: Closure, e:epoch
Intervals- TO: Pitch period (epoch to epoch), TP :Opening interval,
TN: Closing interval, TL: Leakage interval, TC: Closed interval

To obtain the true glottal flow from the speech pressure wave, the filtering effect of the radiation at the lips needs to be cancelled. The radiation characteristics can be relatively accurately approximated by a first order differentiation. The spectral consequence of this differentiation is a relative boosting of higher frequencies by 6 dB per octave. This effect can be easily cancelled by a simple integration of the signal, as this is the inverse of differentiation. If the effect of the lip radiation is not cancelled, the output of the inverse filter will correspond to the differentiated glottal flow, also referred to as the glottal flow derivative.

Inverse filtering based on the speech pressure waveform can yield detailed temporal and spectral information. However, the recording equipment and room are critical, and shortcomings in either condition can lead to inaccurate results. The recording equipment must preserve the phase characteristics of the signal even at very low frequencies, which effectively means that a digital or FM tape recorders is needed unless recording is done straight to computer.

In Figure 1, the glottal onset is at '0' and the flow begins to increase. Both the flow and the derivative signals pass through the zero at this point. At point 'p' the glottal flow is maximum. The maximum negative amplitude in the derivative is called 'Epoch', corresponds to the instant of glottal closure but the folds are not completely closed. After the epochal instant the flow gradually decreases to zero. The point 'c' corresponds to the closure point. Here both the flow and derivative are zero.

The perceived loudness as well as sound pressure level depends on the maximum negative amplitude in the derivative and not on the glottal flow. If the peak of the flow increases then the amplitude also increases. For the same peak flow if the closure is abrupt then the amplitude at Epoch is larger resulting in higher intensity. Thus amplitude at epoch is related to physiological effort. 'En' is maximum negative amplitude at epoch in the derivative. 'Ep' is the maximum positive amplitude in the derivative and the perceived loudness as well as the sound pressure level (intensity in dB) depends on the maximum negative amplitude in the derivative (amplitude at epoch) and not on the peak glottal flow. The three important instants glottal onset, glottal flow peak and glottal closure correspond to the zero crossing in the derivative of the flow, which is the inverse filtered signal. For analysis, specific aspects of the waveform, i.e., those aspects that are thought to be acoustically and perceptually important, and which can be more readily related to the underlying physiological events needs to be measured. So the acoustic parameters open quotient, speed quotient, leakage quotient, EI/EE, spectral roll-off, Ep/En, pitch period (T0), and Harmonic ratio (H1-H0) were be extracted.

The interval between the 'onset' and 'peak' is called the opening phase and is denoted in TP. The interval between the 'peak' and 'closure' corresponds to the 'closing phase' and is denoted by TN + TL. TN corresponds to the interval from peak to the instant of closure, epoch. TL corresponds to the final gesture of the closure. Total interval for which the glottis is open is therefore is, TP+TN+TL. Open quotient = (TP+TN+TL) / T0, it is a frequently used parameter and in terms of the source spectrum, it mainly controls the amplitude. Speed quotient = (TP/TN). Speed quotient signifies the relative interval taken for opening compared to the interval taken for closing. For example speed quotient of 1 means that opening and closing phases are of equal duration. Usually speed quotient is about 2, i.e. the opening phase is longer compared to closing phase. For the same peak flow if TN is shorter then the amplitude at epoch (EN) is larger.

Opening gesture is usually gradual and closing gesture is more abrupt. The time taken for the voice source signal to return from epoch to the baseline, called the return phase is an indication of the closing gesture. When there is an abrupt glottal closure TL - leakage interval will be zero. In the presence of glottal leakage or abduction, TL will be non-zero. The ratio of TL to T0 is called the leakage quotient, LQ.

EE, excitation strength is the negative amplitude at the time point of maximum discontinuity of the differentiated flow. It normally occurs at the maximum slope of the falling branch of the glottal pulse, which typically precedes full closure. At the production level it is determined by the speed of closure of the vocal folds and by the airflow through them. EI/EE - the mean value of the voice source is zero implying that positive area equals the negative area in each cycle. When leakage TL is small, an approximate relationship between EE/EI= TP/TN=SQ. AR- dynamic leakage is the residual flow during the return phase, which occurs from the time of excitation to the time of complete closure.

