ACOUSTIC CHARACTERISTIC OF NEPALI LANGUAGE (LONG TERM AVERAGE SPEECH SPECTRUM)

REG.NO. M9304

AN INDEPENDENT PROJECT WORK SUBMITTED IN PART FULFILMENT FOR FIRST YEAR M.Sc. (SPEECH AND HEARING) TO THE UNIVERSITY OF MYSORE.

ALL INDIA INSTITUTE OF SPEECH AND HEARING, MYSORE 570 006.

DEDICTED TO

MY DEAREST

BUBA (FATHER)

MAA

SISTER

CERTIFICATE

This is to certify that the Independent Project entitled: ACOUSTIC CHARACTERISTIC OF NEPALI LANGUAGE (LONG TERM AVERAGE SPEECH SPECTRUM) is the bonafide work done in part fulfilment for first year M.Sc, (Speech and Hearing) of the student with Reg. NO.M9304.

Mysore May 1994

All India Institute of Speech and Hearing, Mysore 570 006.

CERTIFICATE

This is to certify that this Independent Project entitled: ACOUSTIC CHARACTERISTIC OF NEPALI LANGUAGE (LONG TERM AVERAGE SPEECH SPECTRUM) has been prepared under my supervision and guidance.

Mysore May 1994 Dr. (Miss) S.Nikam,

GUIDE

DECLARATION

I hereby declare that this Independent Project entitled! ACOUSTIC CHARACTERISTIC OF NEPALI LANGUAGE (LONG TERM AVERAGE SPEECH SPECTRUM) is the result of my own study undertaken under the guidance of Dr.(Miss) S.Nikam, Director, All India Institute of Speech and Hearing, Mysore and has not been submitted earlier at any University for any other Diploma or Degree.

Mysore May 1994. Reg. No.M9304

ACKNOWLEDGEMENTS

I would like to extend my sincere thanks to Dr. (Miss) S. Nikam, HOD-Audiology, All India Institute of Speech and Hearing, Mysore, for her invaluable guidance.

I am grateful to the Director, Dr. (Miss) S.Nikam, AIISH for giving me an opportunity to take up this project.

I thank Dr. N.P.Nataraja, Prof, and HOD- Speech Science Department, AIISH for permitting me to use the computer for my project.

Savithri Madam, without your moral and emotional support this project would have been impossible to materialise.

Vanaja Mam and Manjula Mam thanx for everything.

Special thanks to Mr.C.S. Venkatesh, Ms. Rohini, Ms. Sridevi Ms. Lalita, Ms.Sathya for giving final touches to my project.

Rajalakshmi Akka for transforming this project into a presentable form.

Sheela you are the most adorable friend I have.

Hiru, Malu, Gayu, Bhavas, Mona and Gopi it's nice to have you guys around.

Suhasini, Elizabeth, your ever smiling faces always motivate me.

TABLE OF CONTENTS

	Page NQ
INTRODUCTION	1 - 9
REVIEW OF LITERATURE	10 - 25
METHODOLOGY	26-30
RESULTS & DISCUSSION	31 -41
SUMMARY & CONCLUSIONS	42 - 44
BIBLIOGRAPHY	45 - 49
APPENDIX	

INTRODUCTION

Speech is the method of communication by altering the frequency and intensity of voice sound and the intensity of unvoiced sound by the larynx and the lungs. The spectral content of the sound is modified by changing the dimension of the cavities of the mouth and nose which behaves as Helmholtz resonator.

Acoustic characteristic of speech:

Under the classification of acoustic signal speech comes under transient signal. A transient signal is one where the amplitude and frequency component change with respect to time. A transient signal can be analyzed into discrete frequency bands and corresponding intensities with respect to time. The type of analysis give a 3 dimensional picture of speech signal at different time interval.

Any speech spectrum can be obtained depending upon the need of application. The analysis of speech signal is essential in the area of speech transmission, speech perception, speech synthesis and speech identification. Also information is useful in the area of phonetics and psycholinguistics.

Speech is an integral part of the total personality revealing the speaker's environment, social contact and education. Other aspect, such as dress are external but speech is inherent (Mulgrave, Gelman).

What is LTAs

Speech spectrum is concerned with the frequency and intensity of each overtone (harmonic) of the speech wave. This overtone makes to speech intelligibility.

The energy level of each part of spectrum is measured and summed up separately for the speech sequence. The summed energy for each part of spectrum is then plotted. The resulting curve is the long term average speech spectrum.

So far, we have assumed that the vocal tract and therefore the speech spectrum remains in a certain fixed shape for a specific sound produced and changing rapidly to fixed shape appropriate for the next sound.

We were thinking of speech as a sequence of different stationary configuration but the speech wave has very few segments whose principal feature is continuously varying process (Brood and Peterson, 1957).

Long term average speech spectrum reflects several important characteristic of reasonably lengthy period of continuous discourse (Denis and Pinson, 1973).

Such characteristics have been well documented for speech produced by normal men and women.

Long term average spectrum provides information on the spectral distribution of the speech signal over a period of time. Such speech signal represents the product of source and the vocal tract transfer function.

LTAS has been used for studies of the human voice source. The speech signal represents the product of the sound source and the vocal tract transfer function which differ for different sound sequence but it is a varying process. The short term variation due to phonetic structure will be averaged out and the resulting spectrum can be used to obtain information on the sound source (Lofquist and Maderson, 1987).

How to analysis LTAS:

There are a number of methods by which speech can be analyzed spectrally. One such analysis procedure takes a time average of the sound pressure level/cycle across frequency. The measurement is commonly referred to as the long term average spectrum of speech (Formby and Monsen, 1985).

The measurement of LTAS is made by passing energy through a series of continuous band pass filter and interpreting the energy at the output of each filter, these average values are then plotted to arrive at the visual representative. The earliest measurement of LTAS was reported by Bell Telephone Laboratories about 7£ years ago (Crandall and Mackerzince, 1922).

Recently LTAS has been used to study the hearing aid selection (Stermochowicz et al. 1973).

Parameters of LTAS:

- 1. Alpha ratio (α -ratio)
- S. Beta ratio $< \beta$ -ratio)
- 3. Gamma ratio (γ -ratio)

- 4. Energy above 1 KHz
- 5. Energy below 1 Hz
- 6. The frequency of highest peak (F.H.P)
- 7. Energy in the frequency band between 5-8 KHz.
- 8. Graphic display and spectral pattern in frequency range of $0-8\ \mathrm{KHz}$.

CLINICAL APPLICATION

- 1. Acoustic voice sampling has practical advantages in clinical applications as it is non-invasive, does not require close co-operation from the patient and can be made off-line from tape recordings. Such application exclude the screening of large population for early detection of voice pathology, in particular the follow-up of the effects of voice therapy (Lofquist and Manderson, 1987).
- S. Long-term speech spectrum plotted in audiogram format can be quite useful to the clinical audiologist.
- 3. LTAS has been used to assess hearing aid gain, frequency selection and prescription (Cox, Moore, 1988).
- 4. LTAS approach for voice analysis appears to have potential application in the field of speaker recognition and

speaker identification atleast as part of comprehensive system (Massely and Hall ion, 1987).

5. LTAS is a reminder to the clinician of what will b. amplified. The speech spectra also serve as a baseline in prescriptive fitting for the amplified goal.

The speech spectrum is a handy visual aid in counselling the client to explain why they may be having difficulty in understanding certain speech sound categories on level of speech.

- 6. A plot of speech spectrum on an audiogram and sound field threshold obtained with the patient and given hearing aid arrangement can demonstrate the portion of speech spectrum made audible by the hearing aid. Adjustment then can be made to shape the amplified spectrum so that it is maintained within a comfortable listening region for the patient (Wayneel, David, 1987).
- 7. LTAS is a tool in voice research (Ander Lofquist, 1986).
- 8- LTAS is criteria pertinent to the measurement of voice quality (Peterkitzing, 1985).

Hers ife speakers (8 male and 8 female) who spoke 4 different language in the age of 20-40 years were studied. For each speaker 7 parameter of LTAS, namely, Alpha ratio, Beta ratio, Gamma ratio, Energy below 1 KHz, Energy above 1 KHz, Energy in the frequency band 5-8 KHz were measured. The frequency of highest peak were studied by using computer with necessary software.

This study was designed to answer the following questions:

- 1. Is there a significant difference between Nepali and Hindi languages (Indo Aryan) in long-term average speech spectrum?
- f. Is there a significant difference between Nepali language (Indo Aryan) and Kanhada (Dravidian) language in long term average speech spectrum?
- 3. Is there a significant difference between Nepali and Malayalam (Dravidian language) in long term average speech spectrum?

