VOICE ANALYSIS OF PRIMARY (ANGER AND SAD) EMOTIONAL SPEECH UTTERANCES

Sheba Merin Jacob

Register No: 14SLP027

A Dissertation Submitted in Part Fulfilment for the degree of

Master of Science (Speech-Language Pathology),

University of Mysore, Mysuru

ALL INDIA INSTITUTE OF SPEECH AND HEARING

MANASAGANGOTHRI

MYSURU-570006

May, 2016

Certificate

This is to certify that this dissertation entitled **"Voice Analysis of Primary (Anger and Sad) Emotional Speech Utterances"** is a bonafide work in part fulfillment for the Degree of Master of Science (Speech-Language Pathology) of the student (Registration No.14SLP027). 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.

Mysuru

May, 2016

Dr. S. R. Savithri
Director
All India Institute of Speech and Hearing
Manasagangothri, Mysuru- 570006.

Certificate

This is to certify that this dissertation entitled **"Voice Analysis of Primary (Anger and Sad) Emotional Speech Utterances"** is a bonafide work of in part fulfillment for the degree of Master of Science (Speech and Language Pathology) of the student (Registration No. 14SLP027).This has been carried out under my guidance and has not been submitted earlier to any other University for the award of any other Diploma or Degree.

Mysuru,

May, 2016

Dr. T. Jayakumar
Guide
Reader in Speech Sciences
Department of Speech-Language Sciences
All India Institute of Speech and Hearing
Manasagangothri, Mysuru-570006.

Declaration

This dissertation entitled **"Voice Analysis of Primary (Anger and Sad) Emotional Speech Utterances"** is the result of my own study under the guidance of Dr. T. Jayakumar, Reader in Speech Language Sciences, Department of Speech-Language Sciences, All India Institute of Speech and Hearing, Mysuru, and has not been submitted earlier in any other University for the award of any Diploma or Degree.

Mysuru,

May, 2016 Register No. 14SLP027

Acknowledgement

When times for thanksgiving comes, it is always a pleasure to recall the precious faces whom I am indebted to….

J K sir I thank you for all the support and guidance and for being patient with me and helped me to finish this work.

Savithri ma'am…. thank you for providing the opportunity to do this work.

Santhosh sir…thank you for guiding to do all the statistical analysis and for clearing my doubts

A big thanks for all my participants of the study…without whom I couldnot finish my study.

Pappa, Mummy, Chechi, Achachen and Shalu… you are more than I could thank. You people know me more than I know myself.

Nenthrapazhamzz (Akku, Radhika, Bharu, Akshatha, Vidhya, Anchala, Anitta,Merin Chechi)….You made my life easier in AIISH…I wish we could continue to be like this forever, starting off in the morning with hostel coffee and ending with a good movie..

Ajitha … you are my special friend and motivator, You understands me before I speaks….

Sowmya…thankyou for all the motivation and prayers…

Kadi and Bharu …. The best dissertation partners I could ever get…Enjoyed every moment with you guys.

Pavi, Swathi , Yashu , Radhika…The last benchers and latecomers.…Love all the time in class because of you guys…

Varsha, Veepee, Jasiya, Merin, Kirti, Sarga, Meenu, Anju, Anoopa, Sneha, Nainitha, Anu.….The best juniors we had got..Enjoyed all the moments with us…..

Rofina chechi,thankyou for preparing my stimulus…

Serooya chechi, Nazmin chechi, Anitha chechi, Jyotsna chechi, Deepthi chechi.…thankyou for all the support and motivation.

And mostly, thankyou god almighty for all the gifts you have provided me…..

Table of Contents

List of Tables

List of Figures

CHAPTER I

Introduction

The human voice is a great degree adaptable device and most essential method for conveying information between individuals. The specific ability of the human being to produce and perceive vocal speech sounds require an anatomical development of the vocal tract, underlying neural mechanism and capacity for vocal learning and speech perceptual specialization(Fitch 2000).

According to classic theories, voice source and vocal tract function are the two important components of speech production (Fant, 1970; Titze, 1990). Human voice consists of a basic sound and it gets modified according to the specific features of language. These features include phonemes which represent vowels and consonants and also the grammatically significant features like pauses and intonation. Speech consists of segmental aspects and suprasegmental aspect. Segmental aspects carry linguistic information. Paralinguistic and non-linguistic information's are carried by long term suprasegmental aspects.

Human voice can predict non-linguistic information such as age, gender, regional and educational background and the emotional state of an individual. Emotions can be expressed in the linguistic structure and content of speech and it also affects suprasegmental characteristics of speech (Johnstone, 2001). Various aspects of a speaker's physical and emotional state can be identified by voice alone. It includes age, sex, appearance, intelligence, and personality of the individual (Kramer, 1963)

The respiratory system, phonatory system and the resonance system are the key elements in the production of speech. Respiratory system gives a regulated air pressure that runs the phonatory system. Phonatory system majorly constitutes larynx and vocal folds. The vocal folds are placed aside during quiet breathing with low laryngeal resistance and while phonating, air coming up from the lungs encounters the closed vocal folds. Resistance of the vocal folds is taken away by the pressure and flow of the air, which sets them into a pattern of rapid vibration. As the air flows through the glottis, the pressure between the vocal folds reduces (Bernoulli Effect), and causes the vocal folds to close. It leads to the production fundamental frequency (F0) along with its multiples or harmonics. Any difference in the air pressure beneath the vocal cords can lead to change in opening and closing of vocal cords and intern result in change in fundamental frequency, intensity and harmonic energy. Such a change can be expected for high arousal emotions like anger. The resonatory system filters the sound as it passes the vocal tract. It extends from glottis to the oral and nasal cavities. Based on the shape and length of the resonance system, harmonics vary, leading to a highly complicated, radiated speech output Lieberman & Blumstein (1988).

Scherer (2001) defined emotion as an episode of interrelated, synchronized changes in the states of all or most of the five organismic subsystems in response to the evaluation of an external or internal stimulus event as relevant to major concerns of the organism. Johnson & Oatley (1989) reported that emotions are different from moods- emotions emerge suddenly in response to specific stimuli, and it may last for seconds or minutes, whereas moods may last for hours or days and is more uncertain or vague in nature; emotions can broadly be considered as being concerned with changing something, and moods with maintaining something.

