SPEAKER'S FORMANT: AN INDICATOR OF EXPRESSIVE SPEECH IN SOME GROUPS OF PROFESSIONAL VOICE USERS

Register No: 05SLP004

JOHNSI RANI

A Dissertation submitted in part fulfillment for the degree of Master of Science (Speech-Language Pathology)
University of Mysore, Mysore.

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

APRIL 2007

Dedicated to
"Amma-Appa
& my guide"

Certificate

This is to certify that this Dissertation entitled **"Speaker's Formant: An indicator of expressive speech in some groups of professional voice users"** is a bonafide work in part fulfillment for the degree of master of (Speech Language Pathology) of the student (Registration No.05SLP004). 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.

Mysore
April, 2007

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

Certificate

This is to certify that this dissertation entitled **"Speaker's Formant: An indicator of expressive speech in some groups of professional voice users"** has been prepared under my supervision and guidance. It is also certified that this dissertation has not been submitted earlier in any other University for the award of any Diploma or Degree.

Mysore
April, 2007

Ms. K. Yeshoda
Lecturer,
Department of Speech-Language Sciences,
All India Institute of Speech and Hearing,
Mysore-5 70006.

Declaration

This Dissertation entitled **"Speaker's Formant: An indicator of expressive speech in some groups of professional voice users"** is the result of my own study under the guidance of Ms. K. Yeshoda, Lecturer, Department of Speech-Language Sciences, All India Institute of Speech and Hearing, Mysore, and has not been submitted earlier in any other University for the award of any Diploma or Degree.

Mysore
April, 2007

Register No. 05SLP004

ACKNOWLEDGEMENT

I sincerely thank my guide, **MS. K. Yeshoda,** *Lecturer, Department of Speech Sciences, AIISH, for her support and guidance. "Thank you Ma 'am, for your valuable suggestions, unending patience, words of encouragement and motivation throughout the study". You have always been and will be a source of motivation for me, whether be it this study, or any other work. You have made me realize my potentials. Mam, I want to thank you from bottom of my heart.....*

I would like to thank **Dr. Vijayalakshmi Basavaraj,** *Director, AIISH, Mysore, for permitting me to carry out this dissertation.*

It would have been impossible to get through the tough times without you Amma. Your cheerful disposition, calming words and incorrigible giggle have made my life. Thank you so much, for making me realize how important it is to face every tough moment with a smile... and building a great confidence in me.....

Heartfelt thanks to you ***Appa,*** *for teaching me important lessons about life and encouraging me to meet the world with courage and determination. I know I am going to remain your little girl always, but no matter how grown up I may be, I will never outgrow my love for you...*

Mannar, Vetri *&* ***Vitta,*** *I wonder what I would have done without you being there to give me your unbiased suggestions and unconditional love,..... thanks for your valuable guidance, love, care and faith in me..... you all are more than brother and sister's to me..... I am lucky to have brother and sisters like you.....LUV you all*

One person without whom I would not have done anything in my life..... no words to express my feeling for u..... thanks for everything C. K..... be the way you are.....

I express my sincere gratitude to ***all my teachers and staff at AIISH,*** *specially* ***Savithri mam,*** *for laying the foundation for us, imparting knowledge and for moulding us into what we are today.....thank u mam.....*

A special thanks to ***Vasanthalakshmi Mam*** *for all the help rendered during for the statistical analysis.*

A special than to ***Director of Rangayana*** *for permiting me to carry out the study also wanna thank* ***Mr. Ravindra Kumar (AIR)*** *for their valuable help and support.....I wish to thank all my subjects for their extreme cooperation, patience and help, without which, this study would never have been possible..Also a special thanks to* ***all staff who gave their valuable time for being as Judges for this study....***

Special thanks to ***Suma,*** *Your support helped me to be strong and confident in life and deal with the hardships of life.....will miss uyaar.....*

Suma, Rahana, Priya n Supraja*.....how can* ***I*** *forget the wonderful time we spent together at AIISH.....thanks guys for all the moral support, encouragement and I wish you all the best in your endeavors.*

A special thanks to ***Kalai mam, Ananthi mam,& Banu mam, Rani, Powlin, Deepa, Sangitha, KD, J.K, Ram Mohan, n Venugopal sir,*** *for all the timely help rendered during my dissertation and RP work...*

Wanna thank all ***my seniors, who have been a source of motivation for me, I have learnt lot from u all*****.....** *I will truly treasure the time we spent together and it is going to be a part of my 'best memories at AIISH'...I will miss you all.....Thanks a lot...*

Wanna thank all my friends, ***Nambi, Swapna, Radheesh, Anjali, Savitha Anitha, Rohini, Sister, Marly n,****.....thank you for all your help.*

Wanna thank each of my ***classmates*** *(MSc &* ***BSc)*** *.....Thank you all for being such great friends and being so willing to help me when I needed it. I hope all of you have a nice life ahead.....*

I also want to thank all my Juniors, ***specially Padma and Akka, Shijitha, Sangeetha, Gargi,*** *for their constant help. Will miss u guyz.....all the best for future.....*

The ***library staff*** *merits special thanks including* ***Mahadevappa, Lokesh Sir and Nanjunda sir , Raju*** *too, for being very kind and helpful. Last but not the least, thanks to* ***Shivappa*** *and* ***Co.****...for the endless Xeroxes and printouts.*

For you....my strength, my soul but most importantly my friend.... Without you, I am incomplete ....I thank god for you....

Table of contents

List of tables

CHAPTER I

INTRODUCTION

It is not a powerful instrument, this human voice, nor a technically perfect one. A trumpet can blow louder, a violin can play faster, and oboe can spin longer melodies. Still it is unsurpassed in expressivity, depth and soulfulness of tone, truly the queen of instruments (Brodnitz, 1988).

The one faculty that sets man apart from all living organisms - that makes him unique in his ability to think and communicate is 'language'. The one form of communication which man uses effectively in interpersonal relationship is 'speech'. With it he gives form to his inner most thoughts - his dreams, ambition, sorrows and joys, without it, he is reduced to animal noises and empty gestures. In real sense, speech is the key to human existence (Fisher, 1966).

In this regard, considerable effort has been expended over the years by a number of investigators who have attempted to identify the parameters that define this most thrilling of all instruments, the human voice.

The human voice can be, and often is, a most responsive, flexible and, infinitely, variable sound producing instrument, especially when the voice is produced by the artistic level singer or actor, as the art of voice production is a learned behavior. The differential level of learned behaviors for voice production by trained voice users is why the professional voice is so interesting to study, and, for the sake of defining its extraordinary character, it often is compared to the untrained voice (Rothman, Brown & LaFond, 2002).

Anderson (1961) indicated that the superior voice should exhibit: adequate loudness, a clear, purity of tone, pleasing/effective fundamental frequency, an ease and flexibility, a vibrant and sympathetic quality, clearness and ease in diction.

Wilson (1972) opines that a good voice should have the following characteristics: pleasing voice quality, proper balance of oral and nasal resonance, appropriate loudness, a modal frequency level/habitual level suitable for age, size and sex, appropriate voice inflection involving pitch and loudness.

Snidecor (1951) has compared the voice samples of six carefully chosen superior female speakers and compared with the voices of superior male speakers reading the same materials ("rainbow passage") and found the following results; Pitch levels for female voices were found to be placed appropriately 2/3rd of an octave above the pitch levels of male voices. The median pitch levels for each group of subjects were located within the limits of less than 11/2 tones which suggest that the preferred pitch levels for superior speakers may fall within relatively narrow limits. Total pitch ranges were approximately equal in both male and female groups. Voices of females were found to be less variable in pitch than those of males. Compared to the voices of males, voices of females were found to have a less rapid mean rate of pitch change and fewer changes in the direction of pitch per unit of time. Both male and females used slightly slower oral reading rates compared to the median of randomly selected subjects.

Professional Voice Users

Quality is the texture or timbre of a sound or tone that distinguishes it from another tone having the same pitch, duration, and loudness. Voice, vocal quality more precisely is a reliable indicator of the physical, psychological as well as the emotional well being of a person. More number of people are becoming aware of importance of voice and maintenance of superior voice quality to prove oneself in their profession.

According to Stemple (1993), professional voice users are those individuals who are directly dependent on vocal communication for their livelihood. Professional voice users, especially those in speaking profession, require certain qualities in their speech to be successful. Melody, fluency, phrasing, modulation and good breath-support and control have to be superior to maintain speech for professional purposes. The professional voice users are grouped into different category based on training (formal: singers or informal: voice over artists) duration of voice use in a specified manner: teacher and actors etc..

Voice over artists are usually referred to as Broadcasts or AIR announcers. Announcers must have a pleasant and well-controlled voice, good timing, excellent pronunciation, and correct grammar. The most successful announcers attract a large audience by combining a pleasing personality and voice with an appealing style. AIR announcers introduce and close programs, announce station program information, breaks for commercials and public service information. They may also present news, sports, weather, time, and commercials. They also interview guests and moderate panels or discussions.