T0- pitch period is measured between two successive epochs, refers to the time taken for one glottal vibration. Spectral roll-off indicates the smoothness of the glottal closure or the change in the spectral level over an octave change in the frequency. Ep/En is the ratio of energy at peak and epoch and is an indirect measure for the efficiency of voice. Harmonic ratio (H1 - H0) is the ratio of energy at first harmonic and fundamental frequency.

Analysis: The recorded voice samples were transferred on to the computer at a sampling frequency of 16 kHz and using cool edit pro software, the vowel segment /a/ with highest frequency above the singers' base frequency in each emotion was extracted and saved as separate files. These files were then subjected to inverse filtering and appropriate glottal markings were made. A minimum of 5 cycles in each singing was subjected to inverse filtering.

The glottal parameters extracted include T0- pitch period, OQ- open quotient, SQ- speed quotient, LQ- leakage quotient, EI/EE- excitation ratio, AR-dynamic leakage, H0- harmonics, HI-H0 harmonic ratio and spectral roll-off.

Statistical analysis: The data obtained was tabulated and subjected to statistical analysis. SPSS software (version 10) was used to perform the analysis. An independent t-test was done to analyze the variation of glottal parameters across genders within each emotion and paired t-test was done for pair wise comparison of each emotion for each parameter irrespective of gender. Also, mixed ANOVA was used to study the interaction of gender * emotion.

CHAPTER IV

RESULTS

1. Variations across gender

Results of the independent t-test showed significant difference between gender on OQ (Joy, Sorrow $p<0.05$), LQ (Sorrow $p<0.05$), AR (Sorrow $p<0.05$), H0 (sorrow $p<0.05$), and spectral roll-off (Sorrow $p<0.05$). Table 1 shows the mean and SD values of all parameters across emotions and gender and table 2 shows t-values for all parameters.

	JOY				NEUTRAL				SORROW			
	Males		Females		Males		Females		Males		Females	
Parameters	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TO	4.35	.794	3.49	.892	4.39	1.15	3.55	.91	4.21	1.16	3.45	.873
OQ	1.00	2.31	.9375	.102	1.16	.319	1.01	4.17	1.02	.388	.945	.123
SQ	1.44	.388	1.22	.737	1.25	.489	1.30	.472	1.36	.422	1.18	.349
LQ	.247	.183	.312	.261	.215	.122	.241	.126	.385	.303	.188	.145
EI/EE	.855	.379	.673	.248	.531	.207	.596	.221	.555	.244	.813	.272
AR	1.26	.467	1.24	.813	2.01	.897	1.47	.593	1.93	.901	1.30	.290
HO	104.1	4.68	97.67	4.84	100.5	8.73	97.80	5.16	102.1	6.88	96.34	2.44
H1-H0	4.61	3.94	7.50	4.06	5.49	3.34	9.81	5.28	5.87	3.83	8.20	3.75
Roll-Off	7.37	3.28	5.88	1.69	6.65	2.82	6.36	1.12	6.01	1.02	6.53	2.59

Table 1. Mean and SD values of glottal parameters across gender and emotions.

	t value	df	Significance
T0 Joy	2.033	14	.982
T0 Neutral	1.608	14	.495
T0 Sorrow	1.46	14	.217
OQ Joy	**1.75**	**14**	**.005***
OQ Neutral	1.28	14	.071
OQ Sorrow	**1.762**	**14**	**.029***
SQJoy	.733	14	.176
SQ Neutral	.182	14	.858
SQ Sorrow	.947	14	.238
LQJoy	.576	14	.204
LQ Neutral	.423	14	.456
LQ Sorrow	**1.64**	**14**	**.026***
EI/EE Joy	1.12	14	.167
EI/EE Neutral	.605	14	.677
EI/EE Sorrow	1.998	14	.879
AR Joy	.072	14	.344
AR Neutral	1.41	14	.515
AR Sorrow	**1.88**	**14**	**.026***
HO Joy	2.72	14	.645
HO Neutral	.760	14	.122
HO Sorrow	**2.23**	**14**	**.008***
H1 H0 Joy	1.44	14	.948
H1 H0 Neutral	1.95	14	.238
H1 H Sorrow	1.22	14	.736
Roll-off Joy	1.14	14	.531
Roll-off Neutral	**.266**	**14**	**.044***
Roll-off Sorrow	**.528**	**14**	**.050***