DEFINITION OF PARAMETERS STUDIED

1. Alpha ratio (41): It is defined as the energy above 1 KH2 divided by the energy below 1 KHz.

= Energy above 1 KHz

Energy below 1 KHz.

S. Beta ratio (β): It is defined as the energy between 0-2 KHz divided by the energy between S-8 KHz.

= Energy between 0-2 KHz

Energy between 2-8 KHz

- 3. 6amma ratio (δ): It is defined as the energy between 0-1 KHz divided by the energy between 5-8 KHz.
 - = Energy between 0-1 KHz

Energy between 5-8 KHz

- 4. Energy above 1 KHzs It is defined as a sum of energy in frequency range 1-8 KHz divided by number of point in this range.
 - = Sum of energy (1-8 KHz)

Number of point in this range

- 5. Energy below 1 KHz: Sum of energy in frequency range 0-1 KHz divided by number of point in this range.
 - = Sum of energy 0-8 KHz

Number of point in this range

- 6. Energy in the frequency band (5-8 KHz): Sum of energy in frequency range (0-1 KHz) divided by number of point in this band frequency.
 - = Sum of energy of frequency band 5-8 KHz
 Number of point in this band
- 7. The frequency of highest peak: Frequency with maximum amplitude.

(Ref. JISHA, 1991).

REVIEW: OF LITERATURE

There are number of methods by which speech can be analysed spectrally. One such analysis procedure takes a time average of the sound pressure level per cycle across frequency. This measurement is commonly referred to as the long-term average spectrum of speech.

The measurement of the long-term average speech spectrum is made by passing the speech energy through a series of contingenous band pass filter and interpreting the energy at the output of each filter. These average values are then plotted to arrive at the visual representation, a smooth plot by the envelope of the power spectrum of the speech sample (Forby and Monsen, 1982).

Niemaeller and Miller, (197^) studied the spectrum of spoken English language which was measured in 1/3 octave band for 10 men, 10 women and 6 children. They obtained values of peaks for F_{\circ} , F1, F2, F3 and found F0 is deep between F1 and F2 region and at the eardrum the frequency range 1.7 - 4 KHz which is usually associated with F_{a} and many of friction generated component in speech lies only 5-10 dB below the maximum level in the low frequency range.

Dibyandra and Suchita (1991) investigated the effect of age and sex in LTAS for 30 normal subject (15 male and 15 female) by using the vagmi program LTAS and concluded that there was significant difference on Alpha-ratio and Beta ratio but not Gamma ratio between male and female speakers.

They also concluded that mean value of Alpha and Beta ratio for females are greater than that for male ie. energy below 1 KHz is greater than above 1 KHz.

Hicks and Troup (1980) studied the source of spectrum in proforma at singing, they discussed the custom of normalizing voice, source spectra instead of the zero frequency components, which has the most energy. In terms of total energy output the normalize source spectrum in the region of the "singing formant" is clearly much higher for the trained than the untrained voice when normalized the zero frequency component.

Speech power is greater between 100 to 600 Hz where the energy of the fundamental frequency (Fo) of the voice and the 1st formant overlaps. It drops off with increasing frequency above about 600 Hz such that at 10,000 Hz. the level is approximately 50 dB below the peak level measured at lower frequency (Danes and Pinsen, 1963).

The spectra for the group speech are generally comparable both within and across languages. However the general shape of the spectra can be altered depending upon the experimental variables used in given studies (Fant, 1973).

Byrne (1977) studied the spectral differences between samples of male and female speech. According to him, the main differences is that the region of 0.1 KHz and 0.125 KHz is for male speaker whereas female level is much lower although it could be measured precisely in this study. This is explicable by the fact that this frequency region corresponds to the fundamental frequency of male voice.

Weinberg et al. (1930) obtained long-term average spectra using an FFT computing spectrum analyzer. The analyzer was set in cumulative mode for 64 frames with a frequency range 0 to 10,000 Hz and a time window 40 msec. The quantized speech signal was weighted by as harning Fourier transformed and stored for cumulation of function. Each frame represented FFT result for a 40 ms. 64 frames. speech segment. An amplitude spectrum with weighting was plotted at the end of cumulation, thus an amplitude spectrum represented an average long time spectrum derived from a total of 64 "section" made from oral reading by each subject.

Hicks (1980) analyzed source spectrum of professional singing and speaking with normalizing voice source spectra to fundamental frequency component instead of the zero frequency component which has the most energy. According to them the normalized source spectrum in the region of the "singing formant" is clearly much higher for the trained voice than untrained when normalizing to the zero frequency component.

Long term speech spectra were used to objectively differentiate 4 classes of voice according to auditory judgement, normal, mild, moderate, severe degree of hoarseness and also used judgement of degree of roughness and breathiness as well as to differential diagnose based on acoustic analysis.

Yashiyuki Haru (1984) studied in spectral characteristic of airflow signal with oral reading from 5 male and 5 female whose age range was 2-4 years with rainbow passage of reading sample.

They found that the phonetic syllabic spectral component extend to as high as about 25 Hz and 30 Hz. low pass filter. Therefore provides a reasonable compromise which reduces the 4-50 Hz component and very low vocal fundamental frequency.

"Study should be useful in developing strategies of automatic analysis of aerodynamic event during connected utterance".

Byrne (198S) examined the LTAS spectrum for 15 male and 15 female who read passage from magazine and analysis by 1/3 octave band speech spectrum showed decreased energy from 160 Hz to 1 KHz and flat energy from 1 Hz and decreased after 6300 Hz.

So they concluded that the "speech spectrum when analyzed by sound pressure level, 1/3 octave band is more appropriate for the purpose of hearing aid" fitting than analyzed by spectrum level.

The LTAS effect of noise and auditory feedback filtering was found to be similar to those of the changes of vocal intensity according to instruction for both singer and non-singer, when the high frequency components of the auditory feedback was enhanced during reading in noise by both groups of subjects.

Kitzing (1986) obtain long time average spectrum by means of signal analyzes (B&K S033) analyzing 400 frequency lines in the chosen sample using linear average over 1S8 triggered spectra by flat weighting, the analyses were

accomplished in 2 series with different base band frequency range via 0-5 KHz and 0-2 KHz respectively. In order to avoid noise from consonant articulation, unvoiced part of signal were eliminated by a gate, controlled either by EGG signal as by low pass filter device.

Lofquist and Henderson (1987) digitized the speech signal at a rate of 20 Hz and analyzed in frames a 2 ms duration using FFT analysis, pauses, voiceless segment were excluded from the analysis.

The general pattern of loudness matching was similar across subject and subject judgement were found to be highly transmitive. A measure of signal power was found to account only moderately well for the individual data. Incorporating level correction approximately those necessary to achieve equal loudness is recommended in the fitting of programmable hearing aid (Jeny Punch, 1993).

Pseudence Allen, Frederic Weightmare (1998) studied on spectral pattern discriminated by 47 children of 4-9 years age, by using video game formal and the discriminability of the sound was measured both in quite and back ground noise. They concluded that the performance of the younger children

was significantly poorer than the performance of an adult control group.

Yingyongi (1991) analysed spectral vowel which was produced by TEP speakers and normal speakers and compared between normalized spectral energy within a selected high frequency range. For vowel produced by normal speaker, average spectra of vowel could be matched to average spectra vowel produce by TEP speakes by decreasing the spectral stop of their vowel by by 2-3 dB/octave.

Long-term average spectra in male and female showed difference in both low and high frequency band but were similar in mid-frequency region. It was concluded that a single spectra could widely be used to represent spectra of both male and female speakers in that frequency region which is important for hearing aid gain prescription ie 25 Hz through 6300Hz.

The long term RMS 1/3 octave band speech spectrum for male and female talker combined was derived for use in hearing aid prescription (Cox and Moore, 1988).

Allord and Sielson (1985) studied acoustic impedance in free-field. With a micrephones and a spectrum analyzes,

they found the effect of the finite dimension of the panel is negligible. Because of very close spacing between both microphones. This techniques allows precise measurement only for frequencies higher than 500 Hz.

Jaki et al. (1991) studied about dynamic spectral shape feature as acoustic correlates familial stop-consonants-stop. They concluded (1) spectral shape feature are superior to both formants, (2) feature extracted from the dynamic spectrum are superior to feature extracted from static spectrum, (3) feature extracted from the speech signal beginning with the burst onset are superior to feature extracted from the speech signal beginning with the vowel transition.

Frokzaei-Jenson (1976) used the ratio of energy below and above 1 KHz and named it as α -ratio parameter. According to them the amplitude below 1 KHz is independent of the microphone distance, amplitude level etc.

