LONG TERM AVERAGE SPEECH SPECTRUM IN MALAYALAM

Reg. No. **M01**12

Independent Project as a part fulfilment of
First Year M.Sc, (Speech and Hearing),
Submitted to the University of Mysore,
Mysore.

ALL INDIA INSTITUTE OF SPEECH AND HEARING
MYSORE - 570 006

May, 2002

Dedicated to

My Lord Almigthy,

Pappa, Mamma and Jinu

........You mean the world to me

Certificate

This is to certify that this Independent Project entitled ***"LONG TERM A VERAGE SPEECH SPECTRUM IN MALAYALAM"*** *is the bonafide work in part fulfilment for the degree of Master of Science (Speech and Hearing) of the student with Register No. M0112.*

Mysore,
May, 2002

Dr. M. Jayaram
Director
All Indian Institute of Speech and Hearing
Mysore - 570 006.

Certificate

*This is to certify that this Independent Project entitled **"LONG TERM AVERAGE SPEECH SPECTRUM IN MALAYALAM"** has been prepared under my Supervision and Guidance.*

Mysore
May, 2002

Asha Yathiraj.

Dr. Asha Yathiraj
Head of the Department of Audiology
All India Institute of Speech and Hearing
Mysore - 570 006.

DECLARATION

This is to certify that this Independent Project entitled **"LONG TERM AVERAGE SPEECH SPECTRUM IN MALAYALAM"** is the result of my own study under the guidance of Dr. Asha Yathiraj, HOD of Audiology, 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, 2002

Reg. No. M0112

ACKNOWLEDGEMENT

I would like to express my heartful gratitude to **Dr. (Mrs) Asha Yathiraj,** my teacher and guide for her constant guidance and support. Ma'am thanks for being patient even with the most silly mistakes I used to make.

I am grateful to **Dr. M. Jayaram** (Director) AIISH for permitting me to carry out this study.

Pappa and Mamma, you peak above all the rest and I am proud to be your daughter.

Tinu, little sis, you *are* the sweetest thing god has given me.

Sonal, your friendship is a blessing to me.

May the lord bless you and keep you;

and make his face shine upon you

Chaya, your love and patience has not gone unnoticed. You are a,very friendj Thanks, for never saying no, being there whenever I needed you.

Vinu, Delcy

A heart that receive from god has so many riches to give others.

May god draw you deeper in his heart and open new ways for you to deposit the riches of his love.

Rakhi, Amala, you are great gals, thanks for your sincere help in times of necessity.

Chandni, it was soothing sharing the tension with you at 8.pm.!

Ramya - Our friendship is something which I can't define you have always being special in some mysterious way.

Vijayalakshmi & Joby Sir, thanks for all your help during this project work.

Kripal I enjoyed my 1st M.Sc, with you.

Sal, thanks for being such an understanding friend. I value our friendship.

All my classmates : As I always say you guys are simply the best!!

Sapna and Divya Ma'am your moral support definitely boosted my morale.

Deepa you have been a wonderful friend. Thanks for the wake up calls.

Bhuvna, Deema, Mili, Meenakshi its nice to have juniors like you.

A special word of thanks to **Acharya Sir** for helping me with my statistics and **people at Softouch** for their quick and efficient work.

Above everyone to my Lord Almightly for

"If is God that girdeth me
with strength
and makes my way perfect
Thou last enlarged my steps me,
That my feet did not slip "

CONTENTS

LIST OF TABLES

LIST OF GRAPHS

"Behold, we put bits in the horses mouth,
that they may obey us, and we
turn about their whole body
Behold, also the ships, which though
They be so great,
and are given of fierce winds,
Yet are they turned about with a very small helm,
Even so the tongue is a little member,
And so boasteth great things "

The Bible

INTRODUCTION

Speech is the most important sound we hear. We communicate most effectively when we can clearly hear other people speak to us and we speak clearly to them (Skinner, 1984). Speech sounds are recognised by their acoustic properties i.e., frequency composition, relative intensities and changes over time.

The primary goal of any aural rehabilitation and amplification methods is to maximize the use of residual hearing for speech. For a given hearing impaired individual, it is required to assess, as accurately as possible, the audibility of the speech signal throughout the frequency region that contains information necessary for speech recognition.