The expression of emotions plays a key role in human communication. Different theories categorize the various emotions based on their specific individualistic approaches. Certain theories divide emotions as those belonging to the same category and those to different category. Another approach classifies emotions as basic or primary and secondary emotions wherein basic emotions (joy, anger, fear, sadness, disgust, surprise) are defined as those that are not culturally bound and considered to represent evolutionary and survival related patterns of responses. The secondary emotions include uncertainty, diffidence, love and empathy and are affected by culture (Murray and Arnott, 1993)

Darwin, Ekma & Prodger (1872) recognized that voice changes can be caused by emotion. Emotion is conveyed by a subtle combination of features at suprasegmental, segmental, and intrasegmental levels of speech abstraction. Intrasegmental level means vocal voice quality adjustments due to different type of emotions (Murray & Arnott, 1993). According to Scherer (2003) Emotion elicitation methods can control the way emotions are induced and it can be classified into: natural vocal emotion expression, induced emotional expression, and simulated emotional expression.

Voice cues measured from emotional voice can be divided in to those related to (a) Fundamental frequency,.(b) Voice intensity, (c) voice quality and (d) Temporal aspects of speech. Few studies were done during early 1930's including Skinner (1935) who investigated fundamental frequency (F0) and intensity patterns in production of induced happiness and sadness. Fairbanks & Pronovost (1938) had explored fundamental frequency, intensity and durational patterns of a speech passage elicited in five different dramatized expressions. After the early forties, there was little research on the vocal expression of emotions until the 1960s. Later in 1970s computers were used for studying acoustic correlates of emotional speech utterances. The use of computers became common and digital analysis is used for the study of acoustical correlates of emotion in speech. Studies on effect of emotion on acoustic characteristics have shown that average values and range of fundamental frequency differ from one emotion to another.

Cowan (1936) has compared unemotional speech and emotional speech and noted that F0 is normally distributed at average F0 level. He found that unemotional speech has a narrow F0 range compared to emotional speech and happiness gives an increased F0 and F0 range.

Lieberman & Michaels (1962) had done investigations to relate emotional content of speech to its F0 and amplitude using a Vocoder and they found that F0 was not wholly responsible for communication of emotion. Amplitude played a small part in the process, as did the "fine structure" of theF0 and also noted that the emotional modes studied did not all depend to the same degree on the acoustical parameters.

Darvitz (1964) found that speech rate and intensity increased for happiness. William & Steven (1972) for sorrow speaking situation average F0 and F0 range was found to be lower than neutral speaking situation. It was also accompanied by reduction in rate of articulation and increase in duration of utterance.

Laukkanen, Vilkman, Alku and Oksanen (1997) examined variation of formant frequencies in different emotional conditions and found positive emotions has higher F3 and F4 than negative emotions.

Electroglottography (EGG) can be used to study the vocal fold contact pattern in emotional voice. EGG is a noninvasive instrument which measures the time variation of the degree of contact between the vibrating vocal folds during voice production. It calculates contact quotient (CQ), open quotient (OQ) and speed quotient (SQ). EGG can be used to measure whether vocal fold vibration is sensitive to emotional stimuli (Baken & Orkioff, 2000). Increased muscle activity may be reflected as an increase in CQ of EGG and it might reflect vocal fold activity changes due to emotional stimuli.

Johnstone & Scherer (2000) had done experiment of emotional vocal recording using a computer emotion induction task. Acoustic and physiological parameters were analyzed. Results indicate that high arousal emotions are characterized by high F0 and high RMS energy and the glottis closes faster. F0 floor was found to be lowest for the emotions bored and depressed, and highest for happy and anxious speech. He found sadness has decrease in mean F0, F0 range and variability, speed and articulation rate and intensity.

Gobl & Ni Chasaide (2003) investigated the role of voice quality in the communication of emotions, moods and attitudes. Seven different voice qualities were selected according to the listeners perception after the production of different emotional conditions. The voice qualities included harsh voice, tense voice, modal voice, breathy voice, whispery voice, creaky voice and lax–creaky voice. These were synthesised using a formant synthesizer and result suggests that there is association between for few of the emotions and voice quality such as anger and tense voice but for other emotions such an no one-to-one mapping could not be found.

Waaramaa, Laukkanen, Alku & Väyrynen (2008) examined the role of voice source characteristics and formant frequencies in the communication of emotions in monopitched vowel samples [a:], [i:] and [u:]. Student actors (5 males, 8 females) produced the emotional samples simulating joy, tenderness, sadness, anger and a neutral emotional state. Equivalent sound level (L eq), alpha ratio [SPL (1–5 kHz) – SPL (50 Hz–1 kHz)] and formant frequencies F1–F4 were measured. The results found that the capacity of monopitched vowels for conveying emotions differed. Vowels [a:], [i:], [u:] can carry emotional information in various ways, [a:] being the most effective and [u:] the least effective in this function. The spectral structures of vowels as such may be associated with certain emotional colourings of the voice: The open back vowel [a:] was shown to be a successful carrier of tenderness. The front vowel [i:] easily conveyed joy and also sadness. Vowel [u:] was likewise a successful carrier of sadness, and the least effective carrier of the positive emotions of tenderness and joy. Anger tended to be well recognized from each vowel mentioned.

Murphy & Laukkanen (2009) used EGG to differentiate emotions produced by a female speaker expressing joy, anger, sadness, tenderness, and neutral conditions. He found that entire opening phase and final open phase could discriminate emotional expressions with anger and joy having the greatest activity followed by neutrality, sadness and tenderness.

Murphy & Laukkanen (2012) investigated EGG measures related to closed quotient and rate of contact at opening and closing and found that contact quotient, speed quotient and acceleration at vocal fold closure differs across emotional expressions. They have also studied the normalized time of increasing contact of the vocal folds (NTIC), which is a laryngeal measure estimated from EGG using sustained [a:] vowel. However, it was concluded that NTIC was not a significant factor in differentiating between the various emotions.