Table 2: t-values and degrees of freedom for glottal parameters across gender within each emotion.

OQ in joy and neutral emotions were higher in males compared to females. LQ in sorrow, AR in sorrow and HO in sorrow were higher in males compared to females. However, spectral roll-off in sorrow was lower in males compared to females.

2. Variations across emotions:

Results of the t-test indicate significant differences across emotions on EI/EE (joy vs. neutral, neutral vs. sorrow, $p<0.05$). EI/EE was higher in joy and sorrow compared to neutral emotion. Table 3 shows the mean and SD of parameters and table 4 shows the t-values.

Parameters	Joy		Neutral		Sorrow	
	Mean	SD	Mean	SD	Mean	SD
TO	3.9244	.9288	3.9725	1.0954	3.8313	1.0710
OQ	.9700	7.899	1.0881	.2328	.9838	9.394
SQ	1.3369	.5805	1.2794	.4655	1.2756	.3867
LQ	.2800	.2206	.2281	.1208	.2869	.2515
EI/EE	.7643	.32391	.5638	.2104	.6844	.2837
AR	1.2556	.6409	1.7413	.7858	1.6225	.7249
HO	100.92	5.6991	99.169	7.076	99.229	5.8144
H1H0	6.0600	4.1485	7.653	4.8188	7.0394	3.8571
Roll off	6.6256	2.6431	6.5075	2.0785	6.2725	1.9220

Table 3: Mean and SD of all parameters across emotions.

Pairs	t-value	Significance
T0 Joy - T0 Neutral	.307	.763
T0 Joy- T0 sorrow	.816	.427
T0 Neutral -T0 sorrow	1.125	.278
OQ Joy-OQ Neutral	2.012	.063
OQ Joy- OQ sorrow	1.239	.235
OQ Neutral -OQ sorrow	1.759	.099
SQ Joy-SQ Neutral	.393	.700
SQ Joy- SQ sorrow	.386	.705
SQ Neutral -SQ sorrow	.025	.980
LQ Joy - LQ Neutral	1.212	.244
LQ Joy- LQ sorrow	.086	.932
LQ Neutral -LQ sorrow	.991	.337
EI/EE Joy -EI/EE Neutral	2.63	0.03*
EI/EE Joy-EI/EE sorrow	.799	.437
EI/EE Neutral -EI/EE sorrow	2.163	**.047***
AR Joy -AR Neutral	1.903	.076
AR Joy- AR sorrow	1.882	.079
AR Neutral-AR sorrow	.670	.513
H0 Joy-H0 Neutral	1.123	.279
H0 Joy- H0 sorrow	1.591	.132
H0 Neutral- H0sorrow	.050	.961
HI H0 Joy- HI H0 Neutral	1.835	.086
HI H0 Joy- HI H0 sorrow	1.148	.269
H1 H0 Neutral -H1 H0 sorrow	.625	.541
Roll-off Joy- Roll-off Neutral	.229	.822
Roll-off Joy- Roll-off sorrow	.484	.636
Roll-off Neutral -Roll off sorrow	.350	.731

Table 4: t-values across emotions.