The long term average spectrums of speech (LTASS) is a direct measure of the mean sound pressure level of speech as a function of frequency and provides a global representation of the acoustic characteristics of continous discourse (Cox and Moore, 1988). The measurement of LTASS is made by passing energy through a series of continous band pass filters and interpreting the energy at the output of each filter. These average values are then plotted to arrive at a visual representation (Fqfnby and Mqsen, 1982).

Representations of the LTASS have various audiological and acoustical applications. These applications include hearing aid prescription procedures, calculation of articulation index, identification, classification of voice disorders, vocal aging and infant cry analysis.

In rehabilitative audiology the LTASS is widely used in many hearing aid prescription procedures either in the derivation of the prescriptive formulae (Byrne & Fifiled, 1974; Ross, 1975; Davis & Hardick 1981; McCandless & Lyregaard, 1983; Seewald, Rose & Spiro, 1985; Byrne & Dillon, 1986) or in calculating the prescription for the individual client (Cox, 1983; McCandless, 1985; Skinner, 1988). Skinner (1988) suggested that the LTASS be amplified close to the most comfortable loudness level. Others like Cox (1988) proposed that the amplified speech spectrum bisect the long term listening range which is the area between auditory threshold and upper limit of comfortable loudness. Seewald, et al. (1985); Shapiro, (1976) suggested that the speech spectrum be amplified to the specified desired sensation level differing as a function of frequency and hearing loss.

In recent years there has been considerable interest in the application of articulation index (AI) which is derived from the LTASS, in the evaluation and rehabilitation of the hearing impaired. The AI gives an indication of a person's ability to recognise speech. (Hodgson, 1977, 1988; Pavlovic, 1984, 1989; Pavlovic & Studebaker, 1984; Studebaker & Sherbecae 1985). Unaided as well as aided AI can be calculated for an individual. Several authors have suggested formulae to calculate AI (French & Steinberg, 1947; Pavlovic, Studebaker & Sherbecae, 1985; Skinner 1984). These formulae have been used in determining the frequency response of a hearing aid. The hearing aid that results in the highest Al value would be the one selected for the individual (Byrne 1977).

In the area of speaker identification, the long term averaged spectrum has been utilized by numerous investigators (Doherty, 1976; Hollien & Mejewski, 1977; Doherty and Hollien, 1978). Doherty and Hollein (1978) used LTASS approach for identification of twenty-five male speakers who spoke under normal and stress conditions. Correct identification of 100% was obtained for full-band (80-12500 Hz) analyses under laboratory and normal speaking conditions. 80-85% correct identification resulted for band pass conditions (315-3150 Hz). However, relatively long samples (20-90 seconds) appear to be necessary for identification.

The LTASS provides information on the spectral distribution of speech signal over a period of time. Such spectra have been used for studies of the human voice source. Wendler, Doherty & Hollian (1980), in their study claims that measurements derived from long term spectra allow classification of voice disorders, although other studies have contradicted this (Prytz, 1977, Somninen and Hurme, 1982, cited in Wendler et al. 1980).

The differences between the classical trained singers and country singers have also been studied using the long term average speech spectrum by Cleveland, Sundberg, Stone (2001). Their results indicated that the "Singer's formant" around 2.8 Hz seen in trained classical singers were not observed either in the singing or speech of country singers.

Another area where LTASS has found a diagnostic application is as a tool differentiating term and pre-term babies (Goberman and Robb, 1999). A difference between the groups was observed for the 1st spectral peak, which was higher for infants in the pre-term group whereas there was no difference between the groups in the mean spectral energy and spectral tilt.

Linvella and Rens (2001) used LTASS for vocal tract resonance analysis of aging voice. They concluded that with aging the amplitude of the formant peaks decreased, which was attributed to the changes in vocal tract taking place.

With the LTASS having several applications, it is necessary to see if the ones reported in literature can be readily applied in India or whether it varies as a function of various Indian languages.

Aim

The aims of the present study is to derive a long-term average speech spectrum in Malayalam. This would be obtained for adult males, adult females and children. In addition, a common LTASS would be obtained for all the three groups. A comparison of the four LTASS will be made. This would be further compared with the speech spectra obtained in other studies, which are reviewed in literature, to see whether there is any significant difference.