Johnstone, T., Banse, R. and Scherer, K. R. (1995) had done an investigation to test the feasibility of using EGG to measure changes in vocal fold function during emotional speech production. Acoustic and EGG recordings were made for standard phrases with different emotions. The results shows that EGG measurements might provide useful information on the glottal and laryngeal mechanisms responsible for emotional changes to speech, although more precision in measuring EGG would be required to draw conclusions about spectral acoustic characteristics

Waaramaa & Kankare (2013) investigate glottal and filter variables of emotional expressions vary by emotion and valence expressed. Prolonged emotional vowels were produced by professional actors and actresses expressing joy, surprise, interest, sadness, fear, anger, disgust, and a neutral emotional state and found that there is a significant difference between vocal fold contact times for various emotions. For positive emotions CQ of EGG may vary simultaneously and inversely with F3 and F4. Also they concluded that higher formant frequencies were affected by changes in the lower pharynx and larynx.

Waaramaa, Palo & Kankare (2014) studied voice quality differences between freely varying and mono-F0ed vowels expressed by professional actors. The results suggested significant gender differences for CQ of EGG threshold level. SPL, CQ of

EGG, and F4 were used to convey emotions, but to a lesser degree, when F0 was predetermined.

Emotional vocal expressions have traditionally been studied for fundamental frequency (F0), sound pressure level (SPL) (or equivalent sound level, L eq) and duration. Voice quality features due to the glottal characteristics have been much less studied in emotional voice. There is not so much of the studies described the aerodynamic and glottal characteristics. Also, to measure the emotional variation most of studies used prolonged vowel phonation, which may not ideal stimulus for evoking the emotion. Hence, in the present study, speech also used as stimuli to estimated glottal characteristics rather than only phonation. Emotional aspect can vary with language and culture. So the voice quality features arising from glottal characteristics and air pressure in the emotional utterances of vowels will be studied in Malayalam language in the present investigation.

Aim of the Study

The present study aims to investigate few of the aerodynamic and electroglottographic analysis of primary emotional utterances (anger and sad) in normal individuals.

Objectives of the Study

The objectives of the study were to measure

- The variation of EGG quotient in anger and sad emotional utterances.
- Subglottal air pressure variation in anger and sad emotional utterances.

CHAPTER II

Method

Participants

Thirty normal native Malayalam speakers in the age range of 18- 35 year (Mean age-20.6, SD - 2.07) participated in the study. Malayalam is there spoken language since childhood and could speak, read, and write fluently. All of them completed minimum of 12 years of education. All participants had normal speech and language skills, no voice problem, hearing loss and no history of any other neurological or psychological problems.

Exclusion criteria for the group

- Participants, with upper respiratory tract infection, asthma, or allergic diseases at the time of recording.
- Participants with the history of neurological, speech or language disorders.
- Participants with history heavy alcohol or smoking were excluded from the study.

Stimuli

Sentences

Three meaningful Malayalam sentences were selected. The sentences had minimum of two repetitions of bilabials /pa/were presented. This was for the purpose of measuring subglottal air pressure. Rothenberg (1973) suggest that the use of /pa/ syllable, for the indirect measures of subglottal air pressure.

Sentences:

1. /pu:tʃa pa:lu kUdItʃu/ പൂച്ച പാല് കുടിച്ചു

2. /kUṭṭI pu:vu parItʃu/ കുട്ടി പൂവ് പറിച്ചു

3. /paṭṭI pu:tʃaje pIdItʃu/ പട്ടി പൂച്ചയെ പിടിച്ചു

Recording environment

All the recording was done in department laboratory. The laboratory was relatively free from external noise. The participants were seated comfortably.

Instrumentation

Electroglottograph

The Electroglottograph, Model 6103 (Kay elemetrics, NJ) was used for measuring glottal vibratory characteristics and relative contact patterns. All the subjects were explained about the purpose of the study and familiarization was done

Procedure

The participants were seated comfortably in a quiet room and the electrodes were placed on the alae of the thyroid cartilage. Once the electrodes are positioned appropriately, the participants were asked to speak each sentences in two different emotional conditions (anger and sad). The gain of the instrument was adjusted manually and the task was captured directly on the instrument. Audio recorded model sample was given prior to the task.

Practice trials were done prior to the actual recording and all the data were saved on computer memory. The recording that yielded best waveform morphology was retained for extraction of parameters. The curser was positioned at the steady state of vowels and F0, contact quotient (CQ), open quotient (OQ) and speed quotient (SQ) were extracted.

Aeroview

Aeroview system version1.5.0 (Kay telemetric, NJ) was used to measure the sub glottal air pressure for speech task. The aeroview is a computer based portable system consisting of a circumferentially vented pneumotachograph mask coupled to PT-25 B pressure transducer with an intraoral tube attached to it.

Procedure

Before each recording in the aeroview system the surface pneumotachograph mask and intra oral tube was cleaned with antiseptic liquid. The instrument was calibrated as per the instructions in the manual and checked for proper measurements.

The participants were made to sit straight and comfortably in the chair. They were given instructions about the instrument and the placement of the mask. The participants were asked to place the mask firmly on the face and to place the intraoral tube within the oral cavity (with close lips and without the tongue being occluded). They were made to utter each sentence in to different emotional conditions. Two practice demonstration trials were done prior to the actual recording. The demonstration included the validated emotional utters by the native speaker. The data was stored for subglottal air pressure for bilabial stop consonants.

Analysis

The measured EGG and Air pressure parameters were tabulated for further statistical analysis. The data was then subjected for suitable statistical analysis using

SPSS software. The mean and standard deviation was extracted. Repeated measure ANOVA was done to analyze the variation of glottal and air pressure parameters across two emotional conditions.

CHAPTER III

Results

The present study focused on few aerodynamic and electroglottographic parameters of primary emotional utterances (anger and sad) in typically developing individuals. The data was obtained from 30 native typically developing female Malayalam speakers in the age range of 18-35 years. The participants were asked to produce three sentences in two different emotional conditions.