Byrne (1977) studied in spectral analysis of Australian speech, for 30 speaker (15 males and 15 females) and expressed in 1/3 octave band level. They found the spectrum is flat from 1 to 8 KHz and 15 dB greater energy at 0.5 KHz

and lower frequency too. By using average spectrum it would be significant for hearing aid selection and evaluation purpose.

The common practice in hearing aid literature, the speech spectrum which is expressed in spectrum level than 1/3 octave band levels may lead to a different and possibly misleading conceptualization of the relative energy of high and low frequency components. Considering that speech spectra differ greatly from one individual to another, it would seem desirable that speech with a known, preferably average spectrum be used in hearing aid evaluation procedure which involves setting the volume control of a hearing aid to deliver speech at a comfortable listening level.

The long term average speech spectrum was measured at a different recording position 30 cm directly in front of the talker (Reference) and at the tragus of the talkers ear (Ear level position) for 3 groups of subject (Adult male, adult female, and children).

Leonard and Seewald (1991) for all three groups of subjects in LTAS measured at ear level position consisted of more low frequency energy (below 1 KHz) and less high

frequency (2.5 KHz) than did LTAS measured at the reference mic position.

They suggested that the currently used procedures to prescribed hearing aid gain may underestimate the sensation level of a hearing-impaired individuals own amplified speech production at frequency below 1 KHz and over estimate the sensation level of a talker's own speech above E500 KHz.

Result: Hearing aid users are treated more as listeners than as producers of speech, given - the important role of auditory self-monitoring in communication, especially in children.

The long term average spectrum provide information on the spectral distribution of the speech signal over a period of time, such spectra can be used for studies of the human voice and suggest that long-time average spectra are limited in the use for screening purpose but can be useful for documenting changes within a given voice (Ander Lofquist, 1986).

Nataraja (1986) studied three spectral parameter in the voice of dysphonics. They are:

1. The ratio of intensities between 0-1 KHZ to above i-S KHz labelled as AA (AA=01 KHz)

1-5 KHz

- S. Ratio of intensities of harmonics and noise in 2-3 KHz as named it as AC.
- 3. Frequency of 1st formant (AO) which is defined as the frequency with maximum intensity in the range of 300-1000 Hz.

Nataraja <1986) found the following result:

- (1) No significant difference between male and female in terms of AA (Alpha) ratio in normal and dysphonic group and found significant difference between dysphonic group and normal group that is dysphonics had higher intensity in the frequency above 1 KHz than normal.
- (2) There is no significant difference between male and female dysphonics and normal group in terms of AC ratio however a significant difference between the males of the two groups (N dysphonics).
- (3) The 1st formant frequency of the males and females within the dysphonic group showed a significant difference and similar result was found in (N) group.

Thus he concluded that of all the 3 parameters only in AA parameter significant difference in (N) and dysphonic in both males and females.

Lofquist and Manderson (1989) made two measurements on the calculated long-term average spectrum. They are -

- 1) The ratio of energy between 0-1 KHz to 1-5 KHz: According to them this ratio provides a measure of the overall tilt of the sound spectrum. A high value of this ratio indicates that the fundamental and the lower harmonics dominate the spectrum which thus falls-off rapidly. A low value of this ratio shows, on the other hand, that the sound spectrum has a lower spectrum tilt.
- 2) Measurement of the energy between 5-8 KHz: A high level of energy at these frequencies can be associated with noise components of the source in the hypo functional voice (Yanagihara, 1967).

Hartman and Cramon (1984) studied the amount of spectral energy in 1 to 5 KHz range and above 5 KHz.

Since the aim of the analysis is to study the voice source, it is most reasonable to restrict the analysis to voiced portion of the speech signal. In order to find out the effects of inclusion of unvoiced sounds in the speech sample (Lofquist Manderson, 1987) found that the inclusion of voice less segment mostly affects the spectral level above 5 KHz for men voices. However Wonder et al. (1986) found significant difference between the spectras of unmanipulated speech signal and the speech equal signal voiced sound eliminated.

An interesting question about the LTAS analysis concerns the resolution of the measurement. Lofquist and Manderson (1987) addressed this question by analysing recording of the same voice recorded twice on the same day. They found that in all cases, there was a clear difference between the two recordings. They further added that when the voice has been in constant use during the day, the source spectrum has a lower spectral tilt.

Another interesting aspect of LTAS is the amount of speech signal that is to be analyzed to get a stabilized pattern. Lietal, (1969) suggested that after 30 sec. of continuous speech the effect of individual speech sound on

the LTAS will not change significantly regardless of how much more speech is analyzed.

Rashmi (1985) has made an attempt to study the ratio of intensity below and above 1 KHz in the spectral vowel [i] she has concluded that -

- 1) Energy above 1 KHz was less than energy below 1 KHz.
- S) The alpha parameter showed no significant difference between male and female till 9 years. Female group age (9-14 years) and male groups (9-18 years) had shown some change.
- 3) No significant difference between male and female till 9 years was found. After 9 years they showed a change in voice quality both in case of male and female speakers as reflected by changes in ratio. The mean value ranged from 78 to 9E.

Hartman and Cromon (1984) studied the amount of spectral energy in the 1 to 3 KHz range and above 5 KHz in two subgroups of patient. (1) The 1st group showed a voice quality composed of breathiness and tense, which gradually normalized in the follow-up period. (2) The other group initially exhibited (in the follow up) a normal or lax, breathy voice, which subsequently becomes more tense. They found that the

variation of spectral energy and duration of aspiration preceding voice onset, indicates signs of tense and breathy voice production. Thus they concluded that this measure is sufficient to differentiate these 2 groups.

Frit2el and Synberg (1974) studied the preoperative and postoperative long-term average spectra of one case of vocal cord palsy. The pre-operative LTAS exhibited one single peak which was found in the frequency region covered by the fundamental. After the treatment, the fundamental increased by about 6 dB on the average, the partials underlying the first formant by 20 dB and the partials underlying the second formant by 75 dB.

Balaji (1988) found the following result:

- 1) There is significant difference between male and female with reference to α -ratio.
- 2) There is no significant difference between normal-male and dysphonic male and normal female and dysphonic female with reference to α -ratio.
- 3) There is no significant difference between male and female with reference to $5-8~\mathrm{KHz}$.
- 4) There is no significant difference between male and female with reference to frequency of highest peak (F.H.P).

Bopal (986) had studied about $\alpha\text{-ratio}$ (parameter) in adult both males and females age ranging from 16-65 years who were randomly selected.

He reported that there was no significant difference between males and females upto 16-55 years. A significant sex difference was observed between males and females in age range of 56 to 65 years ie. male showing a higher score (value of α -ratio .73) than female (0.73, values of α -ratio). The values ranged from 0.71 to 0.76 in the age range 16-55 years both in the case of males and females.

METHODOLOGY

The present study employed the technique of LTAS to study the acoustic characteristic of Nepali language by comparing between Indo Aryan (Hindi) and Dravadian (Kannada and Malayalam) languages.

The aim of the study was to establish the LTAS characteristic of Nepali language.

SUBJECTS8

16 normal (8 males and 8 females) speaks in the age range of S0 years to 40 years were selected for the study. These subject were grouped into four on the basis of their native language (a) 4 speakers (2 males and E females) whose native language is Nepali (Indo Aryan). (b) 4 speakers (2 males and 2 females) whose native language is Hindi (Indo Aryan). (c) 4 speakers (S males and 2 females) whose native language is Kannada (Dravidian). (d) 4 speakers (2 males and 2 females) whose native language is Malayalam (Dravidian).

These subjects were free from any speech and hearing or ear, nose, throat problem at the time of experiments.

Table-1: Highlighting the details of the subject in terms of age, sex and different native language.

Different language	Male		Female	
	Age	No.of Subjects	Age	No.of subjects
1. Nepali	20-40 yrs.	2	20-40 \	rs. 2
2. Hindi	80-40 yrs.	2	20-40 3	yrs. 2
3. Kannada	20-40 yrs.	2	20-40 3	yrs. 2
4. Malayalam	20-40 yrs.	2	20-40	yrs. 2

MATERIAL

- a) Nepali Passage: Which was taken from Nepali popular

 Newspaper (Gorakha patra) which explained about refugee

 people in Bhutan.
- b) Hindi passage: which was taken from Hindi popular Magazine (Shasita) and written by popular writer, the passage explained about life history of writer.
- c) Kannada passage: which was taken from Kannada Passage which is being routinely used at the Department of Speech Science, All India Institute of Speech and Hearing, Mysore for Speech and Voice Analysis.
- d) Malayalam passage: which was taken from short story book (story about thirsty crow).