Need for the study

The majority of published measurement of the LTASS are for the English language ad spoken in the U.S.A. (Niemoller, Mc Cormick and Miller, 1974; Cox and Moore, 1988; Comelisse, Gagne and Seewald, 1991; Stelmachowicz, Mace, Kopunand Carney, 1993), Australia (Byrne, 1977; Byrne and Dillon, 1986) and England (Boothroyd, 1967, cited in Byrne et al. 1994). There are also measurements for other languages which include: German, Hungarian, Italian, Russian (Tarnoczy, 1956, cited in Byrne et al. 1994), Swedish (Aniansson, 1974, cited in Byrne et al. 1994), Finnish (Kiukaanniemi, 1980, cited in Byrne et al. 1994), Mandarin (Me Coullough, Lew, 1993, cited in Byrne et al. 1994) and Spanish (Banuls Terol, 1971, cited in Byrne et al. 1994). All values of the LTASS are similar but small differences occur between English spoken in different countries and among different languages. However, large differences among individuals within languages is also present (Byrne, 1977; Kuikaanniemi, Soponen and Mattila, 1982).

There is a need to know whether the LTASS of other languages differ from that of the Indian languages. If these is no difference, then the existing LTASS which are being used for various applications such as derivation of different prescription formulae, calculation of Articulation index can continue to be used.

However if the LTASS markedly varies from that mentioned in the literature it may be necessary to alter the test formulae, where LTASS has been applied. This would include having different hearing aid prescriptive formulae for each language and having different values for calculating articulation index.

REVIEW

Speech is a transient signal where the amplitude and frequency component change with respect to time. The analysis of speech signal is essential in the area of speech transmission, speech perception, speech synthesis, speech identification and in the area of phonetics.

There are a 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 long term average spectrum of speech and provide a means of viewing the average frequency distribution of energy in a continous speech sample (Formby and Monsen, 1982).

The first attempts to outline the speech spectrum was made by Crandall and Mackenzie, 1922 (cited in Formby and Monsen, 1982) using a condenser microphone whose output was analysed by a series of resonance circuits. Later studies done by Dunn and White, 1940 (cited in Formby and Monsen, 1982) also used band pass filters to separate the frequency components.

Most estimates of the speech spectrum indicate a maximum intensity around 500 Hz with a decrease in intensity at higher frequencies. However, the overall level and shape of the speech spectrum vary across studies (Byrne, 1977; Byrne and Dillon, 1986; Cox and Moore, 1988; Olsen, Hawkins and Van Tasell, 1987). These differences may be attributed to the following :

a. Sample difference such as language, age, sex and vocal effort of speaker (Byrne, 1977; Cox and Moore, 1988; Cox, Matesich & Moore, 1988; Schwartz, 1970)

b. Procedural differences like the type and length of speech sample, number of subjects, the distance and the relative azimuth of the speaker to the measuring microphone (Olsen, Hawkins and Van Tasell, 1987; Cornelisse, et al. 1989; Boothroyd and Medwetsky, 1992; Studebaker, 1985).

c. Analysis methods such as bandwidth, sampling rate, windowing (Popelka & Engebretson, 1983; Stelmachowicz, Mace Kopun and Carney, 1993; Angew, 1999).

Hence, the experimental variables influencing the speech spectra can be categorised into:

I. Sample related

1. Language
2. Sex
3. Age
4. Vocal effort

II. Procedure related

1. Length of speech sample
2. Number of the trials
3. Position of the microphone
4. Type of recording
5. Reverberant Vs. anechoic sound fields

III. Analysis methods

1. Frequency range

2. Analysis procedure

The manner in which these variables could affect the long term average speech spectrum are mentioned as evidenced from various studies.

I. Sample related differences

1.Language

The language used for the study plays as role in determining the overall level and shape of the speech spectrum. Kiukaannimi and Mattila (1980) compared the LTASS of Finish and English subjects using 50 phonetically balanced words in each language. The results showed a significant difference in the distribution of speech power in Finish at 253-619 Hz while in English it was around 619-3 Hz. The spectrum obtained by Byrne (1977) for Australian English was different from that obtained by Cox and Moore (1988) for American English in the 1-2 kHz frequency region. It could be attributed to the different vowels encountered in Australian and American English respectively. The frequency range 1-2 kHz has the second and third formants of American English vowel /z/ which does not occur in Australian English. Taranczy and Fant (cited in Byrne et al. 1994) compared Swedish. Hungarian and German languages. It was found that there was a significant differences among the languages in the mid-frequency region (700-1500 Hz for males, 1000-2000 Hz for females).