Actors are professionals who rely on voice to convey a repertoire of emotions that complement their performance; they must adapt their quality of voices to suit a variety of roles they assume on stage. These professional voice users have to have a sound knowledge or understanding of their voice, how it works, its sustenance power and versatility.

Bele (2006) investigated dimensionality in voice quality, the effect of different loudness levels on voice quality, for normal and supranormal male voices, presented by a group of professional voice users consisting of teachers and actors. The results revealed that four factors at both normal and loud levels of text reading, overlapped and helped in grouping of the perceptual characteristics. The factors were labeled according to their most characteristic feature. At normal level, the factors were 1) variation and sonority, 2) irregularity 3) degree of noise, and 4) phonatory effort. At loud level, the factors were 1) variation and sonority 2) degree of noise, 3) ringing voice quality 4) phonatory effort.

The results also suggest that sonority, variation in loudness and breathiness were perceived as different and as well as pressedness and the strength of the voice in terms of the acoustic energy it contains. Further it was inferred from the results that text reading at the normal level suggested that the basic dimensions for perceiving voice were pretty much in line with the GRBAS scale.

"Speaker's Formant" (SPF)

The Singer's Formant is associated with 'good voice quality' in opera singing but, more specifically the perception of an opera singer's 'brilliance' (Vennard, 1967),

or 'ring' (Ekholm, Papagiannis & Cagnon, 1998). Bartholomew (1934) was the first to report on the singer's formant, or the singing formant, and later Winckel (1953) commented on the high spectrum level at 3 kHz in spectra of singers' voices.

Sundberg (1990) summarized some of the characteristics specific to singers as: the breathing habits of singers being superior, their phonatory habits being special in that pitch and loudness are independent; they are also special with respect to the distribution and arrangement of formant frequencies. One such arrangement of the formant frequencies within singers voice is the "spectral prominence" in the frequency region between 2500Hz and 3400Hz approximately, which aids in voice projection and enhancement of the voice quality. This region has been termed the "Singer's Formant (Fs)" and helps singers to be heard above the orchestra. The exciting ringing quality of the professional singing voice, i.e., the quality that gives an arresting "edge'* to the voice is also attributed to the presence of "Singer's Formant (Fs)".

It is reported that Singer's Formant is influenced by individual voice types and ranges, the vowels attempted, the pitch and amplitude produced, and individual physiology (Morris & Weiss, 1997). Studies have shown that the precise location of the Fs varies. The Fs is most likely also associated with the projection requirements for stage speakers. If so, it could also rightly be called a Speaker's Formant (Nawka, Anders, Cebulla & Zurakowksi, 1997).

An analogy to Singer's Formant, the Speaker's Formant is expected to be a local maximum energy that is independent from the uttered vowel. Further it should

be more evident in the voice of such professionals as actors or in the voices of other trained speakers who are not only paradigmatic in articulation but also in sound quality (Nawka et al., 1997)

The Speaker's Formant in male voice is located in the critical band between 16 and 17 barks, with borders of 3,150 to 3,700 Hz and a center frequency of 3,400 Hz. Normal male voice also shows a peak in this frequency region, but the peak is less distinctive. Pathologic voices do not bear an energy peak at this part of the spectrum. Nolan (1983) pointed out that this phenomenon was similar to the Singer's Formant. Leino (1993) found a peak around 3,500 Hz as a differentiating feature of good voice quality and named this peak the "Actor's Formant". He proposed the term "Actor's Formant" or "Speaker's Formant" for the grouping of the third, fourth, fifth formants (F3, F4, and F5) at around 3.5 kHz in projected voices of male actors.

The Speaker's Formant is expected to be a strong peak at about 3.5 kHz, found in better speaking voices, which implies that it should be more evident in the trained voices of professionals, such as actors. A limited answer to question as to whether there is an SPF would be to say that a sufficiently small level difference between the strongest spectral peak and the strongest component in the range 3-4 kHz is needed. However, the amplitude of all spectrum overtones is dependent not mainly on how closely the higher formants are clustered, but on the slope of the source spectrum. As for prominence of SPF peak, most likely one has to take the distance between the formants into consideration, as the valleys between the formant peaks actually increase twice as much as the amplitude of the formant when the frequency distance between two formants is halved. It seems appropriate to talk about an SPF whenever

we see a clear peak in Long-Term Average Spectrum at 3-4 kHz in male voices and when this peak is present in individual vowels which is not an occasional phenomenon. It is not necessary for the voice to be good, although typically the phenomenon is observed in better voice quality Bele (2005).

There are speculations about the origin of the "Speaker's Formant" peak. Spectrograms suggest that it simply represents the regular fourth formant. This formant is likely to appear as a clear peak in an LTAS curve provided three conditions are fulfilled. First, the formant must be reasonably stable in frequency. Second, its bandwidth must be fairly narrow. Third, the voice source must produce partials as its frequency. As a resonance phenomenon, it may be generated by a specific glottal configuration where the larynx, isolated from the rest of the vocal tract, would act as an independent resonator, whose resonance frequency would be around 3 kHz.

Need for the study

The ultimate tasks of scientific research are to produce new knowledge thus promoting a deeper understanding and to scatter this new knowledge to society. But most of the studies among professional voice users have been conducted on singers, and teachers. The voice of other elite vocal performers i.e., actors, radio and TV artists are being researched to understand the characteristics of speaking voice.

Although much is known, about singers formant there is either insufficient evidence in the literature, or lack of consensus, or the results are inconclusive, nonexistent, or contradictory with respect to Speaker's Formant. Within the topic of Speaker's Formant there are few studies which correlate the perceptual and acoustic

characteristics of voices in professional voice users leaving so many questions to the speech language pathologist. Review indicates that most studies are on Western population. There is no study on speaker's formant in professional voice user in Indian context. Actors and radio announcers are elite professional voice users as their voices have to encompass certain minimum characteristics to fulfill the requirements of their jobs. Hence this study is envisaged to understand the characteristics of their voices and investigate presence / absence of Speaker's Formant in the two groups of professional voice users: actors and voice over artists.

CHAPTER II

REVIEW OF LITERATURE

The human larynx is a living complex which no instrument is able to imitate, and although the voice is born of a simple, dynamic mechanism, it is not enough to explain its intimate essence which must be sought in every essence of life. Vocal performance has always been of great interest. In ancient Greece, the rhetoric capacity was idolized. The voice has an infinite variety of expressions and is the living material which serves us generously, but accurate diction completes the voice. *Expressive diction* is the soul and life of the word, as the intent look animates and fulfills the potential of the eye. A beautiful voice which may be acoustically harmonious doesn't have much value if it is not subjected to the spiritual and psychic requirements of the word, whether or not the word is potentiated by music. Although singing draws expressive power from the musical requirements which restrict it to a precise rhythm, the speaking voice lives only by its own internal force and intrinsic power, and must draw even more upon the elements of persuasion and commotion (Meano, 1967).

The complex characteristic of human voice has been the objective of a wide range of human interest, including scientific research. Now a days communication skills are important to almost everybody as our society has become more service - oriented. Consequently, this growing impact has induced a lot of research concerning voice assessment and voice training with a comprehensive battery of tests focusing on qualitative and quantitative aspects of vocal performance. There is an increasing segment

of the population which is dependent on vocal communication for its livelihood like actors, singers, lawyers, cheer leaders, etc. They are referred to as professional voice users. Murry and Rosen (2000) define professional voice users as those who require the use of their voice to maintain income.

In work with professional voice users with differing demands on the voice and in relation to research on normal and supranormal voice quality, there is need for a method of perceptual evaluation of normal and supernormal/resonant voices. The term "normal voice" here is related to ordinary speaking voices that are not dysfunctional, and "supranormal voices" are associated with more or less resonant voices.

One of the greatest difficulties found while evaluating voice is the judgment of its quality through a perceptual analysis, that although being the gold standard, involves socioeconomic and cultural aspects as well individual preferences (Medrado, Ferreira & Behlau, 2005; & Bele, 2005). There are several adjectives used in the perceptual evaluation, and the methods applied, due to the natural subjectivity of this process, lead to disagreements among the listeners and to difficulties to reach a consensus on the use of this or that terminology (Bele, 2002). In such context, the acoustic analysis has made possible to guide and complement the speech-language treatment with more objective data.

According to Master, Biase, Pedrosa and Chiari (2006) among various possibilities of acoustical analysis, the long-term average spectrum (LTAS) has been used

in several studies, because it allows "quantifying" the quality of a voice, marking the differences of gender, age, professional voices and dysphonic voices, contributing not only to the evaluation but also to follow-up training or treatment. The LTAS, the mean of several simultaneous spectra, reflects both glottal source and vocal tract filter characteristics in the quality of a voice. It is a particularly useful method because it allows the study of persistent long-term voice factors such as quality. In shorter duration samples, various short-term variations (e.g. due to change of phonemes) would obscure the study of voice quality.