The electoglottographic and aerodynamic data were obtained using Electroglottograph, Model 6103 and Aeroview system version1.5 as per the procedure described in the method. The data was analyzed using SPSS 20 version with following statistical methods:

1. Descriptive statistics (Mean and SD) of electoglottographic and aerodynamic parameters.
2. Repeated measure ANOVA to compare the different emotional conditions and different vowels.
3. Paired 't' test to compare between vowels

The results of the study are explained under the following headings:

1. Test of normality
2. Comparison between sad and anger emotional condition
3. Comparison between vowels in two emotional condition

3.1 Test of Normality

Table 1:

Kolmogorov-Smirnov normality value for electroglottographic and aerodynamic parameters

Normality Test (n=30)					
Parameters	Vowels	Sad		Anger	
		K-S value	Sig	K-S value	Sig
F0	/a/	0.10	0.200	0.10	0.200
	/i/	0.13	0.182	0.14	0.091
	/u/	0.15	0.072	0.11	0.200
CQ	/a/	0.10	0.200	0.16	0.032*
	/i/	0.12	0.200	0.08	0.200
	/u/	0.12	0.200	0.11	0.200
OQ	/a/	0.10	0.200	0.15	0.066
	/i/	0.09	0.200	0.08	0.200
	/u/	0.11	0.200	0.11	0.200
SQ	/a/	0.08	0.200	0.19	0.004*
	/i/	0.16	0.035*	0.12	0.200
	/u/	0.15	0.063	0.08	0.200
Sub –P	/a/	0.09	0.200	0.12	0.200
	/i/	0.14	0.115	0.15	0.060
	/u/	0.12	0.200	0.14	0.099

**P<0.05-nonnormal distribution, F0- Fundamental Frequency, CQ- Contact Quotient, OQ- Open Quotient, SQ- Speed Quotient, Sub –P- Subglottic Pressure*

Kolmogorov-Smirnov normality test was performed to check for normality. The most of the data obtained was found to have normal distribution ($p > 0.05$), hence parametric test was used to compare between the two emotional conditions and different vowel condition. Table 1 shows the K-S value and p- value for each parameter.

Table 2: Mean, Standard Deviation, F- value, p- value for electroglottographic and aerodynamic parameters in sad and anger emotional condition

Parameter	Anger		Sad		F	p value
	M	SD	M	SD		
F0	262.96	39.26	241.19	38.80	64.35	0.000
CQ	44.70	3.32	43.07	3.09	29.86	0.000
OQ	55.20	3.69	56.83	3.01	26.19	0.000
SQ	248.79	52.52	233.83	45.91	06.46	0.000
Sub-P	7.59	1.82	3.50	1.20	35.40	0.000

Note: F0- Fundamental Frequency, CQ- Contact Quotient, OQ- Open Quotient, SQ- Speed Quotient, Sub –P- Subglottic Pressure

The mean and SD was obtained using descriptive statistics and repeated measure ANOVA was used to compare between sad and anger emotional condition. From the table 2 the results show that there is a significant difference between electroglottographic and aerodynamic parameters in sad and anger emotional condition. F0 ($F=64.35$, $p<0.05$) and subglottal pressure ($F=35.4$, $p<0.05$) showed greatest difference between emotional conditions followed by CQ, OQ, and SQ. The F0, CQ, SQ and subglottic air pressure showed high value for the anger condition compared to sad condition. On the other hand OQ showed lower value for anger compared to sad emotional condition.

To find the difference across the vowel in different emotional condition, repeated measure ANOVA is used for individual electroglottographic and aerodynamic parameter.

Table 3:

Mean, SD, F- value and p- value for vowels /a/, /i/, and /u/ in two emotional conditions.

Parameter	**Emotion**	**/a/**	**/i/**	**/u/**	**F value**	**p value**
		M (SD)	M (SD)	M (SD)		
F0	Sad	240.02 (34.99)	233.84 (44.11)	249.72(36.30)	0.93	0.397
	Anger	256.34 (31.36)	255.40 (32.06)	277.14(49.07)	0.93	0.387
Sub-P	Sad	3.28(1.16)	2.99 (0.81)	4.21(1.26)	192.82	0.000*
	Anger	7.37(1.37)	7.62 (1.78)	7.78(1.85)	192.82	0.000*

*Note: F0- Fundamental Frequency, Sub –P- Subglottic Pressure, *p<0.05*

Table 3 shows mean values of F0 and sub glottal pressure for each of the vowels across the two conditions. Repeated measure ANOVA was used to compare between vowels in sad and anger emotional condition. Results shows that there is no significant difference seen across vowels in sad and anger emotional condition for F0 (F=0.93, p-0.397) and there is a significant difference seen across vowels in sad and anger emotional condition for subglottal air pressure (F=35.401, p<0.00).Further pair wise comparison is carried out for subglottal air pressure to see any interaction between the vowels.

Table 4:

Pairwise comparison of different vowels for Sub glottal air pressure.

Vowel	**Anger**	**Sad**
	p value	p value
/a/ Vs /i/	0.256	0.279
/i/ Vs /u/	0.667	0.000*
/u/ Vs /a/	0.493	0.000*

**p<0.05*

Pairwise comparison was carried out to find differences between each vowel. The results found that /u/ and /i/, /u/ and /a/ showed a significant difference (*p<0.005) for sub glottal air pressure in sad emotional condition. There is no significant difference seen in anger emotional condition. Table 4 shows the comparison of vowels across emotions.

Table 5:

Mean, SD, F- value and p value for CQ, OQ and SQ in two emotional conditions

Parameter	**Emotion**	**/a/**	**/i/**	**/u/**	**F- value**	**p value**
		M (SD)	M (SD)	M (SD)		
CQ	Sad	43.40 (3.16)	43.08 (3.07)	42.74 (3.11)	1.026	0.365
	Anger	45.07 (3.17)	44.32(3.39)	44.72(3.48)	1.026	0.360
OQ	Sad	56.51 (3.19)	56.88(3.05)	57.09(2.84)	1.25	0.292
	Anger	54.70(3.50)	55.80(3.51)	55.10(4.08)	1.25	0.291
SQ	Sad	228.71(48.73)	245.86(46.72)	226.92(41.05)	0.79	0.457
	Anger	252.89(58.20)	252.42(50.12)	241.07(49.73)	0.79	0.439

Note: CQ- Contact Quotient, OQ- Open Quotient, SQ- Speed Quotient

Table 5 shows mean values of CQ, OQ and SQ for each of the vowels across the two conditions. Repeated measure ANOVA was used to compare between vowels in sad and anger emotional condition and results shows that there is no significant difference seen between vowels in all the emotional conditions. The values are higher for CQ followed by OQ and SQ.