3. Interaction effect - gender * emotion:

The analysis was done separately for each parameter across gender and emotion and results revealed significant main effect of emotion - {f (2,28) = 4.064, p<0.05} for EI/EE - excitation energy. There was no significant main effect on emotion for pitch period TO {f (2,28) = .543, p=. 567}, open quotient OQ {f (2,28)=3.386, p=0.058}, speed quotient SQ {f (2,28)=0.098, p=0.907}, leakage quotient LQ {f (2,28)=0.597,p=0.557}, dynamic leakage, AR (f (2,28)=2.899, p=0.072), Harmonics H0 (f (2,28)=1.203,p=0.065}, first harmonics to fundamental frequency H1H0 {f (2,28)=1.544,p=0.231} and spectral roll- off RO {f (2,28)=0.158,p=0.0855) between emotions and gender. Since there was a main effect seen for EI/EE excitation energy, Bonferroni multiple pair wise comparison was done and results revealed a significant difference for the pairs EI/EE Joy - EI/EE Neutral and EI/EE Neutral - EI/EE Sorrow. There was a significant interaction effect of gender * emotion for EI/EE - excitation energy {f (2,28) = 4.843, p< 0.05} between gender and emotions. There was no interaction effect seen for pitch period T0 {f (2,28)=. 075, p=. 928}, open quotient OQ {f (2,28) = .388,p=. 682}, speed quotient SQ {f (2,28) = .419,p=. 662}, leakage quotient LQ {f (2,28)=2.877, p=. 073}, dynamic leakage, AR (f (2,28)=1.213,p=. 312), harmonics H0 (f (2,28)= 1.218,p=. 311), first harmonics to fundamental frequency H1H0 (f (2,28)=. 627,p=. 541} and spectral Roll- off {f (2,28)=1.251,p= 302).

CHAPTER V

DISCUSSION

Results showed a significant difference across gender on certain glottal parameters within each emotion. OQ (joy and neutral), LQ (sorrow), AR (sorrow) and H0 (sorrow) were significantly higher in males compared to females. Also spectral roll-off (sorrow) was lower in males compared to females. These results are in consonance with those of Airas & Alku (2006). Airas & Alku (2006) reported that there was a consistent gender difference in the glottal parameters amplitude quotient (AQ) and CLQ, which is defined as the ratio between the duration of glottal closing phase and the fundamental period. In the present study, there were significant gender differences found in the parameters open quotient, leakage quotient, dynamic leakage, harmonics and spectral roll-off. Airas & Alku reported that there was significant difference between male and female subjects in spectral roll - off, depicting smoothness of the glottal pulse. In their study, the subjects selected were professional stage actors and the subjects' task were to recite a text passage of 83 words of Finnish prose in a given emotion and in the present study the subjects' task was to sing kritis depicting various emotions. These changes in the profession of the subjects and the task carried out by the subjects in the present and that of Airas & Alku, could probably have yielded dissimilar results. Additionally, in the present study, there is no significant difference in the pitch period, which is a frequency measure, between male and female Carnatic singers because, the part of the vowel segment /a/, which was analyzed was the portion with highest fundamental frequency in the kriti being sung. However Airas & Alku (2006) found that there was a significant difference in the glottal parameters amplitude quotient and closing quotient, across the emotions anger, neutral, joy, tenderness and sorrow. The results of the present study are poor in consonance with this, where in significant differences are found only for the glottal parameter excitation ratio between the emotion pairs joy-neutral and sorrow-neutral. The possible reason attributed could be, emotions can be expressed well during speech than during singing Carnatic Kritis. Secondly, the subjects taken in Airas & Alku's study were professionally trained stage actors, who can possibly be more adept and prompt in expressing the emotions than Carnatic singers. Thirdly, the age range taken for Airas & Alku's study is 26-45 years, and the age range of the subjects taken for the present study are Carnatic singers in the age range of 18-28 years. so there is a possibility that years of experience that varies with both the group of subjects can be a factor yielding different results.

The results of the present study is also supports the results of Johnstone & Scherer (1999), where in they performed a feasibility study of EGG analysis to the study of emotional voice production and they found out that there was difference in males and females in the EGG parameters across various emotions. It was noted that CLQ closing quotient was smaller indicating the glottis closing faster in females than males in high arousal emotions.