Within the several possibilities of spectrographic analysis, the Long-term average spectrum (LTAS) offers the possibility of quantifying the quality of a voice, pointing the differences between gender, age, professional voices - talked and sung - and dysphonic voices (Leino, 1993; Mendoza, Valencia, Munoz, & Trujillo,1996; Barrichelo, Hener, Dean, & Sataloff, 2001; Haiti, Hans, Vaissiere, Riquet & Brasnu, 2001; Linville & Rens, 2001; Bele, 2002; Camargo, 2002; Sjolander, 2003; Laukkanen, Syria, Laitala, & Leino, 2004; Pinczower & Oates, 2005).

Some particular features of an emission are more stable, such as the voice quality, and become more evident in extended speech samples; this is precisely one of the greatest advantages of using the LTAS (Camargo, 2002). Another advantage is that if the acoustic signal is long enough, the resulting mean spectrum is not affected by differences in the speech sample - subject matters and articulation - indicating a certain degree of

reliability in the comparison between speakers and between studies (Frokjaer- Jensen & Prytz, 1976, Kitzing, 1986, Lofqvist, 1986).

Pinczower and Oates (2005) opine that effective vocal projection is essential for performers like actors and public speakers, etc, making it possible for their voices to be heard by listeners with maximum intelligibility with minimum vocal effort.

Limited literatures are available on voice over artists. Deepa (2004) studied voice and speech characteristic of radio jockeys. Five radio jockeys and five age matched normal subjects participated and the material consisted of sustained vowels /a/, *I\l* and /u/ and five minutes conversation sample. Sustained voice samples were analyzed using MDVP and 31 voice parameters were extracted. In speech frequency range was extracted. The study also involved perceptual experiment. Eight speech language pathologists rated the speech sample on the parameters phrasing, emphasis, modulation and overall performance on a 3 point rating scale. Vital capacity and fluency were also assessed in all the subjects. The results indicated that radio jockeys were superior in most analyzed parameters.

According to Meano (1967) the expressive verbal power of actors who wish to give life to an emotion-to create a mood- is unquestionable pre- eminent in an outstanding stage production which cannot be limited only to the scenery, costumes, or gestures of the characters whom the actors wish to create. The projection skills of actors' performances are of particular interest. For expression to reach its maximum potential,

the actor, singer or speaker must communicate to others the reflection of the emotions which stir within them. To express these emotions effectively, the performers must never lose themselves completely in the sentiments or emotions they wish to convey. They must control projection in objective awareness of its emotional impact while they feel and listen to the created mood inside themselves. In other words, the control is regulated by restraint, dominated by a subtle sense of detachment.

Stage actors also must project their voices over background interference, such as crowd noise, battle sounds, and background music. Actors must be able to adjust their vocal production to their performance space, whether it is a 100- seat theater or a 3000-seat auditorium, while maintaining the intonation range necessary for expressing a wide range of emotions. One of the basic demands made of the actor's voice is the 'test of audibility'. The actor should be able to project without sacrificing the intonation range needed for emotional expressiveness (Acker, 1987).

Kovacic and Budanovac (2002) studied the acoustic characteristics of adolescents actor's and non- actor's voices. The experimental sample consisted of 10 actress and 10 actors, while the control sample included 13 girls and 14 boys. The speech sample consisted of the vowel /a/ in sustained phonation, reading, and spontaneous speech. The results revealed that greater Fo, jitter and shimmer indicated that the acoustic characteristics of the voices of both actress and actors varied more than those of their peers.

Speaker's Formant

In the speaking voice, Leino (1993) relates a peak at 3.5 kHz, the "Speaker's Formant" or "Actor's Formant", and a relatively gentle spectrum envelope declination with the perception of good voice quality in male Finnish actors. With very good voices, the level of the actor's formant from -15 to -25 dB, compared to the strongest component in the spectrum (typically the Fl region below 1 kHz), and the level difference between the peak and surrounding valleys may be much as 10 dB. Thus, the "Actor's Formant" is located about 1kHz higher than the Singer's Formant and weaker in intensity. According to Leino (1993) the Actor's Formant is not an "absolute pre - requisite" for a good voice, but rather a tendency, because some voices judged as very good had a weak actor's formant, whereas some of the poor voices presented strong Actor's Formant.

Nawka et. al., (1997) further identified in the voice spectrum envelope of German actors, an energy increases between 3.150 and 3.700 Hz that would be related to a "loud and shiny" quality of voice and a more gentle fall of actor's voice spectral curve in habitual and strong intensities when compared to normal and moderately rough voices.

Nawka et. al., (1997) in a study of German actors, 5 normal healthy men, 5 patients with hoarseness found statistically significant increase in mean intensity among all three groups of speakers in the F4 region. He reported again that the speaker's formant in male voices is located in the critical band between 16 and 17 Barks, with a center frequency of 3.400 Hz. Normal male voice also show a peak in this frequency region, but peak is less distinctive. Again he reported that the Speaker's Formant is

connected with the sonorous quality of the voice. It increases gradually and is approximately 10 dB higher in professional male voices than in normal male voices at neutral loudness (60 dB).

Both Leino (1993) and Nawka et. a1, (1997) concluded that the speaker's formant is a characteristic of male speaking voice resulting from the voice source and the resonating properties of the vocal tract, and that, this phenomenon is associated with good male voice quality. This implies that it should be more evident in the trained voices of professional, such as actors.

A pilot study was conducted by Raphael and Scherer (1987) in which acoustic analyses were made of actors' voices. Spectral comparisons were made between two types of voice production: the "call" technique as taught by Lassac (1987) and a mode of phonation more like conversational speech. Any differences found between stage actors normal conversational voice and their vocal production for stage might be helpful in the teaching of voice production to stage actors more effectively.

Lassac's "call" technique (1993) appears to attempt to combine in a specific way the laryngeal tone and the voluntary shaping of the vocal tract, especially in regard to the amount of space between the teeth, the degree of stretch in the cheek muscles, and the size and shape of the lip opening to enhance vocal resonance and projection.

It is speculated that Speaker's Formant varies with loud/projected voices. Pinczower and Oates (2005), compared male actor voices on comfortable loudness and in maximum projection level and subjected these voices to acoustic and perceptual analyses. Results emphasized that the spectrum showed higher energy concentration in higher frequencies region around 3.4kHz for the strong emission than the for the ones in comfortable conditions.

Bele (2002), compared Norwegian actors and teachers voices and observed the following differences in the LTAS: actors have emission mechanisms in the strong intensities and, therefore, smaller values in the relation between Fo and F1, the "Speaker Formant" region is stronger for the actors but not as much as referred in the literature. The Actor's Formant seemed to be related to both sufficiently strong overtones and a glottal setting allowing for a lowering F4 and proximity of F3 and F4.According to the author, the perceptual evaluation was more efficient than the LTAS in the differentiation of these voices, leading to the following question: something affects our subjective judgments of vocal quality, something that cannot be objectively measured. The author observes that a peak at 3.5 kHz could also be related to nasal, rough and fry voices, reinforcing the necessity to consider the perceptual analysis analyzing with the LTAS.

Pinczower and Oates (2005) studying Australian actors voices, explored whether acoustic and percepetual features could distinguish a comfortable projected voice from a maximally projected voice. They observed increased acoustic energy in the higher (2-4 kHz) frequency region compared to a lower (0-2 kHz) frequency region in the LTAS of

voice samples when maximal projected than in the comfortably projected voice conditions. These finding offered some preliminary support for the existence of an actor's formant during maximal projection. The authors reported that in the perceptual analysis, the voice emitted in maximal projected voice were evaluated being as more projected and tense, and that there was a correlation between the decline of the spectral curve and perception of projection and tension.

Bele (2005): studied the central concept, the Speaker's Formant (SPF) related to the perceptual characteristics "better normal voice quality" (BNQ) and "worse normal quality" (WNQ). The acoustic analyses done were LTAS and spectrographic measurements of formant frequencies. At very high intensities, the spectral slope was rather quadrangular without a clear SPF peak. The trained voices had a higher energy level in SPF region compared with the untrained, significantly so in loud phonation. The SPF seemed to be related to both sufficiently strong overtones and a glottal setting, allowing for a lowering of F4 and a closeness of F3 and F4 than worse normal quality voices; a finding that was supported by a smaller Ll(3-4kHZ) difference in actors' voices than in teachers' voices. As for the difference between the peak and the preceding valley in BNQ voices with a prominent SPF peak, this was relatively large in this study, visual inspection revealing about 5- 8 db for some of the LTAS.

Leino (1993) however, reported that the level difference between the strongest peak of the spectrum and the SPF was generally 15-25 dB for the good voices and almost 30 dB or more for the poor voices. At the same time, as the peak becomes stronger, the

valleys around it become deeper; the difference between the peak and the preceeding valley was often more than 10 dB in this particular study.

Thus, the review suggests that the concept of Speaker's Formant is relatively popular with more number of studies being done. But the results of these studies reveal that there are lot of issues pertaining to speaker's formant that still needs to be addressed i.e. type of professional voice user groups, types of training, gender, years of experience, etc, because the results could be influenced by all such factors. Therefore, an attempt has been made to investigate whether Speaker's Formant is present in the voices of professional voice users in India. And also to find out if the presence of speaker's formant and perceptual analysis correlate.