On the other hand, from a comparison of the LTASS of French, Dutch, English, Italian and Danish where two males and two female talkers for each language, Pavlovic (1989) concluded that there was no significant effects of either sex or language. The drawback of the study is the limited number of subjects taken for each language. Byrne et al. (1994) also obtained similar findings while comparing 13 languages which included English, Swedish, Danish, German, French, Japanese, Cantonese, Mandarin, Russian, Welsh, Singhalese, Vietnamese and Arabic. They concluded that although different languages use different vowels (formant structures) and the frequency of occurrence of various phonemes differ, these factors appear to have only minor effects on LTASS.

In a study done by Pradhan (1994) a comparison was made between the LTASS obtained in Nepali language with that of Hindi, Kannada and Malayalam. It was observed that there was a significant difference in the energy concentration between Nepali and Malayalam at frequencies below 1 kHz and in the frequency range 1-5 kHz and no significant difference in the 5-8 kHz range. However a significant difference between the other language with respect to Nepali was not seen.

The LTASS in Hindi obtained by Mrinal (1998) showed more power of energy in the frequency region 156-625Hz after which there was a gradual decrease in the intensity. While there were large individual variation in the general shape there was no significant difference between the male and female spectra. However, there was a significant difference in the mean values between his LTASS values and the data reported by earlier investigators at frequency 250-8 kHz.

Overall it appears that the LTASS may vary with language. However, the issue is still not resolved.

2. **Sex**

While comparing the individual differences in long term speech spectrum, Kiukaannimi, Siponen and Mattila (1982) showed that the speech spectrum of females generally slope more at frequencies above first formant while males have a significant flatter spectrum but larger individual variations between the spectra at frequencies above 1 kHz. This variation may be due to the larger difference in the oral cavity, in which F_2 in mainly generated. The fact that males have more speech power at frequencies 1-3 kHz could contribute to the flatter spectra.

Fletcher (1953) while comparing males and females noted some differences. The average F_o for male voice was 125 Hz and female voice was approximately one octave above the male voice. The female voice had considerably more power than male voice above 4 kHz.

Another study by Mendoza, Valencia, Munoz, Tinjieluo (1996) concluded that there is a significant correlation between gender and frequency level. The female voice had more aspiration in the region corresponding to the third formant which made it more breathy.

From these studies it can be noted that there does exist a difference between the LTASS of males and females, this difference being more pronounced in the higher frequencies.

3. **Age**

Niemoller, Mc Cormick and Miller (1974) studied the spectrum of spoken English language which was measured in 1/3 octave bands for 10 men, 10 women and 6 children. They obtained values of peaks for F_o, $F_{1,}F_2$, F_3 and found F_2 is deep between $F_1$and F_3 region. The frequency region 1-4 kHz was about 5-10 dB below the maximum level in the low frequency region and the overall intensity level of children were higher than that of adults. Had the authors scaled the speech samples of the three groups to equalize the intensity levels, such intensity differences may not have been seen. Also, the authors have failed to mention the age range of the children taken for the study.

The long term average speech spectrum of 10 adults males, 10 adults females and 10 children in the age range of 8 years. 6 months to 12 years 5 months were compared by Cornelisse, Gagne and Seewald (1991). The results revealed that the overall level for the group of children was the same as the mean overall level for the group of male subjects. However the mean overall energy of females were 2.5 dB lesser than males. Similar findings were reported by Pearson 1977, cited in Cornelisse et al. 1991.

Rashmi (1985) made an attempt to study the ratio of intensity below and above 1 kHz in the spectral vowel / i / for the age group 4-9 years and 9-14 years. She noted that the energy above 1 kHz was less than energy below 1 kHz. The alpha parameter, which is the ratio of the energy above and below 1 kHz, was not different between males and females till 9 years of age. Further, after 9 years, both the males and female showed a change in voice quality. With respect to the difference in speech spectra due to the age of the subjects taken for the study, no conclusive results were drawn.

While experts like Cornelisse et al. (1991) find no difference in the speech spectrum between the adults and children, others like Niemoller, Mc Cormick and Miller, (1974) conclude that the overall intensity level of children is higher than that of adults. This variation in intensity may not have been seen had the subjects been trained to produce speech at a particular level, or the speech sample had been scaled to overcome individual variations in intensity.