CHAPTER IV

Discussion

The present study aimed to investigate the aerodynamic and electroglottographic analysis of primary emotional utterances (anger and sad) in normal individuals. The objective of the study was to measure the variation of EGG quotient and glottal air pressure variation in anger and sad emotional utterances.

Thirty normal native Malayalam speakers in the age range of 18- 35 year participated in the study. Three meaningful Malayalam sentences were selected and participants were asked to produce these three sentences in sad and anger emotional condition. All the recordings were made using Electroglottograph, Model 6103 and Aeroview system version1.5.0.

4.1 Comparison of Electoglottographic and Aerodynamic parameters across two emotional conditions

The results showed that there is a significant difference between the parameters in two emotional conditions. All the parameters had increased values in anger emotional condition except for open quotient. The OQ has reduced in anger emotional condition. The F0, Contact quotient, Speed Quotient, and Subglottal air pressure has increased in anger emotional condition.

4.1.1 Fundamental Frequency

The F0 has increased from 241.19 Hz in sad emotional condition to 262.96Hz in anger emotional condition ($F=64.35$, $p< 0.05$). This can be due to the increased tension on the extrinsic and intrinsic laryngeal muscle while producing anger emotional condition which can lead to increased F0. Figure 1 shows the mean F0 across two emotional conditions.

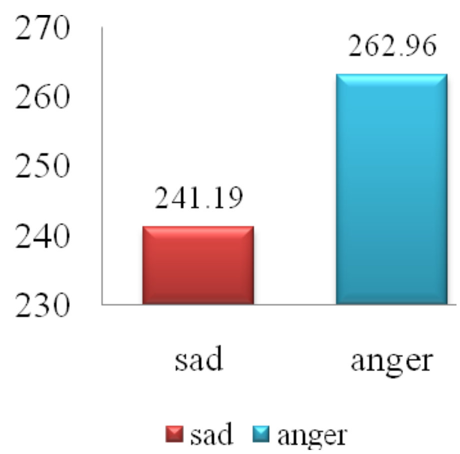

Figure 1: Mean F0 across different emotions

Larynx tends to rise and pharynx tends to narrow during the production of tense voice quality. Thus during the production of high activation emotions like anger the larynx rise results in increased fundamental frequency. This might be one of the reasons why there is increased F0 during the expression of anger emotional condition.

William & Steven (1972) found that for sorrow speaking situation average F0 and F0 range was found to be lower than neutral speaking situation. Johnstone, Banse & Scherer (1995) had also found similar results in an experiment of vocal recording using computer emotion induction task. They found that high arousal emotions are characterized by increased F0 and RMS energy and the vocal folds adducts faster. Johnstone & Scherer (2000) also found in sad emotional condition there is a decrease in mean F0, F0 range and variability, speed and articulation rate and intensity, which can be due to the reduction in high frequency energy.

Gobl & Ni Chasaide (2003) had studied the association of emotion and voice qualities using perceptual testing of synthetic speech samples. The found that emotions with high activation or high power (confidence, interest, happy, angry and stressed) were associated with tense and harsh voice quality and emotions with low activation (relaxed, content, intimate, friendly, sad and bored) has breathy, whispery and creaky voice quality.

Contact Quotient (CQ)

The CQ has increased from 43.07 to 44.70 in anger emotional condition ($F = 29.86$, $p<0.05$). Figure 2 shows the mean CQ across two emotional conditions.

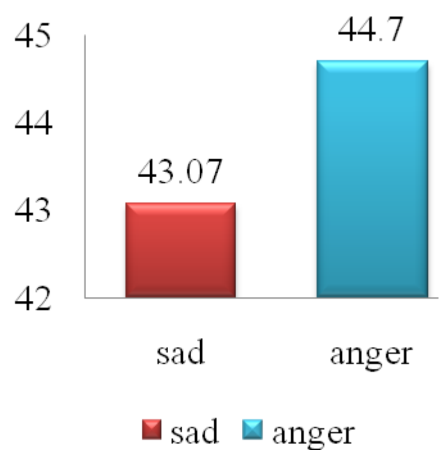

Figure 2: Mean CQ across different emotions

This shows that for the anger emotional condition the vocal folds are tightly adducting than sad emotional condition. Increased CQ values shows increased glottal adduction and stronger contraction of interarytenoids and lateral cricoarytenoid and increased medial compression from the activity of thyroarytenoid muscle. The vocal quality can be similar like of pressed speech.

Murphy & Laukkanen (2012) had also investigated EGG measures related to closed quotient and rate of contact at opening and closing and found that contact quotient, speed quotient and acceleration at vocal fold closure differs between emotional expressions. In sad emotion condition the CQ is lower, indicating reduced glottal adduction. The reduced vocal adduction can lead to a vocal quality similar to that of breathy phonation.

Open Quotient (OQ)

The OQ is the open time divided by the total period length. It has decreased in anger emotional condition from 56.83 to 55.20 ($F= 26.19$, $p<0.05$). Figure 3 shows the mean OQ across two emotional conditions.

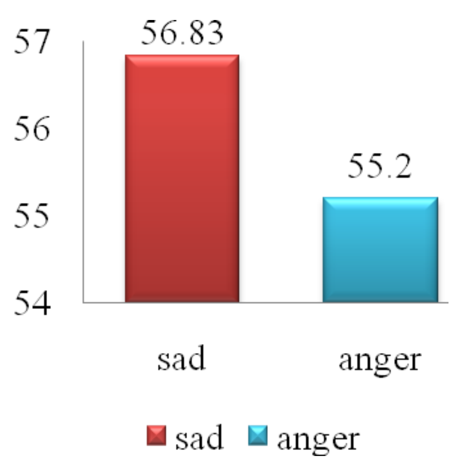

Figure 3: Mean OQ across different emotions

In anger condition OQ is less indicating prolonged closing time and reduced open phase. There can be incomplete closure of the vocal folds in sad emotional condition and there is an increased open phase and it can indicate breathy, whispery or creaky voice quality.