Aim of the study

a) To determine presence of speaker's formant using acoustic analysis in two professional voice user groups; theatre artists and voice over artists (AIR announcers).
b) To investigate whether speaker's formant is present in all the tasks conditions.
c) To compare acoustic and perceptual characteristics.
d) To determine perceptual correlates of 'good speech' in the two groups of professional voice users, which could be indicators of expressiveness.

CHAPTER III

METHOD

Subjects: Thirty seven subjects consisting of two groups of professional voice users (theatre artist and voice over artists) in the age range of 18 to 40 years participated in the study. Among them, 20 subjects were theatre artists, 17 subjects were voice over artists (AIR announcers). Among theatre artists, equal number of males and females were taken i.e. 10 males and 10 females. Among voice over artists, 7 subjects were males and 10 subjects were females. All subjects fulfilled the following criteria;

1. Subjects who don't have any problem in speech, hearing, voice or neurological disease.
2. Subjects who were native speakers of Kannada with proficiency in reading in Kannada.
3. The purpose of the study was explained to the subjects and their consent was taken.

Now onwards in the study voice over artists will be referred to as AIR announcers.

Recording procedure

Participants were given appropriate instruction to perform the following tasks,

1. Phonation of/a/
2. Reading a standard Kannada passage. "Bangaloru namma. . . ." (Developed and standardized at AIISH)
3. Speaking for about 1 min (monologue).

Recording was done by using a high fidelity portable digital mini-disc recorder (Sony MZ-R30). The microphone was placed at a distance of 5"-6" inches from the speaker's mouth. The recording was carried out in a quiet environment.

Analysis

Analysis was carried out in two phases;

I. Acoustic analysis

II. Perceptual analysis

Phase I: Acoustic Analysis

The recorded samples i.e., phonation of /a /, reading, speaking samples were line- fed using 16 kHz sampling frequency. All the samples were subjected to long-term average spectrum analysis (LTAS) of VAGHMI Software of Voice and Speech System, Bangalore.

Extraction of Speaker's Formant

Each sample was displayed as a spectrum. The frequency range of the spectrum was 0-16 kHz. Speaker's Formant is a prominent spectrum envelope peak near 2800 - 4200 Hz. The spectral position of the Speaker's Formant was noted as the highest partial or the mid point of 2 highest partials in the speaker's formant region.

Phase II: Perceptual analysis

Subjects

Five qualified female speech language pathologists in the age range of 25-45 years were considered as judges for the perceptual analysis. These professionals had minimum of five years of clinical experience.

Procedure

The judges were asked to listen to the audio recorded of all subjects speech sample, individually and rate each of the task, i.e. speaking and reading separately. The samples could be heard as many times as possible by the judges. They were provided with score sheets containing the parameters to be rated. The judges were asked to rate the speaking and reading samples using a three point scale. 1= Below Normal, 2= Normal and 3= Excellent. The score sheet for perceptual analysis is given in Appendix 1.

CHAPTER IV

RESULTS

A total of thirty seven subjects consisting of both groups of professional voice users (Theatre artist and AIR announcers) participated in the study. Among them, 20 subjects (54.05 %) were Theatre artists and 17 subjects (45.94 %) were AIR announcers. In Theatre artists, equal number (10 each) of males and females were present but in AIR announcers, 7 subjects were males (41.17 %) and 10 subjects were females (58.82 %).

Aims of the study were

1. To determine presence of Speaker's Formant using acoustic analysis in Theatre artists and voice over artists (AIR announcers).
2. To investigate the presence of Speaker's Formant in all the task conditions.
3. To compare acoustic and perceptual characteristics.
4. To determine perceptual correlates of 'good speech' in the two groups of professional voice users, which could be indicators of expressiveness.

Results will be discussed under the following headings

I. Acoustic analysis
II. Perceptual analysis
III. Correlation between perceptual and acoustic analysis
IV. Perceptual correlates of 'good speech' and presence of Speaker's Formant in both groups of professional voice users.

I. Acoustic analysis

Presence of Speaker's Formant in all the three task conditions i.e. phonation, reading and speaking in the total number of subjects (in percentage) is shown in Table 1. When both group of subjects (Theatre artists and AIR announcers) were considered together, the subjects could be divided into two groups on the basis of presence or absence of Speaker's Formant. Group I consisted of subjects in whom Speaker's Formant was present in all the three task conditions and Group II consisted of subjects in whom Speaker's Formant was absent in any one of the task condition.

Groups		Task Conditions (in %)		
		Phonation	Speaking	Reading
Group I	Presence of SPF(24)	64.86	64.86	64.86
Group II	Overlapping subjects (13)	61.53	38.46	69.23

Table 1: Subjects with presence of Speaker's Formant in three task conditions

In Group I, 64.86 % of subjects had Speaker's Formant in all the task conditions. In Group II, Speaker's Formant was found in 69.23 % of subjects in reading task condition only, 61.53 % of subjects in phonation task only and 38.46 % of subjects in speaking task condition only.

The presence of Speaker's Formant was compared across the task conditions in all subjects irrespective of the professional voice users groups (Table 2).

Conditions	Mean	Standard Deviation	F	Sig (2- tailed)
Phonation	3305.75	473.68	11.86	0.000
Speaking	3898.08	490.15		
Reading	3653.25	510.37		

Table 2: Mean, Standard deviation of Speaker's Formant across three task conditions.

Higher mean values were obtained for speaking followed by reading and phonation. Repeated measures of ANOVA revealed significant difference between the task conditions [$F(2, 46) = 11.86$, $p < 0.001$]. From the Bonferroni test of pair wise comparisons it was found that phonation was significantly different from reading ($p < 0.001$) and speaking ($p < 0.05$).

The professional voice user groups were studied individually for the tasks and the results are depicted in Table 3.

Conditions	Groups	Mean	Standard Deviation	t
Phonation	TA	3259.16	334.87	0.474
	AIR	3352.33	593.52	
Speaking	TA	3836.50	490.59	0.583
	AIR	3959.66	543.76	
Reading	TA	3682.08	503.15	0.282
	AIR	3624.41	497.35	

Table 3: Mean, Standard deviation and't' values of Speaker's Formant in three task conditions across professional voice user groups.

When both the professional voice user groups were compared separately for presence of Speaker's Formant and across task conditions using one-way repeated measures of ANOVA, significant difference was found. From the Table 3, it can be

observed that the mean value of the Speaker's Formant was highest for speaking followed by reading and finally phonation across the three task conditions in both groups of professional voice user. Independent t- test revealed that there was no significant difference ($p > 0.05$) across all the three task conditions.

Further, an attempt was made to understand the gender wise distribution of subjects in both groups for presence of Speaker's Formant. Table 4 shows the distribution of subjects with presence of Speaker's formant in both theatre artists and AIR announcers.

Groups	Males	Females
Theatre artists	60	60
AIR announcers	85.71	60

Table 4: Presence of Speaker's Formant in Theatre artists and AIR announcers (in %)

In Theatre artists, 6 subjects, 60 % each in males and females revealed Speaker's Formant. In AIR announcers 6 males (85.71 %) and 6 females (60 %) had Speaker's Formant. Overall, a total of 64.86 % of subjects in both groups were found to have Speaker's Formant in all three task conditions.

Table 5 shows the mean and standard deviation of Speaker's Formant in three task conditions for both professional voice user groups.

Groups	Conditions	Mean	SD	F	Sig. (2-tailed)
Theatre artists	Phonation	3259.16	334.87	6.169	0.007
	Speaking	3836.50	490.59		
	Reading	3682.08	503.15		
AIR announcers	Phonation	3352.33	593.52	5.670	0.01
	Speaking	3959.66	543.76		
	Reading	3624.41	497.35		

Table 5: Mean, Standard deviation and 'F' values of Speaker's Formant in three task conditions for both professional voice user groups.

Significant difference was found between the task conditions in AIR announcers ($F(2,22) = 5.670$, $p < 0.01$) as well as in Theatre artists groups ($F(2,22) = 6.169$, $p < 0.01$). From Bonferroni's results, in Theatre artists, the mean Speaker's Formant value for phonation was significantly different from both reading and speaking at 0.05 levels. In AIR announcers, the mean Speaker's Formant value for phonation was significantly different from speaking at 0.01 level.

Table 6 shows the mean and standard deviation of Speaker's Formant in three task conditions for Theatre artists. The mean value of Speaker's Formant was higher in females in all task conditions in Theatre artists. However significant difference was found only for speaking task condition at 0.05 level.

Conditions	Groups	Mean	Standard Deviation	t	Sig (2- tailed)
Phonation	Male	3156.166	237.339	1.073	0.309
	Female	3362.166	406.102		
Speaking	Male	3552.500	346.157	2.401	0.037
	Female	4120.500	464.865		
Reading	Male	3565.166	654.711	0.791	0.447
	Female	3799.000	309.062		

Table 6: Mean, Standard deviation and 't' values of Speaker's Formant in Theatre artists.