The results of the present study imply several gender differences. Firstly higher OQ indicate higher TP/TN/TL or lower T0. It suggests higher amplitude in males compared to females. Also glottal abduction is more in males compared to females as indicated by higher LQ. The data also suggest smooth glottal closure in males compared to females.

Results also indicated significant difference between emotions. EI/EE was significantly higher in joy and sorrow compared to neutral emotion. This indicates that the amplitude of the epoch was high in these conditions compared to neutral emotion. ***Also, the results indicated significant interaction effect of gender * emotion.*** This was higher in females in neutral and sorrow emotions compared to males. This implies reduced amplitude of the epoch or negative amplitude.

Results of the present study has contributed to the field of musical acoustics. Future studies on comparison of vocal behavior at different FO in singers are warranted.

CHAPTER V

SUMMARY AND CONCLUSIONS

Classical singing is an art that a person develops after rigorous training and practice. Apart from lyrics used, vocal characteristics can express differences in emotions of the song rendered by these trained singers. Any discussion of the emotional response to music is faced at the very outset with the fact that very little is known about this response and its relation to stimulus. Evidence that it exists at all is largely based upon the introspective reports of the listeners and the testimony of the composers, performers and critics. It is a well known fact that vocal gestures relate to the physiology of larynx and voice quality differences produced by laryngeal features are one of the fundamental factors differentiating vocal emotional expression. Till date, no study has been specifically devoted to a detailed exploration of study of variations in vocal fold behavior during expression of emotions in classical singing. Therefore this study analyzed the vocal fold behavior during expression of various emotions in Carnatic singers, specifically at high pitch.

Sixteen Carnatic singers, eight males and eight females, in the age range of 18-28 years participated in the study. The subjects had passed the junior music examination and had no history of any laryngeal pathology, vocal abuse or vocal misuse behavior and the subjects had been practicing regularly.

Subjects were instructed to sing sapasa. Ra:gas depicting three emotions - joy, sorrow and neutral - were selected. Singers were instructed to sing kritis with these ra:ga:s. They were instructed to sing the kriti, till they were satisfied that they have expressed the intended emotion. All the singings were audio-recorded using a Digital mini-disc recorder and were transferred directly on to the computer at a sampling frequency of 16 kHz. Two musicians judged the emotions of these ra.gas and only those that were judged 100% to express intended emotions were considered for further acoustic analysis. VAGHMI software (Voice & Speech Systems, Bangalore) was used for the vocal fold behavior analyses. Each singer's base frequency was extracted from sa of sapasa using F0 program of VAGHMI. Vowel /a/ segment at highest frequency above the singer's base frequency in each emotion was selected and saved as separate files. These were then subjected to inverse filtering, and appropriate glottal markings were made.

The glottal parameters assessed were TO- pitch period measured between two successive epochs, refers to the time taken for one glottal vibration, OQ- open quotient which is the proportion of the period during which the glottis is open to the total period, SQ- speed quotient measures the ratio of opening and closing durations, LQ- leakage quotient which is the time taken for the voice source signal to return from epoch to the baseline, EI/EE- excitation ratio excitation strength is the negative amplitude at the time point of maximum discontinuity of the differentiated flow, AR- dynamic leakage is the residual flow during the return phase, which occurs from the time of excitation to the time of complete closure, H0- harmonics, Harmonic ratio-Hl - H0 is the ratio of energy at first harmonic and fundamental frequency and spectral roll-off which indicates the smoothness of the glottal closure.

The data obtained was tabulated and subjected to statistical analysis. SPSS software (version 10) was used to perform the analysis. An independent t-test was done to analyze the significant difference across genders and paired t-test was done to analyze the significant difference across emotions. Also mixed ANOVA was used to study the interaction effect of gender and emotion.