EXPERIMENT SET-UP FOR LTAS MEASUREMENT

Samples chosen for each language were recorded to computer. The sample thus recorded were digitized into PC-AT using a IS bit A-D (Analogue to Digital) and D-A) Digital to Analogue) converter with a sampling frequency of 16 KHz. The digitized signal was analyzed for LTAS with the LTAS program using auto-correlation method developed by Voice and Speech System, Bangalore.

INSTRUMENTATION:

The instrumentation used was as follows:

- i) Directional mic.
- S) A/D converter (Analogue to Digital converter)
- 3) Computer with video display
- 4) Output printer

Diagram of set-up is given in Figure-1

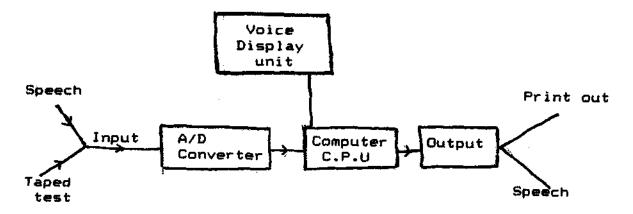


Fig.1: BLOCK DIAGRAM OF INSTRUMENTAL SET-UP.

PROCEDURE:

The procedure consisted of 3 parts (A) Recording (B)

Analysis (C) Comparision.

A) Recording: Subject were enstrusted to read the passage at a comfortable pith and loudness for a duration of 50 sec. and the stable portion of the sample for 10 sec was analyzed for LTAS.

Speech samples we. recorded individual in a quiet room. Recording was made using a hi-bias metal cassette using a professional philips stereo cassette deck (type F611S) and dynamic microphone (MD 433 Sennheisen) with flat frequency response from 50-15,000 Hz. The mic-to-mouth distance was approximately 15 cm (6"-8") for all subjects.

The level indicator of the tape recorder was adjusted to see that the speech level was reasonably steady and was at a suitable level to ensure good recording. This recorded material was subjected to spectral analysis.

B) <u>Spectral analysis:</u> For carrying out spectral analysis the program "LTAS" with autocorrelation was used.

The spectral analysis were carried out for the frequency range 0-8 Hz to study 1.5 kHz etchergy of different parts of frequency (500 Hz, 1K, 1.5 kHz etchergy)

The spectral analysis of data is obtained in the following way (Samples enclosed in Appendix-1).

- 1) Graphic display and spectral pattern in frequency range of 0 Hz 8 KHz was studied.
- S) A data of the energy of all the different point which were analyzed by the computer was noted.
- 3) Alpha ratio (α)
- 4) Beta ratio (β)
- 5) Gamma ratio (δ)
- 6) Energy in the frequency band 5-8 KHz
- 7) Energy below 1 KHz
- 8) Energy above 1 KHz
- 9) Fo (frequency of highest peak (energy)

These data were then subjected to appropriate statistical analysis to find out if there were acoustical differences in the Nepali language and other Indo Aryan (Hindi) and Dravidian language (Kannada Malayalam) through LTAS parameter.

Comparison: Acoustic characteristics of spectral pattern in frequency range of 0 Hz - 8 KHz (each 500 Hz external) and value of δ , β , δ ratio and Fo of Nepali language (Indo Aryan (Hindi) and Dravidian language (Kannada and Malayalam).

RESULTS AND DISCUSSION

The present study was undertaken to compare the acoustic characteristics (long-term average speech, of Nepali Language (Indo-Aryan Language) with Hindi Language (Indo-Aryan Language) and Nepali language with Malyalam (Dravidian and Nepali language with Kannada language) language in terms of the following parameters.

- 1. α ratio
- 2. ß ratio
- 3. δ ratio
- 4. Energy between frequency band 5-8 KHz
- 5. Frequency of highest peak (FHP)
- 6. Energy below 1 KHz
- 7. Energy above 1 KHz
- S. Graphic display and spectral pattern in frequency range of $0-8\ \mathrm{KHz}$.

To data obtained was subjected to statistical analysis and the results have been discussed under following heading:

1 α ratio: (Energy between 0-1KHz/i-S KHz):

There was no significant difference between the α ratio of Nepali and Hindi language (Indo-Aryan) and α -ratio of Nepali (Indo-Aryan) language and Malayalam and Kannada

(Dravidian) language on T-test at 0.05 level. However some difference existed between languages. While the mean of α -ratio Hindi and Malayalam language were high and Nepali and Kannada were low. This range of α -ratio for Hindi and Malayalam was wide compare to that of Nepali and Kannada language.

Table-1: Shows the mean, standard deviation and range of α -ratio for all the 4 languages.

Language	Mean	S.D	Range	
Nepal	4.70835	6.3073	1.237 - 15.70838	
Hindi	21.24963	11.7642	8.8368 - 33.566	
Malayalam	34.99059	43.68	8.3439 - 88.4489	
Kannada	14.87624	19.246	4.8129 - 28.91782	

β-ratio:

There was significant difference between the β - ratio of Nepali and Hindi (Indo Aryan) and Nepali and Malayalam

(Dravidian language at 0.05 lev3el on T-test. Moreover there was no significance difference between the $\beta\text{-ratio}\,\text{of}\,Nepali$ Thand Kannada language. The $\beta\text{-}$ ratio showed the same trend Theas the $\alpha\text{-}$ ratio, ie. mean and range of $\beta\text{-}$ ratio of Hindi and

Malayalam were high and Nepali and Kannada were low.

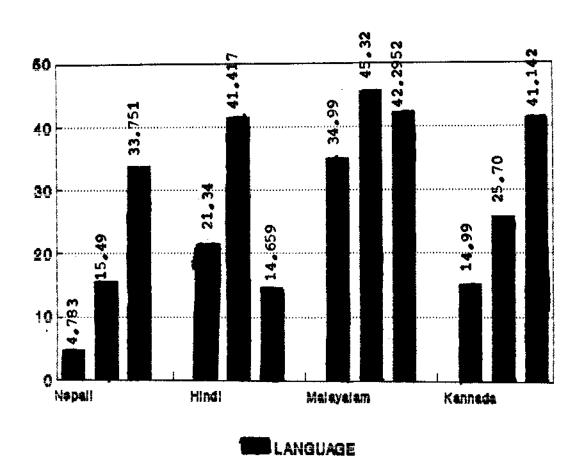


Fig. shows Comparison and 3 different rates in four different languages.

Table-2: Shows the mean, SD, and range of β -ratio for all the four languages.

Language	Mean	S.D	Range		
Nepali Hindi Malayalam Kannada	15.493366 41.417782 45.32828 25.7058	8.9541 10.268 13.693 17.1306	5.1836 - 28.5024 58.16972 - 30.5784 30.18417 - 66.972 11.4839 - 48.1294		

δ -ratio:

There was no significant difference between the δ - ratio of the Nepali and Hindi language (Indo-Aryan). On the t-test <.05> level. However, there were significant difference between the δ -ratio of Nepali and Kannada, Nepali and Malayalam (Dravidian) language the mean of β -ratio of Nepali, Malayalam and Kannada were high and Hindi language was low. The range of the δ -ratio for Nepali, Malayalam, Kannada was wide compare to that of Hindi language.

Table-3: Shows the mean, SD, and range of δ -ratio for all four languages.

Language	Mean	S.D	Range
Nepali	35.7510	15.00193	49.9306 47.03368
Hindi	14.6349	8.3098	4.366098 - 28.2789
Kannada	46.295	16.78	28.9185 66.0471
Malayalam	41.1412	15.678	18.3705 60.2371

Energy of FHP: (Frequency of Highe.t Peak>i

There were significant difference, between the frequency of higest peak of Nepali and Hindi and Nepali and Malayalam language at 0.05 level is 't' test. However, there was no significant difference between Nepali and Kannada language. While mean, SD and range of FHP of Nepali and Kannada were high and Hindi and Malayalam were low.

Table-4: Shows the mean, SD and range of FHP for all the four languages

Language	Clean	S.D	Range		
Nepali	578.15	147.38	406.3 - 812.95		
Hindi	573	40.61614	250 - 243		
Kannada	350.49	31.S49	218 - 281.8		
Malayalam	382	10.43	250 - 500		

Energy between 5-8 KHz:

There was significant difference between the energy between 5-8 KHz and Nepali and Hindi (Indo aryan) language and Nepali and Malayalam (Dravidiana) language at the level of 0.05 using the t-test. However, some difference existed between Nepali and Kannada languages.

In mean, range and SD had found similar results in all the four languages.