Table 7 shows the mean and standard deviation of Speaker's Formant in three task conditions for AIR announcers. It is evident that females had highest values for all task condition in AIR announcers. This result was similar to Theatre artists. But when females and males were compared no significant difference was noticed although the mean values were higher for females.

Conditions	Groups	Mean	SD	t	Sig (2- tailed)
Phonation	Male	3218.666	490.502	0.765	0.462
	Female	3486.000	701.089		
Speaking	Male	3735.666	621.414	1.507	0.163
	Female	4183.666	379.361		
Reading	Male	3218.666	328.140	0.489	0.635
	Female	3697.166	651.015		

Table 7: Mean, Standard deviation and 't' values of Speaker's Formant in AIR announcers.

II. Perceptual analysis

Perceptual analysis was carried out for all the subjects for speaking and reading task only. Results are summarized and the alpha values for each parameter in speaking and reading are shown in Table 8.

Parameters		Speaking	Reading
Voice	Quality	0.73	0.77
	Pleasantness	0.77	0.71
Articulation	Intelligibility	0.86	0.76
	Pronunciation	0.75	0.86
Fluency	Continuity	0.72	0.88
	Rate	0.82	0.88
	Effort	0.70	0.86
Prosody	Intonation	0.82	0.80
	Stress	0.61	0.75
	Rhythm	0.73	0.80

Table 8: Inter-judge reliability on various parameters for speaking and reading.

Reliability coefficient alpha was calculated to find inter-judge reliability between five judges. Since alpha values are greater than 0.70 for all parameters [except rate (0.61) in speaking task], the judgments could be considered reliable. For each of the parameters, the majority rating by judges were considered and correlated with acoustic measures. When subjects were considered as a group, irrespective of the gender, inter-judge reliability was highest for the parameter intelligibility (0.86) compared to other parameters, in speaking task. It was followed by intonation and rate (0.82). In reading task, inter-judge reliability coefficient was highest for continuity and rate, i.e. the alpha value was 0.88 for both parameters when compared

to all the other parameters. It was followed by pronunciation and effort, i.e. the alpha value was 0.86 in both the parameters.

Table 9 shows percentage ratings across gender in Theatre artists and AIR announcers across parameters for speaking and reading tasks. In general, it was observed that most of the perceptual parameters were rated as 'excellent' for both Theatre artists and AIR announcers. Within professional voice users groups, percentage rating was highest for AIR announcers when compared to Theatre artists. AIR announcers were rated 'excellent' in all the parameters except for rate in speaking task. In AIR announcers, only rate was rated as 'normal' in speaking task condition. In theatre artist, rate, effort and rhythm were rated as 'normal' in speaking task condition. None of the parameters were rated as below normal, except quality (3.33%) in speaking task in Theatre artists.

No	Parameters	Group	Speaking (in %)			Reading (in %)		
			Below normal	Normal	Excellent	Below normal	Normal	Excellent
1	Quality	TA	3.33	33.33	63.33	0	18.33	81.66
		AIR	0	16.66	83.33	0	11.66	88.33
2	Pleasantness	TA	0	23.33	76.66	0	26.66	73.33
		AIR	0	23.33	76.66	0	28.33	71.66
3	Intelligibility	TA	0	39.99	59.99	0	33.33	66.66
		AIR	0	26.66	73.33	0	11.66	88.33
4	Pronunciation	TA	0	31.66	68.33	0	13.33	86.66
		AIR	0	21.66	78.33	0	19.99	79.99
5	Continuity	TA	0	40.00	60.00	0	39.99	59.99
		AIR	0	16.66	83.33	0	16.66	83.33
6	Rate	TA	0	59.99	39.99	0	66.66	33.33
		AIR	0	54.99	44.99	0	34.99	64.99
7	Effort	TA	0	73.33	26.66	0	78.33	21.66
		AIR	0	39.99	59.99	0	66.66	33.33
8	Intonation	TA	0	44.99	54.99	0	59.99	39.99
		AIR	0	23.33	76.66	0	25.00	75.00
9	Stress	TA	0	50.00	50.00	0	23.33	76.66
		AIR	0	23.33	76.66	0	11.66	88.33
10	Rhythm	TA	0	76.66	23.33	0	38.33	61.66
		AIR	0	41.66	58.33	0	48.33	51.66

Table 9: Gender-wise ratings (in %) for Theatre artists and AIR announcers across speaking and reading tasks.

In reading task, it was observed that most of the parameters were rated as 'excellent' in both groups of professional voice users. Within reading task, high percentage rating was found for most of the parameters in AIR announcers when compared to Theatre artists. Rate, effort and intonation were rated as 'normal' in theatre artist whereas in AIR announcers, rate and effort were rated as 'normal' when compared to all the other parameters. None of the parameters were rated as 'below normal' in reading task condition.

Table 10 shows gender-wise ratings (in %) across Theatre artists and AIR announcers for all the parameters in speaking task.

Parameters	Gender	Theatre artists (in %)			AIR Announcers (in %)		
		Below normal	Normal	Excellent	Below normal	Normal	Excellent
Quality	M	0	16.66	83.33	0	13.33	86.66
	F	6.66	50.00	43.33	0	20.00	80.00
Pleasantness	M	0	6.66	93.33	0	23.33	76.66
	F	0	40.00	60.00	0	13.33	86.66
Intelligibility	M	0	33.33	66.66	0	26.66	73.33
	F	0	46.66	53.33	0	26.66	73.33
Pronunciation	M	0	30.00	70.00	0	16.66	83.33
	F	0	33.33	66.66	0	26.66	73.33
Continuity	M	0	40.00	60.00	0	13.33	86.66
	F	0	40.00	60.00	0	20.00	80.00
Rate	M	0	63.33	36.66	0	56.66	43.33
	F	0	56.66	43.33	0	53.33	46.66
Effort	M	0	86.66	13.33	0	33.33	66.66
	F	0	60.00	40.00	0	46.66	53.33
Intonation	M	0	36.66	63.33	0	20.00	80.00
	F	0	53.33	46.66	0	26.66	73.33
Stress	M	0	40.00	60.00	0	13.33	86.66
	F	0	60.00	40.00	0	33.33	66.66
Rhythm	M	0	63.33	36.66	0	23.33	76.66
	F	0	90.00	10.00	0	60.00	40.00

Table 10: Gender-wise ratings (in %) across Theatre artists and AIR announcers in speaking task.

It was noticed that males in both groups of subjects had higher ratings for most of the parameters when compared to females in speaking task. In general, it was noticed that most of parameters were rated as excellent by most of the judges. In theatre artist, rate, effort and rhythm were rated as 'normal' when compared to other parameters. None of the parameters were rated as 'below normal' except quality

(6.66%) in female Theatre artists. Within AIR announcers, rate was rated as 'normal' in both males and females and rhythm was rated as 'normal' in females.

Parameters	Gender	Theatre artists (in %)			AIR Announcers (in %)		
		Below normal	Normal	Excellent	Below normal	Normal	Excellent
Quality	M	0	26.66	73.33	0	13.33	86.66
	F	0	10.00	90.00	0	10.00	90.00
Pleasantness	M	0	23.33	76.66	0	26.66	73.33
	F	0	30.00	70.00	0	30.00	70.00
Intelligibility	M	0	33.33	66.66	0	6.66	93.33
	F	0	33.33	66.66	0	16.66	83.33
Pronunciation	M	0	6.66	93.33	0	23.33	76.66
	F	0	20.00	80.00	0	16.66	83.33
Continuity	M	0	43.33	56.66	0	16.66	83.33
	F	0	36.66	63.33	0	16.66	83.33
Rate	M	0	66.66	33.33	0	33.33	66.66
	F	0	66.66	33.33	0	36.66	63.33
Effort	M	0	73.33	26.66	0	53.33	46.66
	F	0	83.33	16.66	0	80.00	20.00
Intonation	M	0	73.33	26.66	0	20.00	80.00
	F	0	46.66	53.33	0	30.00	70.00
Stress	M	0	23.33	76.66	0	10.00	90.00
	F	0	23.33	76.66	0	13.33	86.66
Rhythm	M	0	30.00	70.00	0	63.33	36.66
	F	0	46.66	53.33	0	33.33	66.66

Table 11: Gender-wise ratings (in %) across Theatre artists and AIR announcers in reading task.

Table 11 shows the percent ratings for Theatre artists and AIR announcers across parameters in reading task. In general, majority of the judges rated all subjects as 'excellent' for most of the parameters in reading task irrespective of the professional voice user groups. Also none of the judges were rated as 'below normal'

on all the parameters. Within professional voice users, males were rated as 'excellent' in most of the parameters when compared to females. In theatre artist, rate and effort were rated as 'normal' in reading task irrespective of the gender. In AIR announcers, effort was rated as 'normal' in both males and females, but rhythm was rated as 'normal' in males

III. Correlation between perceptual and acoustic analyses

Correlation was studied across the subjects as follows,

Group I: Presence of Speaker's Formants in all task condition.