Results showed a significant difference across gender on certain glottal parameters within each emotion. OQ (joy and neutral), LQ (sorrow), AR (sorrow) and H0 (sorrow) were significantly higher in males compared to females. Also spectral roll-off (sorrow) was lower in males compared to females. These results are in consonance with those of Airas & Alku (2006). Airas & Alku (2006) reported that there was a consistent gender difference in the glottal parameters amplitude quotient (AQ) and CLQ, which is defined as the ratio between the duration of glottal closing phase and the fundamental period. In the present study, there were significant gender differences found in the parameters open quotient, leakage quotient, dynamic leakage, harmonics and spectral roll-off. Airas & Alku reported that there was significant difference between male and female subjects in spectral roll - off, depicting smoothness of the glottal pulse. In their study, the subjects selected were professional stage actors and the subjects' task were to recite a text passage of 83 words of Finnish prose in a given emotion and in the present study the subjects' task was to sing kritis depicting various emotions. These changes in the profession of the subjects and the task carried out by the subjects in the present and that of Airas & Alku, could probably have yielded dissimilar results. Additionally, in the present study, there is no significant difference in the pitch period, which is a frequency measure, between male and female Carnatic singers because, the part of the vowel segment /a/ , which was analyzed was the portion with highest fundamental frequency in the kriti being sung. However Airas & Alku (2006) found that there was a significant difference in the glottal parameters amplitude quotient and closing quotient, across the emotions anger, neutral, joy, tenderness and sorrow. The results of the present study are poor in consonance with this, where in significant differences are found only for the glottal parameter excitation ratio between the emotion pairs joy-neutral and sorrow-neutral. The possible reason attributed could be, emotions can be expressed well during speech than during singing Carnatic Kritis. Secondly, the subjects taken in Airas & Alku's study were professionally trained stage actors, who can possibly be more adept and prompt in expressing the emotions than Carnatic singers. Thirdly, the age range taken for Airas & Alku's study is 26-45 years, and the age range of the subjects taken for the present study are Carnatic singers in the age range of 18-28 years. So there is a possibility that years of experience that varies with both the group of subjects can be a factor yielding different results.

The results of the present study is also supports the results of Johnstone & Scherer (1999), where in they performed a feasibility study of EGG analysis to the study of emotional voice production and they found out that there was difference in males and females in the EGG parameters across various emotions. It was noted that CLQ closing quotient was smaller indicating the glottis closing faster in females than males in high arousal emotions.

The results of the present study imply several gender differences. Firstly higher OQ indicate higher TP/TN/TL or lower T0. It suggests higher amplitude in males compared to females. Also glottal abduction is more in males compared to females as indicated by higher LQ. The data also suggest smooth glottal closure in males compared to females.

Results also indicated significant difference between emotions. EI/EE was significantly higher in joy and sorrow compared to neutral emotion. This indicates that the amplitude of the epoch was high in these conditions compared to neutral emotion. ***Also, the results indicated significant interaction effect of gender * emotion***. This was higher in females in neutral and sorrow emotions compared to males. This implies reduced amplitude of the epoch or negative amplitude.

Results of the present study has contributed to the field of musical acoustics. Future studies on comparison of vocal behavior at different F0 in singers are warranted.

REFERENCES

Airas, M., & Alku, P. (2006). Emotions in vowel segments of continuous speech: analysis of the glottal flow using the normalized amplitude quotient *Phonetica, 63,26-46.*

Ananthapadmanabha, T.V.,(2008). Voice source characterization by Inverse Filtering. *Proceedings of the national workshop on voice: Assessment and Management,* 1-25.

Banse, R., & Scherer, K.R.,(1996). Acoustic profiles in vocal emotion expression. *Folia phonetricia,70,6*14-636

Boominathan, P., Scherer, R.C., Shrivastav, R., Rajashekhar, B., Ravikumar, A., & Nagarajan, R. (2006). Carnatic singing - some aspects. *Symposium: Care of professional voice,* Philadelphia, USA.

Bunch, M. (1982). *Dynamics of singing voice.* New York: Springer-Verlag.

Childers, D. G., Naik, J. M., Larar, J. N., Krishnamurthy, A. K.& Moore, G. P. (1983)."Electroglottography, speech and ultra-high speed cinematography, in *Vocal Fold Physiology and Biophysics of Voice,* edited by I. Titze and R.Scherer (Denver Center for the Performing Arts, Denver), pp. 202-220.