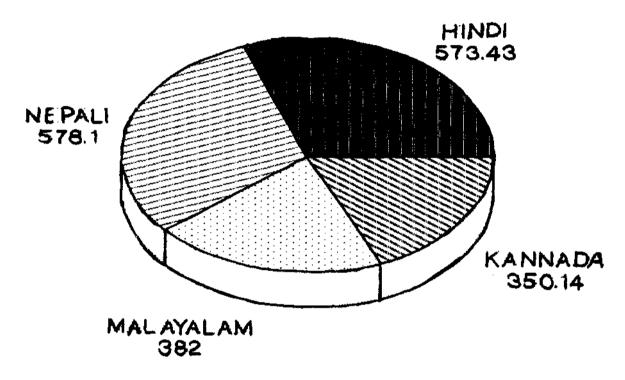


Fig. Shows the mean, SD and range of FHP for all the four languages.

INTRODUCTION

Asked to select the most precious of the five senses, few people would name hearing. Yet of all man's links to the outside world, hearing seems to be the essential sense, the one that makes man peculiarly human. Hearing is a late development in evolution but it has become the sentinel of our senses, always on the alert. The acquisition and monitoring of speech, the detection of potential danger, the elementary feeling of existing in a living universe - all depend upon the auditory modality.

Throughout waking life, the mar receives an uninterrupted stream of messages from the outside world ie audible messages which are screened, sorted and acted upon or filed away. How precious hearing is becomes clear when it is lacking. Ear, a sensitive organ may be damaged due to infection, exposure to noise, congenital malformations,

OtOtOXIC liiuya, .-___is_ __,

many more resuu *ii

· I · * * · • • j

Table-5: Shows the mean, SD, and range of Energy between 5-8 KHz for all four languages.

Language	Mean	S.D	Range
Nepali	68.048013	6.59037	75.10248 - 62.74112
Hindi	58.95443	3.65635	55.4504 - 63.2045
Kannada	58.8485	5.2729	51.7536 - 64.88528
Malayalam	56.356.75	3.4303	52.8974 - 61.8765

Energy below 1 KHz:

There was no significant difference between the Nepali and Hindi languages at the level of 0.05 level on t-test in Energy below 1 KHz. However there was significant difference between Nepali and Kannada and Nepali and Malayalam language at the level of 0.05 on t-test.

In mean, range and SD of were found similar results in all the four languages.

Table-6: Shows the mean, SD and range of Energy below 1 KHz for all the four language.

Language	Mean	S.D	Range		
Nepali	87.82867	1.10693	89.725 - 86.941		
Hindi	89.744335	.1796	89.5323 - 89.966		
Kannada	89.762485	.216	89.4997 - 89.952		
Malayalam	89.54936	.2854	89.1802 - 89.85843		

Energy above 1 $^{\rm KHz}$

There were significant differences existed between Nepali and Hindi languages and Nepali and Kannada and Nepali and Malayalam languages at t-test at 0.05 level on Energy above 1 KHz.

Table-7: Shows the mean, SD and range of energy above KHz for all the four languages \bullet

Language	Mean	S.D	Range		
Nepali	84.41293	3.8038	86.7524 - 77.76445		
Hindi	75.7466	4.4227	80.07014 - 68.9993		
Kannada	75.92969	3.81664	70.48 80.2809		
Malayalam	78.73417	S.595	78.089 82.35806		

Tafale-8: Shows the result of t-tett cm 0.05 level

Nepali vs. Hindi Language at 0.05 level

	α-ratio	β-ratio	δ -ratio	FHP	5-8 кнz	Below	Above
	u lacio	ртасто	0 10010	FIII	J 0 10112	energy	energy
						< 1KHz	> 1KH
D	10.47	11.18	17.95	161.82	5.4373	.9152	4.839
FD	7.4	7.86	8.79	113.95	3.844	.647	3.368
t	2.24	3.305	2.12	4.868	S.3656	.29608	2.657
Level 0.05	_	++	-	++	-	_	++
Level 0.01 _	-	-	-	++	-	-	-
Nepali vs. Ka	annada						
SD	9.134	9.34	16.75	197.53	6.06	.933	3.758
SED	6.458	6.603	11.848	104.308	4.24	.6596	2.253
t	1.56	1.5466	31.09	1.872	1.872	2.6695	2.530
Level 0.05	_	_	•+	_	++	++	++
Level 0.01			++		-		
Nepali vs. Ma	alayalam						
SD CE	36.0	18.37	123.01	6.89113	.8472	13.35	4.22
SED	25.3	12.99	86.97	4.972	.598	9.438	2.98
t	1.18	29.96	3.77	1.8883	2.944	3.16	2.846
Level 0.05		++	++	++	_	++	++
Level 0.01			++	++			B

NOTE: - No significant differences between 2 languages ++ Significant differences between 2 languages.

DISCUSSION

The result indicates significant difference between Nepali and Hindi (Indo-Aryan) languages, Nepali and Malayalam (Indo-Aryan vs. Dravidian) as β -ratio at 0.05 level of 't' test.

This may be because of Nepali sample which had been chosen for the study consisted of more fricatives sounds (Eg. Sangyogita Rastra Sangh) with more of high frequencies.

There is a striking similarity between Nepali (Indo-Aryan) and Kannada language on most of the parameters.

This may be because of most of the Nepali speaker of present study were residing (studying) in Karnataka State and hence they might be influenced by Kannada language or they are exposed to Kannada language.

 $\alpha\text{-ratio}$ - There is no significant difference between Nepali and Hindi language and Nepali and Kannada and Nepali and Malayalam language.

This may be because of no influence of difference language in this parameters.

There is high value of mean and wide range in Hindi and Malayalam language and narrow range and low mean in Nepali and Kannada language. And SD is very high in Malayalam.

 β -ratio: There was significant difference between Nepali and Hindi language and Nepali and Malayalam language, but there were not significant difference between Nepali and Kannada language may be because of all Nepali speakers who were studying in Karnataka State and expose to the Kannada language.

Mean is very low in Nepali language and high in Malayalam and Hindi language in β -ratio, SD and range found to be similar.

 δ -ratio: There is no significant difference between Nepali and Hindi language and-in δ -ratio on t-test at 0.05 level.

May be because of geographically and genetically similarity of both Indo-Aryan language.

Frequency of Highest Peak (FHP): There was significant difference between Nepali and Hindi and Nepali and Malayalam at 0.05 level in t-test of FHP parameters, but there was no significant difference between Nepali and Kannada language.

May be because of all Nepali speaker exposed to Kannada language.

SD is very high in Kannada with Nepali language when as very low in Hindi and Malayalam language.

Mean is very low value in Malayalam and Kannada language and more in Nepali and Hindi language.

Range is high in Nepali and Kannada and low in Hindi and Malayalam language.

(5-8 KHz): Energy between There was no significant difference between Nepali-Hindi language, and Nepali Malayalam language in this parameters but there was significantly difference between Nepali and Kannada language.

Range, mean and SD are similar value in those all four language.

Energy below 1 KHz: There was no significant difference between Nepali and Hindi language.

May be both language comes under the Indo-Aryan which are genetically and geographically similar.

But significant difference between Nepali and Kannada,
Nepali and Malayalam may be becaust of dissimilarities of
Indo-Aryan and Dravidian language.

May be this parameters is depends on particular language characteristics.

Energy above 1 KHzs There was significantly difference between Nepali and Hindi and Nepali and Kannada and Nepali and Malayalam language.

May be because these parameters contain energy only above 1 KHz which was influenced of different language. So this parameters were different in different languages.

SD, mean and range are similar to all four languages.

Table-5 (Energy between 5-8 Hz), Table-6 (Energy below 1 KHz)

and Table-7 (Energy more than i KHz): Shows similar result

in all SD, Mean and range aspects in all these four

languages.

May be this parameter are not effected (influence by other languages).

SUMMARY AND CONCLUSION

The present study was carried out to study the characteristic (long-term average speech spectrum) of Nepali language by comparing with Hindi (Indo-Aryan), Malayalam (Dravidian) and Kannada (Dravidian) language.

The parameters of LTAS were studied in four (2 male and 2 female) normal speakers in each language. (Nepali, Hindi, Malayalam, Kannada)>

LTAS parameters are:

- 1. Alpha α -ratio (Alpha)
- 2. β -ratio (Beta)
- 3. δ -ratio (Gamma)
- 4. Energy between frequency band 3-8 Hz
- 5. Frequency of highest peak (FHP)
- 6. Energy below 1 KHz
- 7. Energy above 1 KHz.

These parameters were analysed using computer Vagmi programme vagmi developed by Voice and Speech System, Bangalore (Dr.Anantapadmanabhan).

The data was further subjected to statistical analysis and following results were obtain.