Murphy & Laukkanen (2009) examined discrimination of emotions produced by a female speaker expressing joy, anger, sadness, tenderness, and neutral conditions using the electroglottogram (EGG) signal. Entire opening phase and final open phase could differentiate emotional expressions with anger and joy having the greatest activity followed by neutrality, sadness and tenderness.

Murphy & Laukkanen (2012) investigated EGG measures related to closed quotient and rate of contact at opening and closing and found that contact quotient, speed quotient and acceleration at vocal fold closure differs between emotional expressions

Rising of F0 typically increases the SPL by increasing the glottal closing speed or speed of flow cessation in a pulse, and consequently decreasing OQ. In the present study there is also an increase of F0 in anger emotional condition which can also attribute for the decrease in OQ.

Speed Quotient (SQ)

SQ is the opening time divided by the closing time of the glottis. Figure 4 shows the mean SQ across to emotional condition.

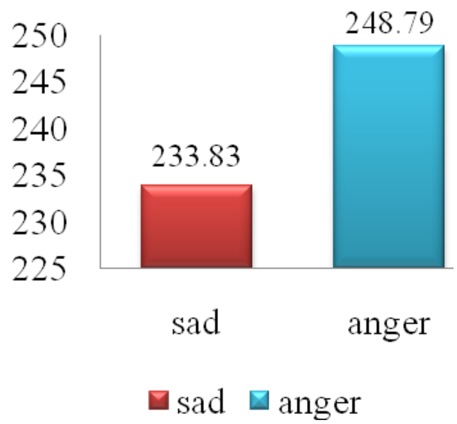

Figure 4: Mean SQ across different emotions

The mean value for the Speed Quotient (SQ) has increased in anger emotional condition from 233.83 to 248.79 ($F= 06.46$, $p <0.05$). Speed Quotient (SQ) is an important parameter of vocal fold vibration. Holmberg (1988) & Dromey (1992) defined it as the ratio between closing duration and opening duration, and thus a measure of the symmetry of the glottal pulses. As creaky phonation pulses are usually shorter in closing and longer in opening, their pulse shape is skewed, while breathy phonation pulses, with similar closing and opening durations, show a more symmetrical shape. Esling (1984) found that skewness of EGG signals is a useful indicator of phonation types. But more recent studies (e.g. Keating, Esposito, Garellek, Khan, & Kuang (2012) on phonation contrasts across languages) have not found that EGG SQ reliably varies with phonation in the way that CQ does. In the present study it has increased in anger emotional condition.

The variation in SQ can be more associated with increase in SPL. The rising in tone intensity can lead to increased SQ values as reported by Von Leden, Moore &

Timcke (1960), Kitzing & Sonesson (1974). In anger emotional condition the increase in F0 may lead to increase in loudness and this may attribute for the increase in the SQ values.

Sub Glottic Pressure (Sub – P)

The subglottic pressure has increased in anger emotional condition. The mean value in sad emotional condition was 3.50 cm H_20 which has increased to 7.59 cm H_20 ($F=35.40$, $p<0.05$) in anger emotional condition. Figure 5 shows the mean sub glottis pressure across to emotional condition.

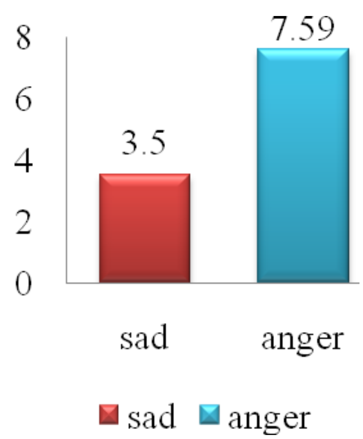

Figure 5: Mean sub glottis pressure across different emotions

The increased or raised subglottic pressure will produce high F0 or can presume that high F0 is associated with high subglottic pressure and low F0 is associated with low sub glottis pressure (Ladefoged, 1967; Stathopoulos & Sapienza, 1993).In the present study also we can find that there is a increase in both F0 and subglottic pressure anger emotional condition.

Isshiki (1961) suggests that minimal subglottic pressure required for phonation increases with the rise of F0, indicating that the resistance at the glottis to the air current increases with the rise of F0. On low F0 phonation, increased exhalation or blowing results in the slightly increased F0. At the register-boundary, increased exhalation lowers the F0 that may be attributed to the consequent change in vibrating mass of the vocal cords. Vibrating mass of the vocal cords during phonation is mainly dependent on the tension and shape of the vocal cord, both of which are controlled by the internal and external laryngeal muscles, but to some extent is also affected by the subglottic pressure especially at the so-called register-boundary.

It has been suggested that when the vocal folds are relatively thick and slack, Bernoulli forces and subglottal pressure have greater influence on the vibration pattern (Atkinson, 1978). F o has been found to be more sensitive to changes in subglottal pressure in the low F0 region than at higher frequencies (Stevens, 1974).

4.2. Comparison of vowels across emotional condition for F0 and subglottic air pressure

Comparison between sad and anger emotional condition shows that there is no significant difference between vowels seen in sad and anger emotional condition for F0. The results showed that the F0 was higher for vowel /u/ - 277.14 Hz, followed by /a/ - 256.34 Hz and /i/- 255.40 Hz in anger emotional condition. Similar trend was seen in sad emotional condition.

According to Meyer (1974) due to intrinsic frequency, the fundamental frequency for high vowels /i/ and /u/ can be higher than the low vowels /a/and /ae/.

Shadle (1985) had calculated the F0 in sentences for different vowels and had found that /u/ and /i/ had higher F0 values compared to /a/. The values were /u/- 225 Hz, /i/- 225 Hz and /a/- 215 Hz in sentence initial position and /u/- 175 Hz, /i/- 174 Hz and /a/- 175 Hz in sentence final position. In the present study also we can find a similar trend wherein the values were higher for /u/ followed by /i/ and /a/ in both emotional condition.

Waaramaa, Laukkanen, Alku & Väyrynen E (2008) have shown that vowels [a:], [i:], [u:] can carry emotional information in various ways, [a:] being the most effective and [u:] the least effective in this function. The spectral structures of vowels as such may be associated with certain emotional colourings of the voice: The open back vowel [a:] was shown to be a successful carrier of tenderness. The front vowel [i:] easily conveyed joy and also sadness. Vowel [u:] was likewise a successful carrier of sadness, and the least effective carrier of the positive emotions of tenderness and joy. Anger tended to be well recognized from each vowel mentioned. In the present study such a correlation could not be obtained.