4. Intensity

a. Overall intensity level

The intensity at which the speaker is recording his speech could vary the speech spectrum obtained. Schwartz (1970) compared the LTASS of oral and whispered speech. Seven young adult females were taken and the results obtained revealed not only a reduction in intensity of whispered spectrum but also a marked flattening in the spectrum. Others like Pearson (1977) cited in Cornelisse et al. (1991) and Keikaannimi et al. (1982) have concluded that difference between soft and average speech are small. Byrne et al. (1994) concluded that greater vocal effort resulted in an increase in the relative mid to high frequency content of speech.

The overall level and configuration of speech energy chosen to represent everyday speech have resulted in different speech spectra, (Angello & Wagner, 1974; Brandt, Ruder & Shipp, 1969; Dillon, 1988). While the speech spectrum for the Byrne and Tonnison (1976) procedure is based on measurement of Australian English spoken at 68 dBSPL, the speech spectrum for Cox procedure (1983) is based on American English at an overall level of 70 dBSPL. The spectrum for the Pascoe procedure (1978) is the sensation level at which normal hearing persons hear third octave bands of speech when overall level of broad band speech is 65 dBSPL.

b. Back ground Noise and Spectrum intensity

Lau and Wong (2000) studied the effects of background noise on LTASS. Recordings were done in quiet (30dB), babble noise (50dB), traffic noise (65dB) and restaurant noise (80dB). In noise, spectral level increased by 3dB for every 10dB increase in noise regardless of the type and spectral shape of noise and the peak shifted from the low to the mid frequencies. A high level of energy at the frequencies 5-8 kHz has been associated with the noise components of the source (Yanagihara, 1967).

These studies thus indicate that the speech spectra obtained would depend on the overall intensity level measured, with the background noise levels controlled which would inturn determine the amount of gain which needs to be provided to the hearing impaired.

Procedural Variables

1. Length of Material

One of the most important variable concerning use of LTASS is the length of material necessary to obtain a stable signal. Lofquist and Manderson (1987) analysed the reading of a standard text during increasingly shorter time segments. They did not find differences in the signal when the analysis time was reduced to half of the text's total length. Li and Hug hes (1974) suggests a sample of 30-40 sec to reach stability in LTASS. The effects of individual sounds in LTASS is not significant after 30-40 sec of continuous speech.

While some studies have reported no difference between the spectra for talkers reading a two minute passage and repeating the "Joe...........Lawn" (3.4sec) passage, Holte, Margolie, (1987), Byrne (1977), have contradicted these finding. According to them "Joe.........Lawn" passage is not representative of the phoneme distribution in English and therefore a longer passage should be considered.

Lofquist et al. (1987) suggested analysis be made of reading a standard text to minimize variations due to phonetic structure and provide a more natural speech sample than sustained phonations. If the whole signal is analysed deleting the pauses, the ratio of energy in the frequency range 0 - 1 kHz to 1-5 kHz was 0.32 and the relative energy level in 5-8 kHz was 732. The corresponding values for the voiced only and voiceless only condition were 0.34 and 0.07 and 478 and 1732 respectively. The inclusion of the voiceless segments therefore affects the spectrum level above 5 kHz.

It should be noted that in every day speech, both voiced-voiceless speech sounds occur. Hence, a sample that include only voiced or voiceless speech sounds will not represent everyday speech. From the above research it can be noted that it is also necessary that the passage should be long enough, representing the phonemes of the languages.

2. Number of trails

Lofquist and Manderson (1987) studied 7 males and 10 women of the age range of 20-25 years who had to read a passage that lasted for 3 minutes in normal intonation and reading rate. The passage was read five time during a two-week period. The results indicated that the greatest difference between sessions are present in the lowest frequency of the spectrum, in the frequency band around 4 kHz. This was attributed to the fact that the configuration of F_4 has been associated with labial radiation particularly determined by the shape and size of lips that can be modified by subjects in subsequent readings of the text. Similar results were obtained by Owens (1961).

They had also found differences in LTASS in relation to the time of day taken for measurement. There was a reduction in the formant peaks towards the end of the day. This could be attributed to the stability of the voice by the end of the day. Nittrouers, McGowan, Milenkovic & Bechler, 1990 have contradicted there findings. However, subsequent studies have not been conducted to study the effects of this variable.