Group II: Presence of Speaker's Formant in any one task condition

Spearman's rank correlation between ratings of perceptual parameters and presence Speaker's Formant in speaking and reading for Group I was done and is shown in Table 12.

Results showed that perceptual ratings of most of the parameters were significantly correlating with Speaker's Formant for reading and speaking. Results revealed that parameters with correlation between perceptual ratings and Speaker's Formant were different in both task conditions. In speaking task condition, irrespective of the professional voice user groups (Theatre artist and AIR announcers) the results indicated that the perceptual ratings for pleasantness ($r = 0.783$, $p < 0.001$) highly correlated with presence of Speaker's Formant. On the other hand rhythm ($r = 0.007$, $p < 0.05$) correlated least significantly with presence of Speaker's Formant in speaking task. Intelligibility ($r = 0.730$), pronunciation ($r = 0.705$) and quality ($r = 0.689$) were significantly correlating at $p < 0.001$, and intonation ($r = 0.646$) and stress ($r = 0576$) were significantly correlating at $p < 0.01$. For the parameter continuity ($r =$

.468) correlation was significant at $p < 0.05$. There was no correlation found between rate, effort and presence of Speaker's Formant in all the task conditions.

Parameters		Speaking	Reading
Voice	Quality	0.689***	0.608**
	Pleasantness	0.783***	0.536**
Articulation	Intelligibility	0.730***	0.272
	Pronunciation	0.705***	0.716***
Fluency	Continuity	0.468*	0.653**
	Rate	0.384	0.218
	Effort	0.287	0.179
Prosody	Intonation	0.646**	0.508*
	Stress	0.576**	0.537**
	Rhythm	0.007	0.307

Table 12: Spearman's rank correlation between perceptual parameters and Speaker's Formant in speaking and reading in Group I. *** = correlation coefficient is significant at 0.001 level (2- tailed), ** = correlation is significant at the 0.01 level (2- tailed). * = correlation coefficient is significant at 0.05 level (2- tailed).

In reading task, pronunciation ($r = 0.716$) correlated most significantly with presence of Speaker's Formant at $p < 0.001$. Continuity ($r = 0.653$), quality ($r = 0.608$), stress ($r = 0.537$) and pleasantness ($r = 0.536$) correlated significantly with the presence of Speaker's Formant at $p < 0.01$. Intonation (0.508) correlated with presence of Speaker's Formant only at $p < 0.05$.

Spearman's rank correlation between perceptual ratings and presence of Speaker's Formant across Theatre artists and AIR announcers. (Group I)

The correlation between perceptual ratings and presence of Speaker's Formant across both task conditions i.e. speaking and reading in Theatre artists and AIR announcers is shown in Table 13.

Parameters		Speaking		Reading	
		TA	AIR	TA	AIR
Voice	Quality	0.753**	0.649*	0.641*	0.583*
	Pleasantness	0.808***	0.744***	0.474	0.585*
Articulation	Intelligibility	0.819***	0.552*	0.154	0.418
	Pronunciation	0.768**	0.685*	0.648*	0.857***
Fluency	Continuity	0.362	0.481	0.710**	0.753**
	Rate	0.122	0.552	0.389	0.220
	Effort	0.367	0.196	0.324	0.171
Prosody	Intonation	0.573**	0.481	0.530	0.666*
	Stress	0.416	0.649*	0.512	0.585*
	Rhythm	-0.154	0.140	0.024	0.563

Table 13: Spearman's rank correlation between perceptual parameters and Speaker's Formant in speaking and reading across professional voice groups (TA = Theatre artists, AIR = AIR announcers). ***= correlation coefficient is significant at the 0.001 level (2- tailed), ** = correlation is significant at the 0.01 level (2- tailed). * = correlation is significant at the 0.05 level (2- tailed).

In Theatre artists intelligibility (r = 0.819) and pleasantness (r = 0.808) correlated highly with presence of Speaker's Formant at $p < 0.01$ level. Pronunciation (r = 0.768), quality (r = 0.753) and intonation (r = 0.573) were correlating with presence of Speaker's Formant in reading task conditions. In speaking task, when perceptual ratings for AIR announcers was compared for

correlation with presence of Speaker's Formant, the results indicated that pleasantness ($r = 0.744$) was highly correlating at $p < 0.001$ level with Speaker's Formant. Pronunciation ($r = 0.685$), quality ($r = 0.649$), stress ($r = 0.649$) and intelligibility ($r = 0.552$) were correlating with the presence of Speaker's Formant at $p < 0.05$. No correlation was obtained for perceptual ratings for continuity, rate, effort and rhythm.

In reading task, within Group I, when theatre artists were considered, results revealed that perceptual ratings for continuity ($r = 0.710$) was highly correlating with the Speaker's Formant at $p < 0.01$ with the Speaker's Formant. But perceptual ratings of pronunciation ($r = 0.648$) and quality ($r = 0.641$) correlated at $p < 0.05$ with the Speaker's Formant. In AIR announcers for reading task, results show that perceptual ratings for pronunciation ($r = 0.857$) was significantly correlating at $p < 0.001$ with presence of Speaker's Formant. Intonation ($r = 0.666$), stress ($r = 0.585$), pleasantness ($r = 0.585$) and quality ($r = 0.583$) were highly significant at $p < 0.05$. Continuity ($r = 0.563$) did not correlating at $p< 0.05$ level but it was correlating at $p < 0.1$ with Speakers Formant in reading. Overall, there was no significant correlation between rate, effort and rhythm and Speaker's Formant in speaking and reading across theatre artist and AIR announcers groups.

Table 14 shows Spearman's rank correlation between perceptual ratings of parameters and presence of Speaker's Formant in Group II. Five subjects who had presence of Speaker's Formant in speaking and 9 subjects with presence of Speaker's Formant in reading were considered for the respective task conditions.

Parameters		Speaking (5)	Reading (9)
Voice	Quality	0.707	0.000
	Pleasantness	0.707	0.000
Articulation	Intelligibility	0.866	0.000
	Pronunciation	0.707	0.000
Fluency	Continuity	0.707	-0.411
	Rate	0.577	0.725*
	Effort	0.707	0.725*
Prosody	Intonation	0.866	0.725*
	Stress	0.866	0.866**
	Rhythm	0.289	0.822**

Table 14: Spearman's rank correlation between perceptual ratings of parameters and presence of Speaker's Formant in speaking and reading for Group II. ** = correlation is significant at the 0.01 level (2- tailed). * = correlation is significant at the 0.05 level (2- tailed). Highlighted values correlated at significance level of 0.1.

In speaking task condition, perceptual ratings for intelligibility (0.866), intonation (0.866) and stress (0.866) were not correlating with Speaker's Formant at $p < 0.05$ but it correlated at $p < 0.1$.

In reading task, results indicated that perceptual ratings for rhythm and stress were significantly correlating with presence of Speaker's Formant at $p < 0.01$. Perceptual ratings for effort, rate and intonation were significantly correlating with presence of Speaker's Formant at $p < 0.05$, in reading task. There was no correlation between perceptual ratings for quality, pleasantness, intelligibility and pronunciation with the presence of Speaker's Formant in reading task. Differences between theatre artists and AIR announcers could not be done as the number of subjects was less in the Group II.

IV. Perceptual correlates of 'good speech' and presence of Speaker's Formant in both professional voice users. (Group I)

Table 15 shows the results of the comparisons between the obtained perceptual ratings 'normal' and 'excellent' with the values of the Speaker's Formant for subjects in Group I for speaking.

Table 15 shows the mean and standard deviation of Speaker's Formant with respect to the perceptual ratings of 'normal' and 'excellent' in speaking task conditions. 'Mann Whitney test' revealed that perceptual ratings of quality, pleasantness, intelligibility and pronunciation significantly correlated with the existence of Speaker's Formant in speaking at $p < 0.001$. Mean value was higher for 'excellent' ratings in all the parameters when compared to normal ratings. When the speaking task was rated as "excellent" the value of the Speaker's Formant was higher than the values obtained when rated as 'normal'.

Parameters	Ratings	No. of subjects	Mean	SD	Z
Quality	Normal	5	3210.000	322.562	3.306***
	Excellent	19	4079.157	378.926	
Pleasantness	Normal	9	3420.222	356.399	3.757***
	Excellent	15	4184.800	347.086	
Intelligibility	Normal	6	3276.833	331.909	3.501***
	Excellent	18	4105.166	372.002	
Pronunciation	Normal	6	3304.166	355.250	3.234***
	Excellent	18	4096.055	386.391	
Continuity	Normal	4	3269.250	501.183	2.247*
	Excellent	20	4023.850	419.232	
Rate	Normal	7	3579.714	416.063	1.842
	Excellent	17	4029.176	496.907	
Effort	Normal	10	3713.500	644.008	1.376
	Excellent	14	4029.928	358.864	
Intonation	Normal	4	3148.750	337.221	3.099**
	Excellent	20	4047.950	394.343	
Stress	Normal	7	3469.000	391.501	2.763**
	Excellent	17	4074.764	450.824	
Rhythm	Normal	17	3893.411	531.334	0.032
	Excellent	7	3909.428	495.422	

Table 15: Mean and Standard deviation for speaking. ***= correlation coefficient is significant at the 0.001 level (2- tailed), ** = correlation is significant at the 0.01 level (2-tailed). * = correlation is significant at the 0.05 level (2- tailed).