- 1. There is no significant difference between Nepali and Hindi language, Nepali and Kannada and Nepali and Malayalam language in α -ratio.
- 2. There was significant difference between Nepali and Hindi languages and Nepali and Malayalam languages but there were no significant difference between Nepali and Kannada languages in β -ratio.
- 3. There is no significant difference between Nepali and Hindi languages in δ -ratio.
- 4. There was significant difference between Nepali and Hindi and Nepali and Malayalam in FHP parameter.
- 5. There is no significant difference between Nepali vs. Hindi language and Nepali vs. Malayalam language in Energy between (S-8 KHz). There is significantly different between Nepali and Kannada language.

- 6. There is no significant difference between Nepali vs.

 Hindi language in Energy below 1 KHz. But there is significantly difference between Nepali and Kannada and Nepali and Malayalam language.
- 7. There is significantly difference between Nepali and Hindi, Nepali and Kannada and Nepali and Malayalam languages in Energy above 1 KHz.

Limitation:

- 1. There was limited number of sample.
- S. Only few parameter has been taken.

BIBLIOGRAPHY

- Agnello, J., and Wagner, W. (1974): Speech spectra and perception in the design of hearing Aid. Chapter 3, 85-106. In Interpreting hearing aid technology. (Ed. Donnelly, K.). C.C.Thomas, Springfield, Illinois.
- Balaji, J. (1987): Long-term average spectrum and electroglottography in dysphonics. Unpublished Masters Dissertation, University of Mysore, Mysore.
- Benson, R.W., and Hirsh, I.J. (1953): Some variable in audio spectrometry. Journal of Acoustical Society of America, 25, 499-505.
- Berger, K.W., Klich, R.J., Millin, J.R. (1982-84): The speech spectrum in hearing aid fitting. Audicible, Vol-31, 34, 18.
- Besnd Weinberg, Yashiyuke, Morti and Bonnie, E. Smith (1980):

 Long time spectral and intensity characteristics of oesophageal speech.

 Journal of Acoustical Society of America, Vol.67, 1981.
- Bryne, D. (1977): The speech spectrum some aspect of its significance of hearing aid, selection and evaluation. British Journal of Audioloav 11, 40-46, 1977.
- Cox and Moore, J N. (1988): Composite speech spectrum for hearing aid gain prescription. Journal of Speech Hearing Research, 31, 102-107.
- Dibyendu Bhattacharya, Suchita Vaidya, and Thermar, S.(1991), Effects of age and sex on long-term average spectrum. Journal of Indian Speech and Hearing Association, 8-48.

- Farmby, C., Monsen, R.B. (198B): Long-term average speech spectrum for normal and hearing-impaired adolescent. Journal of Acoustical Society of America, 71, 196-206.
- Flanogen, J.L., and Christenson, S.W. <1980>: Computer studies on parametric coding of speech spectra. Journal of Acoustical Society of America, Vol.68, 420.
- Harish, C.M., and Waits, W.M. (1965): Measurement of speech spectra recorded with a close-talking microphone. Journal of Acoustical Society of America, 37, 936-927.
- Hicks, P., and Trocip, G.N. (1980): Source spectrum is professional singing some points of Be Made. Pholia Phonetic, 32, 23-28.
- Hollien, H., and Mojedieski, W. (19) the Speaker identification by long-term spectra under normal and distorted speech condition.

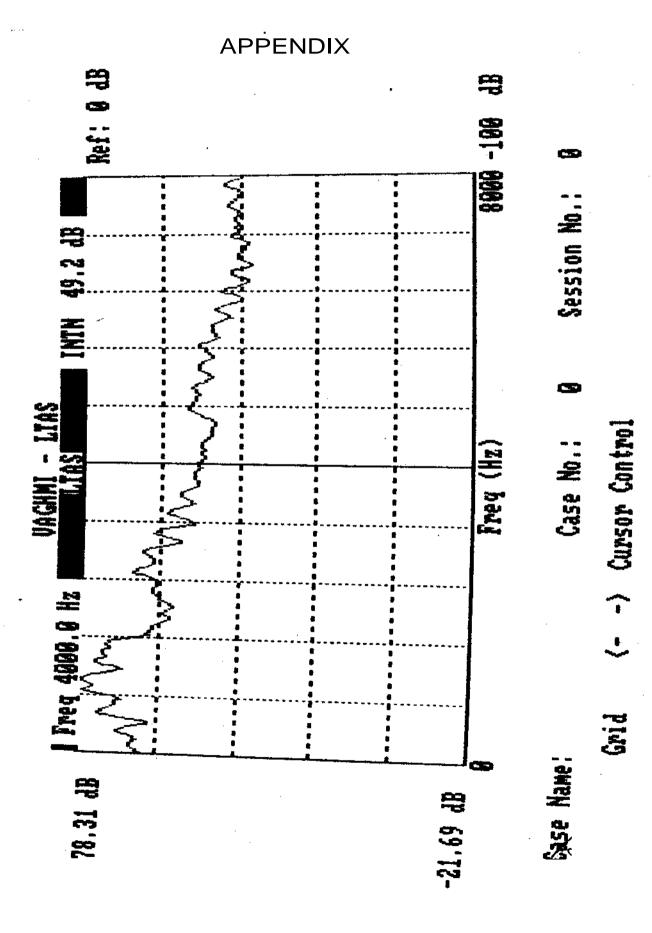
 Journal of Acoustical Society of America, 62,979-980.
- Jean, R., Allorer and Benita Sietan (1985): Measurement of acoustic impedance in a free field with three mice and a spectrum analysis. Journal of Acoustical Society of America, Vol.27, 1619.
- Jerry Punch and Brad Rakerd (1993): Loudness, matching of signals spectrally shaped by a simulated hearing aid. Journal of Speech Hearing and Research, 36, 2, 352.
- Joellen Wilbus, Robert, McDonald (1992): Non-linear analysis of cyclically correlated spectral spreading in modulated signals. Journal of Acoustical Society of America, Vol.92, 1, 29, 1992.

Karl-Swami (1987): On spectral coarticulation in stop vowel, stop syllable implication for automatic speech recognition. Journal of phonetic, Vol.5, 85-100.

- Kiukoannienu,: H., Siponen, P., Mattricle, P. (1982):
 Individual differences in the long-term speech spectra. Pholia Phonetic, Vol.34, 21-28.
- Komm, C.A., and Derk, D.D. (1982)! Some effect of spectral shaping on recognition of speech by hearingimpaired listener. Journal of Acoustical Society of America, Vol.71, II, 1211.
- Leonard, E., Cornelise, Jean-Pierse Sagne and Richard C Secwald (1991. Ear level recordings of the long-term average spectrum of speech. Ear and Hearing, Vol.12, No.1, 47-54, 1991.
- Lofquist, A. (1986): Long-term average spectrum criteria pertinent to the measurement of voice quality. Journal of Phonetic, 14, 477-482.
- Lofquist, A. (1986): Long-time average spectrum is a tool in voice research. Journal of Phonetic, 14, 471-476.
- Lofquist, A., Manderson, B. (1987): Long-term average spectrum of speech and voice analysis. Pholia Phonetic, 39, 831-229.
- Mariken tes keuss, Joots, M., Festen, and Reinics Plamp (1992), Effect of spectral envelope smearing on speak receptive. Vol.91, 887
- Niemeller, A.F., McCormick, and Land Miller, J.D. (1974): On the spectrum of spoken English. Journal of Acoustical Society of America, 55-461.
- Olsen, W.D., Hawkins, D.B., and Van Tasel, D.J.(19): Representation of the long term spectra of speech. Ear and Hearing, 5(suppl), 100-102, 1987.

- Patrivie-Bramming and Johan Sunberg (1988): Spectrum factor relevant to phonetogram measurement. Journal of Acoustical Society of America, Vol.88, S33S.
- Patricio, B., Stelmochowisz, Annel, Moc Judy G Kapun, and Edward Corney (1993): Long term and short-term characteristic of speech implications for hearing aid selection for young children. Journal of Speech Hearing and Research, Vol.36, 609.
- Pausowang-Gelfer, A., Parker Massey, K., and Mallien, H. <1987): The effects of sample dewation and timing on speaker identification occurring by means of long-term spectra. Journal of Phonetic, 17, 3S7-338.
- Prakash, S.G. (1989): Few objective measurement of voice quality in hearing-impaired children.
 Unpublished Master Dissertation, University of Mysore, Mysore.
- Rudimase, H.W., Clark, K.C., Carlson, F.O., Eisenstein, J.C., and Walker, R.A. (1958): Voice measurement with an audio spectrometer. Journal of Acoustical Society of America, S0, 503-513.
- Ruth, Salebky Kamen and Watson, B.C. (1991): Effects of long term tracheastomy on spectral characteristic equal production. Journal of Speech Hearing and Research, 34, 1057.
- Sadaoki-Furvi (1986): On the role of spectral transition for speech perception. Journal of Acoustical Society of America, 80, 1016.
- Schutte, H.N., Miller, R.(1984): Resonance balance in register categories of the singing voice: A spectral analysis study. Pholia Phonetic, Vol.86, 289-295.