3. Position of Microphone

One important procedural variable is the angle of incidence of the recording microphone to the talker's mouth. Cornelisse, Gagne and Seewald (1991) investigated the extend of differences between the LTASS measured at the ear level of a talker and that of when the microphone is placed directly in front

i

of the talker. The results indicated a decrease of 10 dB in the level of high frequency energy (above 2500 Hz) of the LTASS obtained at ear level position relative to that in front. This is most likely due to the directional radiation characteristics of the mouth where high frequency energy is attenuated to a greater extend than low frequency energy.

In a study by Dunn and Farnsworth, 1939 (cited in Cornelisse et al. 1991) comparison was made between the LTASS obtained from a microphone positioned directly in front of the talker's mouth and at 80 different microphone placement around the head. An attenuation of the band levels in the high frequencies was observed as the microphone was moved from directly in front to behind the talker. Similar findings have been reported by Studebaker (1985).

Thus, it reveals that the LTASS measured at the ear-level of a talker consists of more low frequency energy and hence hearing impaired listeners would prefer less gain at low frequencies.

4. Type of Recording

Two types of multi-talker speech spectra have been employed. They are the sequential spectra and simultaneous spectra.

A simultaneous spectrum is obtained by measuring the long term average RMS spectrum produced by several talkers talking together and the recording been done at one time. This type of spectrum has been employed in hearing aid prescription by Cox (1983) and Pascoe (1978). In contrast, a sequential spectrum is obtained by measuring the long term RMS speech spectrum for each of the several individual talkers and arithmetically averaging the obtained levels across talkers. This has been used by Byrne and Tonnison (1976).

For the purpose of hearing aid gain prescriptions, the sequential type of spectrum seems to be more appropriate because it accurately represents the average levels in the speech of individual talkers.

A major drawback of using a simultaneous spectrum is that as the recording is done with a number of individuals talking simultaneous, the angle of incidence of the recording microphone is difficult to compute as it will vary for each individual. Moreover editing the pauses during analysis of this procedure is difficult, as the time where one individual paused would be overlapping with the time another individual produced an utterance.

5. Reverberant / diffuse sound fields Vs anechoic sound fields

Owing to the directionality of the human mouth / head and torso, the high frequencies compared to low frequencies will be mainly radiated in the frontal direction, while the diffuse field will represent a spatial integration of radiation in all directions (Dunn and Farnsworth, 1939, cited in Byrne et al. 1994).

Furthermore, most room surfaces have a greater absorptivity for high frequency sounds than for low frequency sounds. Reverberant of diffuse fields will tend to have relatively weaker high frequency components when compared to anechoic sound fields. Consequently the relative high frequency content will be higher in the frontal / anechoic conditions (Byrne et al. 1994).

Analysis Methods

1. Frequency range

For many years, the range from 500-2000 Hz was called the speech range probably because individuals with normal sensitivity across the range would understand simple speech in quiet (Pascoe, 1978). However studies done by Skinner and Miller (1983) and Pascoe (1978) highlighted the utility of frequency at least as high as 6000 Hz when functioning in noise.

There is great variations in the delimitation of the frequency to be employed while carrying out measurements. Some have chosen constant frequency range (Kitzing, 1986) while others have chosen a logarithmic scale that stimulates response of human ear (Wendler, Doherty and Hollian, 1980) or the area of spectrum with a clear acoustic and physiological correlation (Lofquist, Manderson, 1987).

Majority of the studies have utilized there systems of measurement in order to compare normal and pathological voices or to determine post surgical effects. At present there is lack of sufficient criteria to decide which is the most appropriate procedure for measurement. However it is advisable to include a frequency range from 250 Hz till about 8 kHz for understanding speech in quiet as well as in the presence of noise. While some have analysed frequency bandwidth restricted to frequency below 3 kHz. (Pittam, 1985; Pritz, 1976, cited in Lofquist et al. 1987) others have extended this band upto 16 kHz (Wendler, Doherty and Hollian, 1980).