Subglottic air pressure was higher in anger emotional condition. The vowel /a/- 7.37(1.37) cm H_2O had higher pressure compared to /i/- 7.62 (1.78) cm H_2O and /u/- 7.78(1.85) cm H_2O in anger emotional condition. According to Smitheran and Hixon (1981) normal speaker's air pressure ranges between 4 and 5 cm H_2O for normal conversations or moderate level phonations. For females in the age range of 17-30 years had an air pressure of 6.09 cm H_20 for the vowel /i/. The study by Isshiki, Nobuhiko (1961) suggest that when vowel production was controlled by seeing a sound level meter so as to keep equal phon level, the subglottic pressure was different according to the vowels with /i/ having the maximum pressure followed by /u/, /e/, /o/, /a/. Higher supraglottal impedance in high vowels, leading to slower discharge of air pressure, elevated larynx position in high vowels, effects of stress on tense vowel might be the reason for high pressure in /i/ vowel. In the present study we can see that there is no much difference seen between the vowels in emotional utterance. This could because of the lesser duration of the vowels in the production of emotional speech utterances.

In a study done by Demolin, Gerbers, & Hassid (2014) found a significant difference seen in sub glottal pressure in the production of vowels. The difference in subglottal pressure between /a/ and /i/ is 1.65 cm hPa and between /a/ and /u/ is 1.76 hPa and they attribute this to variations in the resistance provided by the vocal folds to the outgoing air (the glottal impedance) or to variations in the stiffness of the vocal tract walls.

In summary, the F0, CQ, SQ and subglottic pressure has increased in anger emotional condition OQ has decreased in anger emotional condition. Larynx tends to rise and pharynx tends to narrow during the production of tense voice quality. Thus it might lead to the increase in F0 and increased subglottic pressure in anger emotional condition. There is also a difference noted between vowels in sad emotional condition for subglottic pressure and it can be due to the difference height and placement of tongue during production of vowels.

CHAPTER V

Summary and Conclusion

The present study aimed to investigate few aerodynamic and electroglottographic parameters of primary emotional utterances (anger and sad). The data was obtained from 30 native typically developing female Malayalam speakers in the age range of 18-35 years.

The electoglottographic and aerodynamic data were obtained using electroglottograph, Model 6103 and Aeroview system version1.5. The participants were asked to produce three sentences in two different emotional conditions. The F0, Contact Quotient, Open Quotient and Speed Quotient were measured using EGG and Subglottic pressure was calculated using Aeroview system. The data was analyzed using SPSS 20. version.

The results revealed that there is a significant difference between electroglottographic and aerodynamic parameters in sad and anger emotional condition. F0 and subglottal pressure showed greatest difference between emotional conditions followed by CQ, OQ and SQ. The F0, CQ, SQ and Subglottic air pressure showed high value for the anger condition compared to sad condition. On the other hand OQ showed lower value for anger compared to sad emotional condition. Increase in F0 can be due to the increased tension on the extrinsic and intrinsic laryngeal muscle while producing anger emotional condition which can lead to increased F0. The increased or raised subglottic pressure will produce high F0 or can presume that high F0 is associated with high subglottic pressure and low F0 is associated with low sub glottis pressure.

Increased CQ values shows increased glottal adduction and stronger contraction of interarytenoids and lateral cricoarytenoid and increased medial compression from the activity of thyroarytenoid muscle where as the OQ has reduced in anger emotional condition. Comparison between the vowels revealed that only for sub glottis pressure there is a difference seen between the vowels. For F0, CQ, OQ, and SQ there was no much difference seen between the vowels.

In conclusion, there is significant difference seen between electroglottographic and aerodynamic parameters in sad and anger emotional condition. There is an increased subglottal pressure and F0 while producing anger emotions when compared to sad emotional condition and it showed a significant difference between vowels in sad emotional condition.

Clinical Implications

- The aerodynamic and EGG parameters can be used for differentiate between the emotions; however more investigations are required to estimate the magnitude of change.
- The current investigation could provide insight into glottal characteristics in emotional voice and variation in the acoustic and aerodynamic characteristics of voice.

Limitations of the study

- The present study addressed changes in aerodynamic and electroglottographic parameters for primary emotional utterances. However the results could not be generalized as the study considered only female subjects.
- Reliability test was not carried out for the emotional utterance.
- Only primary emotions sad and anger were taken for the present study.

Future Directions

- The future study can include more emotional condition.
- Both genders can be included in the future study.
- Perceptual judgment of the participant's utterance might reduce the error in the research.

References

Atkinson, J. E. (1978). Correlation analysis of the physiological factors controlling fundamental voice frequency. *The journal of the Acoustical Society of America*, *63*(1), 211-222.

Baken, R. J., & Orlikoff, R. F. (2000). *Clinical measurement of speech and voice*. Cengage Learning.

Bombien, L., Mooshammer, C., Hoole, P., & Kühnert, B. (2010). Prosodic and segmental effects on EPG contact patterns of word-initial German clusters. *Journal of Phonetics*, *38*(3), 388-403.

Cowan, M. (1936). *Pitch and intensity characteristics of stage speech* (Vol. 1). Department of Speech, University of Iowa.

Darwin, C., Ekman, P., & Prodger, P. (1998). *The expression of the emotions in man and animals*. Oxford University Press, USA.

Davitz, J. R.(1964). *A review of research concerned with facial and vocal expression of emotion.* The Communication of Emotional Meaning, edited by J. R. Davitz Ney York: Mc Graw Hill

Demolin, D., Gerbers, S., & Hassid, S. (2014). Vowel and consonant effects on subglottal pressure. *The Journal of the Acoustical Society of America,135*(4), 2294-2295.

Dromey, C., Stathopoulos, E. T., & Sapienza, C. M. (1992). Glottal airflow and electroglottographic measures of vocal function at multiple intensities. *Journal of Voice*, *6*(1), 44-54.

Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K.,.& Scherer, K. (1987). Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, *53*(4), 712.