- Stevens, S.S., Egan, J.P., and Miller,G.A. (1947): Methods of measuring speech spectra. Journal of Acoustical Society of America, 19, 771-780.
- Sundeera, J., Ternstron, S., Perkins, W.H., and Gramming, P. (1988): Long-term average spectrum analysis of phonatory effects of noise and filtered auditory feedback. Journal of Phonetic, 16, 203-220.
- Wayneo, Olsen, David, B., Hawkins, and Sianne, J., Van Jasel (1887): Representation of the long-term spectrum of speech. Ear and Hearing, 8, 1005.
- Wendeer, J., Boberty, E.T., and Hollien, H. (1980): Voice classification by means of long-term average spectrum. Pholia Phonetic, 32, 51-60.
- Yashiyuki Harri (1986): Some spectral characteristic of air flow signals associated with oral reading. Philia Phonetic, Vol.38, 157.
- Zaki, B., Massair and Zaharian, S.A. (1991): Dynamic spectral shop feature as acoustic correlates for emitted stop consonant. Journal of Acoustical Society of America, Vol.89, 6978.
- Zolewski, Mojewski, W., and Mallien, H. (1975): Crosscorrelation of long-term average speech spectra as a speaker identification technique. Acoustica, 3, 20-24.



: Showing LTAS for Nepali language

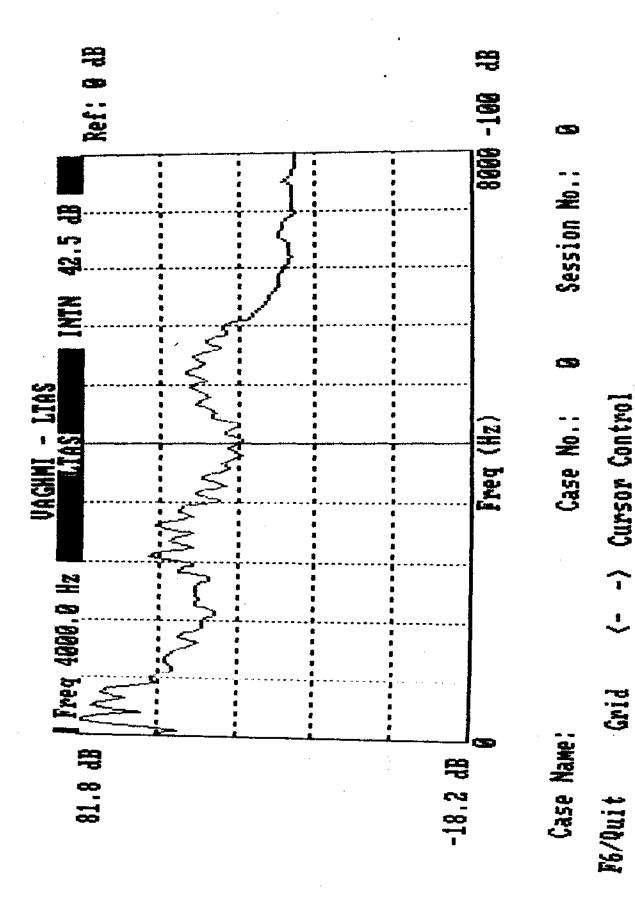


Fig. 2: Showing the LFAS for Malayalem language

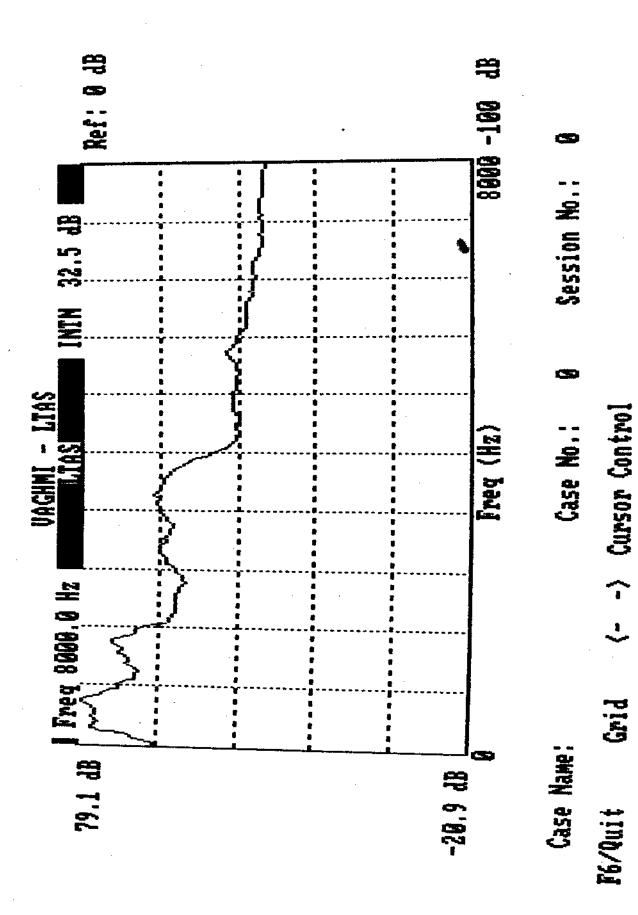
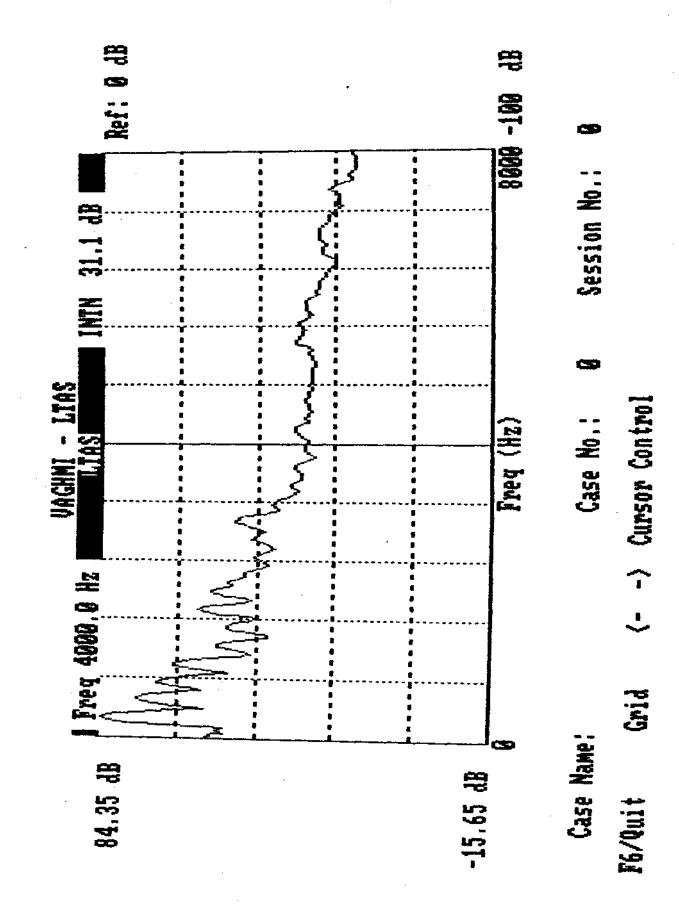


Fig. 3: Showing LTAS for Kannada language



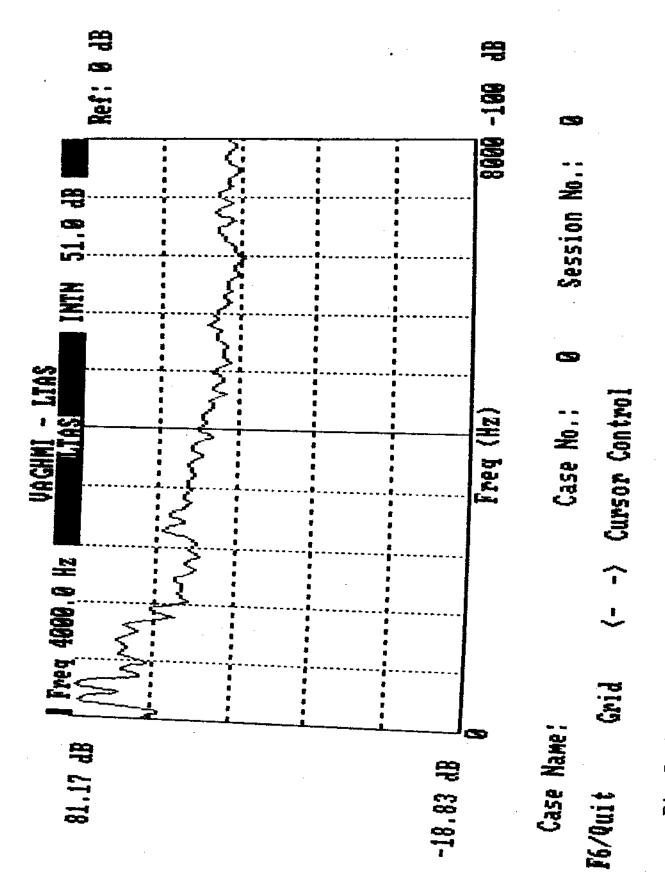


Fig. 5: Showing the LTAS for Malayalam language.