Christer, G., & Chaider, A.,(2003). The role of voice quality in communicating emotion, mood and attitude. *Speech communication*,40,189-212.

Cummings, K. E., & Clements, M.A.(1995) Analysis of glottal excitation of emotionally styled and stressed speech. *The Journal of the Acoustic Society of America,98,88-98.*

Fairbanks, G. (1940)recent experimental investigations of vocal pitch in speech. *The Journal of the Acoustic Society of America,*

Fairbanks, G., & Provonost, W.(1938). Vocal pitch during simulated emotion. *Science,88,382-381.*

Fairbanks, G.,& Provonost, W. (1939). An experimental study of the pitch characteristic of the voice during expression of emotion. *Speech Monography ,6,87* -104.

Fairbanks, G.,& Hoaglin, L.W.(1941).An experimental study of the durational characteristics of voice during the expression of emotion. *Speech Monography,* 8,85-91.

Hanson, D. G., Gerratt, B. R., & Berke, G. S. (1990). Frequency, intensity and target matching effects on photoglottographical measures of open quotient and speed quotient. *Journal of Speech and Hearing Research,33,45-50.*

Henrich, N., d'Alessandro, C, Castellengo, M., & Doval, B. (2004). "On the use of the derivative of electroglottographic signals for characterization of nonpathological phonation,". *The Journal of the Acoustical society of America,* 115,3,1321-1332.

Kaiser, L. (1962). Communication of affects using single words. *Synthese, 14,300-* 319.

Kitzing, P., and Sonesson, B. (1974). "A photoglottographical study of the female vocal folds during phonation,". *Folia Phoniatricia* 26,138-149.

Koufman, J. A., Radomski, T.A., Joharji, G.M., Russel, G.B., & Pillsbury, D.C (1996). "Laryngeal biomechanics of the singing voice".*Otolaryngology- head and neck surgery,* 115(6),527-37.

Laukkanen, A.., Vilkman, E., & Alku, P.(1996). Physical variations related to stress and emotional state: A preliminary study. *Journal of Phonetics*,,24,.313-335.

Laukkanen, A.., Vilkman, E., & Alku, P.(1997).on the perception of emotions in speech : the role of voice quality. *Folia phonetricia,22,l 57-168* .

Leonard, B. (1993). *Emotion and meaning in Music.* Chicago: Macmillan & co.

Petterson, V., & Westgaard, R.H. (2005) 'The activity patterns of neck muscles in professional classical singing' *Journal of voice,* 19(2), 238-51.

Satyanarayana, R.(1983). *Karnatakadalli sangeetha kale.* Bangalore: karnataka Sahitya parishat.

Schrer, K.R.,(2000). Vocal effect expression: a review and a model for future research.

Schrer, K.R.,(2003). Vocal communication of emotion: a review of research *Paradigms.Speechcommunication.40,227-256*

Skinner, E.R.,(1935). A calibrated recording and analysis of the pitch, force, quality of vocal tones expressing happiness and sadness. *Speech Monography,2,8l-* 137.

Strube, H.W.,(1974). Determination of instant glottal pressure from the speech wave. *Journal of acoustical society of America.* 56,1625-1639.

Sundberg, J.(1984). Voice source characteristics in six premier country singers. *Journal of voice.* 13,168-183.

Sundberg, J. (1990). What's so special about carnatic singers', *Journal of voice,* 4(2), 107-119.

Titze, I. R., Baer, T., Cooper, D., & Scherer, R. (1984). "Automatic extraction glottographic waveform parameters and regression to acoustic andphysiologic variables," in *Vocal Fold Physiology: Contemporary Research in Clinical Issues,* edited by A. J. Bless DM (College Hill, San Diego), pp. 146-154

Williams, C.E., & Stevens, K.N.,(1969). On determining the emotional state of pilots during flight: an exploratory study. *Aerospace media,40,1369-1372.*

Williams, C.E., & Stevens, K.N.,(1972). Emotions and speech; some acoustical correlates. *The Journal of the Acoustical Society of America,52,*1238- 1251.