Esling, J. H. (1984). Laryngographic study of phonation type and laryngeal configuration. *Journal of the International Phonetic Association*, *14*(02), 56-73.

Fairbanks, G., & Pronovost, W. (1939). An experimental study of the pitch characteristics of the voice during the expression of emotion. *Communications Monographs*, *6*(1), 87-104.

Fant, G. (1970). *Accoustic theory of speech production: with calculations based on X-ray studies of Russian articulations*. Mouton & Company.

Fitch, W. T. (2000). The evolution of speech: a comparative review. *Trends in cognitive sciences*, *4*(7), 258-267.

Gobl, C., & Ní, A. (2003). The role of voice quality in communicating emotion, mood and attitude. *Speech communication*, *40*(1), 189-212.

Holmberg, E. B., Hillman, R. E., & Perkell, J. S. (1988). Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *The Journal of the Acoustical Society of America*, *84*(2), 511-529.

Isshiki, N. (1961). Voice and subglottic pressure. *Studia phonologica*, 1, 86-94

Johnson-Laird, P. N., & Oatley, K. (1989). The language of emotions: An analysis of a semantic field. *Cognition and emotion*, *3*(2), 81-123.

Johnstone, T. (2001). *The effect of emotion on voice production and speech acoustics*. University of Western Australia.

Johnstone, T., & Scherer, K. R. (2000). Vocal communication of emotion. In M. Lewis & J. Haviland (Eds.). The handbook of emotion (pp. 220-235). New York: Guildford

Johnstone, T., Banse, R. and Scherer, K. R. (1995). Acoustic Profiles from Prototypical Vocal Expressions of Emotion. Proceedings of the XIIIth International Congress of Phonetic Sciences, 4, 2-5

Keating, P., Esposito, C., Garellek, M., Khan, S., & Kuang, J. (2010, July). Phonation contrasts across languages. In *Poster presented at the 12th Conference on Laboratory Phonology*.

Kitzing, P., & Sonesson, B. (1974). A photoglottographical study of the female vocal folds during phonation. *Folia Phoniatrica et Logopaedica*, *26*(2), 138-149.

Kramer, E. (1963). Judgment of personal characteristics and emotions from nonverbal properties of speech. *Psychological Bulletin*, *60*(4), 408.

Ladefoged, P. (1967). *Linguistic phonetics* (Vol. 6). Phonetics Laboratory, University of California.

Laukkanen, A. M., Vilkman, E., Alku, P., & Oksanen, H. (1997). On the perception of emotions in speech: the role of voice quality. *Logopedics Phoniatrics Vocology*, *22*(4), 157-168.

Lieberman, P., & Blumstein, S. E. (1988). *Speech physiology, speech perception, and acoustic phonetics*. Cambridge University Press.

Lieberman, P., & Michaels, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. *The Journal of the Acoustical Society of America*, *34*(7), 922-927.

Mooshammer, C. (2010). Acoustic and laryngographic measures of the laryngeal reflexes of linguistic prominence and vocal effort in Germana). *The Journal of the Acoustical Society of America*, *127*(2), 1047-1058.

Murphy P, Laukkanen AM.(2012) Investigation of normalised time of increasing vocal fold contact as a discriminator of emotional voice type. In *Multimodal signals: Cognitive and algorithmic issues,* 90 – 7

Murphy, P. J., & Laukkanen, A. M. (2009). Electroglottogram analysis of emotionally styled phonation. In *Multimodal signals: Cognitive and algorithmic issues*, 264-270.

Murray, I. R., & Arnott, J. L. (1993). Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion. *The Journal of the Acoustical Society of America*, *93*(2), 1097-1108.

Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America*, *53*(6), 1632-1645.

Scherer, K. R. (2000). Psychological models of emotion. *The neuropsychology of emotion*, *137*(3), 137-162.

Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech communication*, *40*(1), 227-256.

Scherer, K. R., Schorr, A., & Johnstone, T. (Eds.). (2001). *Appraisal processes in emotion: Theory, methods, research*. Oxford University Press.

Shadle, C. H. (1985). Intrinsic fundamental frequency of vowels in sentence context. *The Journal of the Acoustical Society of America*, *78*(5), 1562-1567.

Skinner, E. R. (1935). A calibrated recording and analysis of the pitch, force and quality of vocal tones expressing happiness and sadness; and a determination of the pitch and force of the subjective concepts of ordinary, soft, and loud tones. *Communications Monographs*, *2*(1), 81-137.

Smitheran, J. R., & Hixon, T. J. (1981). A clinical method for estimating laryngeal airway resistance during vowel production. *Journal of Speech and Hearing Disorders*, *46*(2), 138-146.

Stathopoulos, E. T., & Sapienza, C. (1993). Respiratory and laryngeal function of women and men during vocal intensity variation. *Journal of Speech, Language, and Hearing Research*, *36*(1), 64-75.

Stevens, K. N. (1974). Theory of Vocal-Cord Vibration and Its Relation to Laryngeal Features. *The Journal of the Acoustical Society of America*, *55*(2), 383-383.

Titze, I. R. (1990). Interpretation of the electroglottographic signal. *Journal of Voice*, *4*(1), 1-9.

Von Leden, H., Moore, P., & Timcke, R. (1960). Laryngeal vibrations: Measurements of the glottic wave: Part III. The pathologic larynx. *AMA Archives of Otolaryngology*, *71*(1), 16-35.

Waaramaa T , Kankare E. (2013). Acoustic and EGG analyses of emotional utterances. *Folia Phoniatrica et Logopaedica*; 38 : 11 – 8 .

Waaramaa T, Laukkanen A-M, Alku P, Vä y rynen E. Monopitched expression of emotions in different vowels. *Folia Phoniatrica et Logopaedica*. 2008;60:249 – 55.

Waaramaa, T., Palo, P., & Kankare, E. (2015). Emotions in freely varying and mono-pitched vowels, acoustic and EGG analyses. *Logopedics Phoniatrics Vocology*, *40*(4), 156-170.

Williams, C. E., & Stevens, K. N. (1972). Emotions and speech: Some acoustical correlates. *The Journal of the Acoustical Society of America*,*52*(4B), 1238-1250.