Table 16 shows the results of the comparisons between the obtained perceptual ratings 'normal' and 'excellent' with the values of the Speaker's Formant for subjects in Group I for reading.

Parameters	Ratings	Number of subjects	Mean	Standard deviation	Z
Quality	Normal	5	3148.400	169.35	
	Excellent	19	3786.105	459.55	2.917**
Pleasantness	Normal	6	3242.50	200.910	
	Excellent	18	3790.16	483.50	2.569**
Intelligibility	Normal	7	3456.285	447.434	
	Excellent	17	3734.352	496.314	1.303
Pronunciation	Normal	7	3163.571	220.182	
	Excellent	17	3854.882	423.179	3.433***
Continuity	Normal	9	3261.666	297.972	
	Excellent	15	3888.200	431.563	3.133**
Rate	Normal	15	3588.333	493.697	
	Excellent	9	3761.444	493.068	1.044
Effort	Normal	17	3631.411	539.173	
	Excellent	7	3706.285	375.653	0.858
Intonation	Normal	11	3435.846	417.193	
	Excellent	13	3910.181	457.889	2.435*
Stress	Normal	7	3249.714	343.583	
	Excellent	17	3819.411	447.790	2.574**
Rhythm	Normal	12	3484.750	402.613	
	Excellent	12	3821.750	527.543	1.474

Table 16: Mean, standard deviation for reading. ***= correlation coefficient is significant at the 0.001 level (2- tailed), ** = correlation is significant at the 0.01 level (2- tailed). * = correlation is significant at the 0.05 level (2- tailed).

Table 16 shows mean and standard deviation of the parameters rated as 'normal' and 'excellent' in professional voice users. Perceptual ratings of parameters 'normal' and 'excellent' were compared with the mean values of the Speaker's Formant of both professional voice users. 'Mann Whitney test' indicated that some of

the parameters showed significant difference between perceptual ratings of 'normal' and 'excellent' with Speaker's Formant of reading. Pronunciation parameter correlated highly and was significant with Speaker's Formant in reading. Perceptual ratings for quality, pleasantness, continuity and stress showed significant difference at $p < 0.01$ with the Speaker's Formant. There was no significant difference between perceptual ratings of effort, rate and rhythm with the Speaker's Formant in reading. Higher mean value was obtained for 'excellent' ratings when compared to 'normal' ratings.

CHAPTER V

DISCUSSION

The main aim was to study the presence of Speaker's Formant in Theatre artists and AIR announcers and the relation between the perceptual parameters and acoustic findings. The results of the present study revealed several points of interests.

1. Presence of Speaker's Formant in both groups of professional voice users (Theatre artists and ***AIR*** *announcers).*

LTAS analysis revealed that a total of 64.86 % of the subjects had Speaker's Formant. A clear peak in the frequency region between 2800 - 4200 Hz was noticed in both the professional voice user groups. These results are in consensus with the findings of Leino (1993), who has reported grouping of third, fourth and fifth formants at around 3.5 kHz in the projected voices and used the term Speaker's Formant. Studies did on the voices of actors in various Western languages have revealed the presence of Speaker's Formant (Leino, 1993; Munro, 2002; Bele, 2002; Pinzower & Oates, 2005).

Further, it is reported that Speaker's Formant is evident in the voices of such professionals as actors or in the voices of other trained speakers with superior voice quality (Nawka et al., 1997). Bele (2002) compared actors and teachers and reported that Speaker's Formant was stronger for actors when compared to teachers.

Presence of Speaker's Formant is explained in terms of the superior resonatory mechanize in elite professional voice user groups, including Theatre artists and AIR

announcers. The region of the Speaker's Formant can be regarded as a distinguishing resonatory effect of the vocal tract with only minimal change in frequency, which is characteristic of the fourth formant. It may be resulting from the glottal source and the additional resonatory properties of the vocal tract that these professional voice user groups adapt to project their voice (Nawka et al., 1997).

In the present study, it was found that even females had Speaker's Formant but frequency value was higher compared to males. Acker (1987) reported presence of Speaker's Formant in phonation in both genders. This confirms the view of Acker (1987) that the resonatory characteristics in female professional speakers could be similar to that of male professional speakers.

Speaker's Formant was studied in three different task conditions i.e. phonation reading and speaking in this study. Speaker's Formant was found in 64.86 % of total subjects in all the task conditions. The following studies have supported these findings. Master et. al., (2006) found that Speaker's Formant in male Brazilian actors' text reading. Speaker's Formant was found in reading (Bele, 2002, 2005), connected speech (Nawka et al., 1997) and in phonation (Acker, 1987).

It is opined that a professional speaker may employ a different speaking strategy, and even a different technique for production of public speech, when compared to a normal conversation speech. Manipulation of the spectral balance may be the only strategy available for the speaker in order to obtain a 'carrying voice'.

2. Perceptual ratings of speaking and reading

Majority of subjects in both the professional voice users were rated as 'excellent' in most of the parameters. This indicates that the parameters of voice, articulation, fluency and prosody are important contributors for professional voice users. More specifically, quality, pleasantness, pronunciation, continuity, and stress were seen to be the characteristic features of the voices in these two groups of professional voice users. Hence, it can be speculated that these parameters may be effective for vocal projection.

When Theatre artists were considered separately, a total of six parameters in the speaking and reading were rated as 'excellent'. They were quality, pleasantness, intelligibility, pronunciation, continuity and intonation. Eight parameters in the speaking and reading tasks were rated as 'excellent' in AIR announcers. This indicates that AIR announcers were better in terms of their voice when compared to the Theatre artists. Therefore, it can be opined that the semiformal training AIR announcers undergo for voice helps them in maintaining a well-controlled speech, which has good timing, excellent pronunciation and correct grammar. They also attract a large group of audience by combining a pleasing personality and voice with an appealing style. But in contrast, though the voice of Theatre artists are equally good, they use their voice concentrating mainly on loudness and clear pronunciations than quality of voice. Also, they are required to use voice constantly with lot of variations depending on the roles they portray. Though the professional demands of each profession are different, general vocal hygiene and conservation measures would be of help in both groups for maintenance of'professional voice'.

3. Correlation between the ratings of perceptual parameters and Speaker's Formant

Most of the perceptual parameters correlated significantly with the presence of Speaker's Formant in speaking task in Theatre artists. For AIR announcers most of the perceptual parameters correlated significantly in reading task. This indicates that Theatre artists were good in speaking whereas, AIR announcers in reading. This finding supports the differential professional demands on voice. Theatre artists need to project their voice while speaking on stage, therefore they have better voice modulation in speaking. AIR announcers usually have script backed speech tasks to perform in their profession, such as announcing, interviewing, anchoring, etc. Hence, they perform better in reading tasks compared to spontaneous speech by emphasizing more on quality, pronunciation, intelligibility and intonation of speech.

Leino (1993) opined that presence of Speaker's Formant indicated "good speaking voice" quality in professional Finnish actors. Leino (1993) and Nawka et. al., (1997) preferred the term "good speaking voice" instead of "projection". They opined "projection" was related to "good professional speaking voice".

Results indicated that some subjects in both groups of professional voice users (Theater artists and AIR announcers) had Speaker's Formant in either one of the task conditions. In these subjects only a few perceptual parameters correlated with SPF. This indicates that SPF was not present in all tasks and hence, they could have been rated as "normal" on most parameters.

4. Perceptual correlates of 'good speech' and presence of Speaker's Formant in both groups of professional voice users.

When both 'normal' and 'excellent' rating in perceptual analysis were considered for correlation, it was revealed that there was significant correlation between the ratings and the frequency value of the Speaker's Formant. i.e. their value became higher as their ratings increased. When subjects were rated as 'excellent', the SPF value was higher. This shows that clustering of formants increased the amplitude of the peak/s in the Speaker's Formant region emphasizing a glottal source as the origin of Speaker's Formant.

The finding of the present study implies that Speaker's Formant might be an indicator of good speech in both theatre artists and AIR announcers. A high correlation between presence of Speaker's Formant and perceptual parameters of voice, articulation, fluency, and prosody emphasize the importance of these parameters in good speaking voice. Hence all these perceptual parameters may be considered to evaluate "good speaking voice" and further acoustically confirm presence of Speaker's Formant.

CHAPTER VI

SUMMARY AND CONCLUSION

A total of thirty-seven female and male subjects from two groups of professional voice users (Theatre artist and AIR announcers) participated in the study. The present study aimed at discerning the presence of speaker's formant from acoustic analysis; ascertain parameters of speech from perceptual analysis and correlate Speaker's Formant with perceptual parameters in all subjects. The subjects performed three task conditions i.e. phonation of /a/, reading a standard Kannada passage and speaking (monologue about 'self). Long-term Average Spectrum was extracted for the samples in all task conditions. The Speaker's formant (a peak in the frequency region between 2.8 to 4.2 kHz) was noted for each sample separately when present.