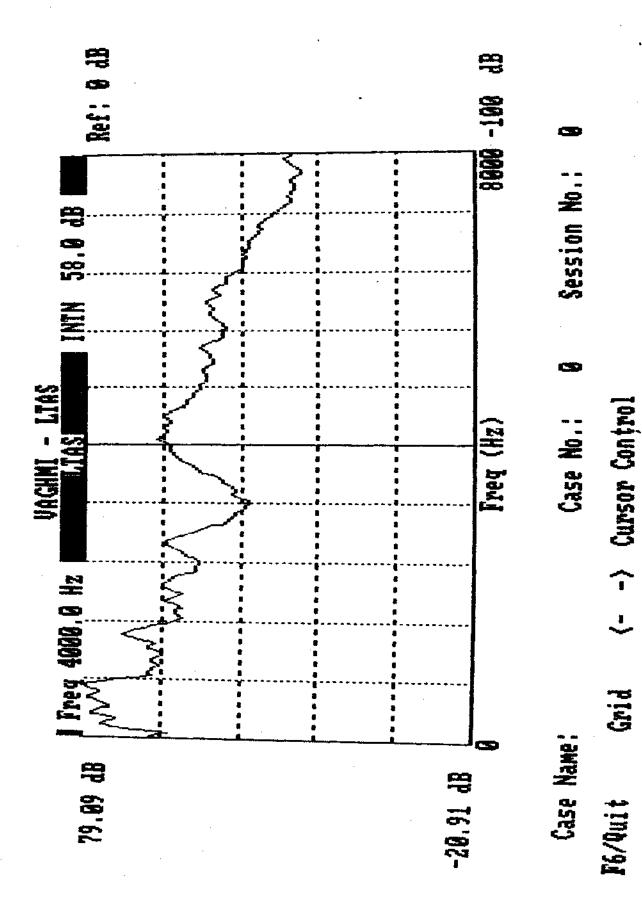


Fig. 6 Showing the LTAS for Hindi Language

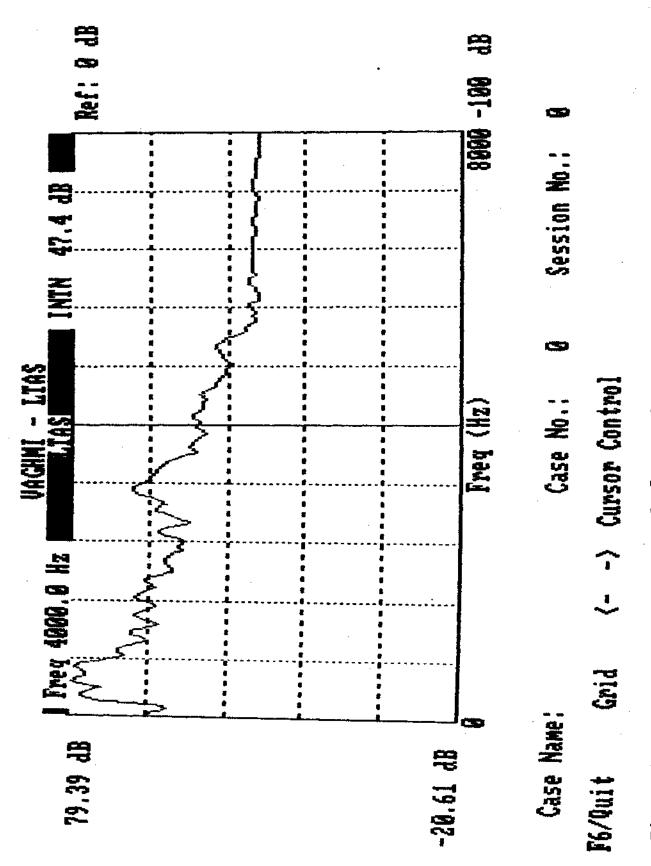


Fig : Showing the LTAS for Kannada Language.

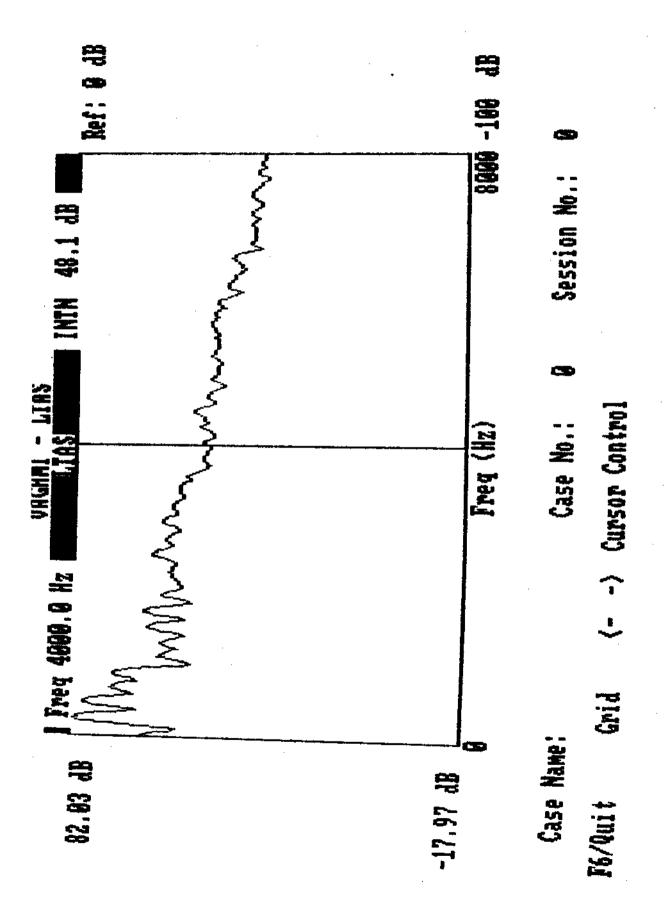


Fig.: 7 Showing the LTAS for Nepali Language

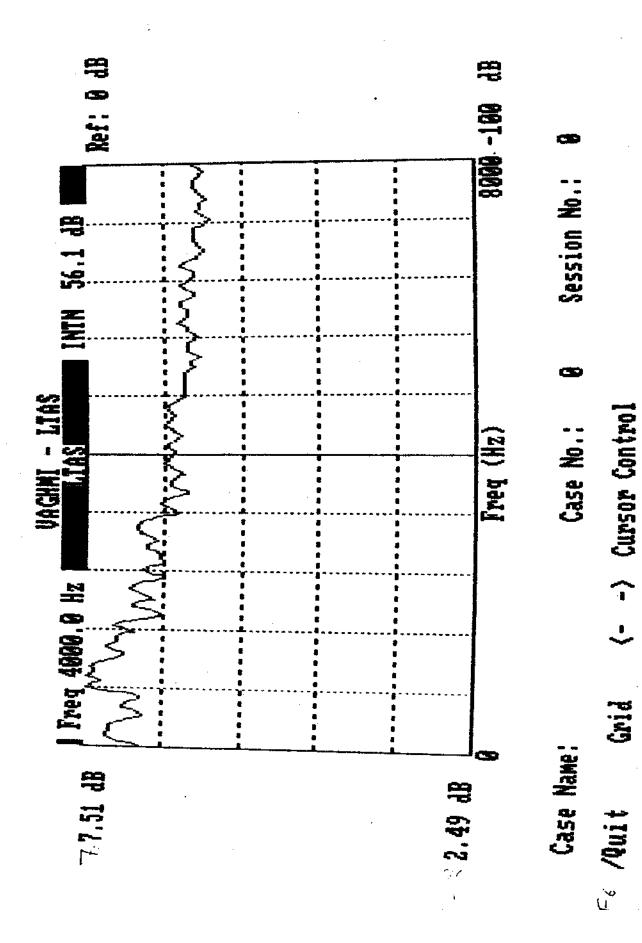


Fig: @ Showing the LTAS for Nepali language.