The Speaker's Formant is reported to be a strong peak about at 3.5 kHz found in better speaking voices (Acker, 1987; Leino, 1993; Nawka, et. al., 1997; Bele, 2002 & 2005) implying that SPF is more evident in the trained voices of professionals, such as actors, voice-over artists etc. In the present study, Speaker's Formant was evident in most subjects (both Theatre artists and AIR announcers). But the range of the Speaker's Formant value wider (2800 - 4200 Hz). SPF was noted in males and females in both professional voice user groups.

Perceptual analysis involved rating of speaking and reading samples of all subjects individually by five experienced Speech Language Pathologists. Results revealed that most of the judges rated the samples (speaking and reading) as

'excellent' on most parameters. Of the parameters, quality, pleasantness, intelligibility, pronunciation, continuity, intonation and stress were rated as "excellent" in speaking and reading task conditions for both groups of subjects.

The correlation between acoustic and perceptual analysis revealed high correlation between presence of SPF and quality, pleasantness, intelligibility, pronunciation, continuity, intonation and stress. Theatre artists were rated as "excellent" in more parameters in speaking task and on the other hand AIR announcers were rated as "excellent" in reading task conditions. None of the parameters were rated as 'below normal'.

Further, correlation between perceptual determinants 'good speech' and Speaker's Formant was done. It was found that the subjects, in whom SPF was present, were rated as "excellent" in most of the perceptual parameters. Hence, these parameters could be speculated as perceptual correlates of "good speech".

Future direction: In general, the results of the present study revealed that presence of SPF is an indicator of "good speaking voice". Also, this could be confirmed by the perceptual correlates. However, speakers have to be trained in large numbers on these perceptual parameters, and then using acoustic analysis checked for presence of SPF. The results of such a study would confirm the findings of the present study.

REFERENCES

Acker, B. F. (1987). Vocal tract adjustments for the projected voice. *Journal of Voice,* 1 (1), 77-82.

Anderson, V.A. (1942). *Training the speaking voice.* 2nd edition. New York: Oxford University Press.

Barrichelo, V. O., Hener, J. R., Dean, C. M, & Sataloff, R. T. (2001). Comparison of singer's formant, speaking ring, and LTAS among classical singers and untrained speakers. *Journal of Voice,* 15 (15), 344-350.

Bartholomew, W. T. (1934). A physical definition of "good voice" quality in the male voice. *Journal of Acoustical Society of America,* 6 (2), 25-33.

Bele, I. V. (2002). *Professional speaking voice: A perceptual and acoustic study of actor's and teacher's voices.* Master's dissertation submitted in part fulfillment of M. Sc. (Speech Language Pathology) to the University of Oslo: Norway.

Bele, I. V. (2005). Reliability in perceptual analysis of voice quality. *Journal of Voice,* 19 (4), 555-573.

Bele, I.V. (2006). Dimensionality in voice quality. *Journal of Voice,* In press.

Bele, I.V. (2006). The speaker's formant. *Journal of Voice,* 20 (1), 555-578.

Brodnitz, F. (1988). *Keeping your voice Healthy* (2nd ed.). Texas: PRO-ED, Inc.

Camargo, Z. A. (2002). Cited in Master, S., Biase, N. D., Pedrosa, V., & Chiari, B. M. (2006). The long-term average spectrum in research and in the clinical practice of speech therapists. *Journal of Voice,* 18(1), 111-120.

Deepa, D. (2004). *Voice and speech characteristics of Radio Jockeys,* Unpublished Master's dissertation submitted in part fulfillment of M. Sc. (Speech Language Pathology) to the University of Mysore.

Ekholm, E., Papagiannis, G. C, Cagnon, F. P. (1998). Relating objective measurement to expert evaluation of voice quality in Western classical singing: Critical perceptual parameters. *Journal of Voice,* 12 (2), 182-196.

Fisher, H.B. (1966). *Improving voice and articulation.* Boston: Houghton Miffliss Company.

Frokjaer-Jensen, B., & Prytz, 1976; Kitzing, 1986, & Lofqvist, (1986). Cited in Master, S., Biase, N. D., Pedrosa, V., & Chiari, B. M. (2007). The long-term average spectrum in research and in the clinical practice of speech therapists. Pro-Fono Revista De Atualixacao Cientifica, Barueri (SP), 18 (1), 111-120.

Haiti, D. M., Hans, S., Vaissiere, J., Riquet, M, & Brasnu, D. F. (2001). Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *Journal of Voice,* 15(2), 351-361.

Kovacic, G., & Budanovac, A. (2002). Acoustic characteristic of adolescent and actors' and non-actors' voices. *Folia Phoniatrica Logopadica,* 54, 125-132.

Laukkanen, A. M, Syrja, J., Laitala , M., & Leino, T. (2004). Effects of two-month vocal exercising with and without spectral biofeedback on student actor's voice. *Logopedics Phoniatrica Vocology,* 29 (2), 66-74.

Leino, T. (1993). Long-term average spectrum study on speaking voice quality in male voices. *SMAC93* Proceedings of the Stockholm Music Acoustics Conference, July 28-August 1, Stockholm, Sweden.

Lessac, A. (1997). *The use and training of the human voice: A practical approach and voice dynamics.* New York: Mayfield Publishing Company.

Linville, S., & Rens, J. (2001). Vocal tract resonance analysis of aging voice using the long term average spectra *Journal of Voice*, 15 (3), 323- 330.

Master, S., Biase, N. D., Pedrosa, V., & Chiari, B. M. (2006). The long-term average spectrum in research and in the clinical practice of speech therapists. *Pro-Fono Revista De Atualixacao Cientifica,Barueri (SP),* 18 (1), 111-120.

Meano, C. (1967). *The human voice in speech and song.* Charles C Thomas. Publisher.

Medrado, R., Ferreira, L. P., & Behlau, M. (2005). Voice Over: Perceptual and acoustic analysis of vocal features. *Journal of Voice,* 19 (3), 340-349.

Mendoza, E., Valencia, N., Munoz, J., & Trujillo, H. (1996). Differences in voice quality between men and women: Use of the long-term average spectrum. *Journal of Voice,* 10(1), 59-66.

Morris, J., Weiss, R. (1997). The singer's formant revisited: pedagogical implications based on a new study. *Journal of Singing*, 53 (3), 21-25.

Murry, T., & Rosen, C. A. (2000). Voice handicap index in singers. Journal of *Voice,* 14 (3), 370-377.

Nawka, T., Anders, L., Cebulla, M., & Zurakwski, D. (1997). The speaker's formant in male voices. *Journal of Voice,* 11 (4), 422-428.

Nolan, F. (1983). *The Phonetic Bases of Speaker Recognition.* Cambridge, MA: Cambridge University Press.

Pinczower, R.., & Oates, J. (2005). Voice projection in actors: The LTAS features that distinguish comfortable acting voice from voicing with maximal projection in male's voice. *Journal of Voice,* 19 (3), 440-453.

Raphael, B., & Scherer, R. (1987). Voice modifications of stage actors: acoustic analyses. *Journal of Voice,* 1 (1), 83-87.

Rothman, H. B., Brown, W. S., Jr., & LaFond, R. (2002). Spectral changes due to performance environment in singers, nonsingers, and actors. *Journal of Voice,* 16 (1), 323-332.

Sjolander, P. (2004). Spectrum effects of subglottal pressure variation in professional baritone singers. *Journal of Acoustic Society of America,* 115 (3), 1270-1273.

Snidecor, J.C. (1951). The pitch and duration characteristic of superior speakers during oral reading. *Journal of Speech Disorder,* 16(1), 44-52.

Stemple, J. C. (1993). *Voice therapy: Clinical studies.* St. Louis, MO: Mosby - Year Book, Inc.

Sundberg, J. (1990). What's so special about singers? *Journal of Voice,* 4 (2), 111-113.

Vennard, W. (1967). *Singing: The Mechanism and the Technique.* New York: NY: Carl Fisher.

Weiss, R., Brown, W. S., Jr., & Morris, J. (2001). Singer's Formant in sopranos: Fact or Fiction? *Journal of Voice,* 15 (4), 457-468.

Wilson, k. (1972). *Voice problem of children. Baltimore:* The Williams & Wilkins Company.

Winckel, F. (1953). Physikalische Kriterrien fur objective Stimmbeurteilung. *Folia Phoniatrica (Basel), 5* (3), 232-252.

APPENDIX 1: Score sheet for perceptual analysis.

Perceptual analysis

Name:

Age/ gender:

Number of years of clinical experience:

	Parameters	**l=Below Normal**	**2= Normal**	**3= Excellent**
1	**Voice**			
	Quality			
	Pleasantness			
2	**Articulation**			
	Intelligibility			
	Pronunciation			
3	**Fluency**			
	Continuity			
	Rate			
	Effort			
4	**Prosody**			
	Intonation			
	Stress			
	Rhythm			