In the light of clinical evidence there is recent emphasis in hearing aid design on extending high frequency range of hearing aids. Much of the energy of voiceless phonemes /s/ /sh/ *IV* , /ch/ are above 4 kHz, (Zeng, Turner, 1990; Boothroyd & Medwetsky, 1992). The reception of high frequency sounds have both phonemic and morphemic implications as these sounds form plural and past tense makers. Thus even though the spectrum level decreases towards the higher frequencies, it definitely plays a role in speech intelligibility and hence should be considered for analysis.

Analysis procedure

Kryter (1962 } cited in Studebaker and Hockberg, 1985) compared two analysis procedures which were sound pressure level in 1/3rd octave bands and decibels per frequency within each band. It was found that the former was more appropriate for the purpose of hearing aid fitting.

In order to assess the effects of varying analysis time, Lofquist and Manderson (1977) analysed the readings of a standard passage in successively smaller portions. When the time was reduced to 1/3, 1/4, 1/5 of the original analysis, variations occur between measurements. These variations were most likely due to the differences in the segmental structure of the analysed parts. Other sources of variation were the onset and offset of voicing, transition between voiced and unvoiced segments as well as stress and pitch patterns.

Weinberg (cited in Sadaoki-Furvi, 1986) obtained the long term average spectra using Fast Fourier Transformation computing spectrum analyser. The analyser was set in cumulative mode for 64 frames with a frequency range 0-10,000Hz and a time window of 40 sec. Each frame represented the FFT result for a 40 msec speech segment. An amplitude spectrum with weighting function was plotted at the end of cumulation. Thus, an amplitude spectrum represented an average long term spectrum derived from the total of 64 sections made from oral reading by each subject.

At present, there is not sufficient criteria to decide which is the most appropriate analysis method. However, Mendoza, Munoz and Naranjo, (1996) have suggested analysing the greatest number of levels possible in the frequency range with the aim of determining if the signal remains constant throughout the spectrum.

From the review, it is evident that the long term average speech spectrum obtained varies depending on a number of parameters including measurement and analysis methods. Hence care should be taken to obtain a comprehensive speech spectrum representative of everyday speech ensuring proper control of all the variables wherever possible.

It would be ideal to take a speech sample which is representative of all the phonemes in the language under study and give a number of trials to the subjects so that while recording the continuity of the sample is maintained. A sequential spectrum would be preferred since it would accurately represent the intensity level at which an individual speaks as the distance and angle of the microphone from the site of recording can be made constant. However while recording it is undesirable to position the microphone directly in front of the mouth because of breath noises from plosive sounds. Care should also be taken so that the noise levels are well below that of speech and it does not interfere with the recordings. Thus, it would be desirable to establish a standard LTASS, controlling as many variables as possible.

METHOD

The below mentioned method was adopted in order to obtain the long turn average speech spectrum of Malayalam.

Selection of subjects

- 15 adult males and female each in the age range of 20-40 years and 15 children (8 boys, 7 girls) in the age range of 8-10 years were taken for the study.
- The mother tongue of all the subjects was Malayalam. Care was taken to ensure that there was equal representation of the dialects spoken in North and South Kerala.
- All the subjects had normal speech and language and no history of any hearing or otological problem.
- A minimum education level of high school pass for all the adults subjects and a completion of II standard for the children was required.

Selection of Material

The reading material was a 2.5min. passage from a Malayalam course book of the II standard. The passage consisted of a description of the life of Gandhiji. The selection was such that the passage represented all the phonemes of the language. The phonemes were chosen according to the frequency in which they occur as given by Ramakrishnan et al (1962). All the words in the passage were familiar to the subjects. Care was taken so that the material was unemotional and could be easily read by everyone.

Instrumentation

Recordings were made using a Sony portable minidisk recorder (MZ R-70) with a unidirectional microphone (F 500).

The tape recorder sample were downloaded into a computer. The software programme creative sound blaster 16 was used for editing the software program. The Audiolab software was used to compute the LTASS.

Environment

The recordings were carried out in a quiet environment where the ambient noise levels were well below the speech levels.

Procedure

Recording

The subjects were instructed that they had to read the passage with minimal errors. To ensure that the subjects read the passage fluently, a typed copy of the passage was distributed to all the subjects in advance. In order to let them be familiar with the material, the subjects read the material aloud 2-3 times. Samples were recorded individually in a quiet room with the microphone placed 15 cm in front of the mouth of the speaker. The subjects were instructed to read the passage at a comfortable pitch and loudness i.e., with normal vocal effort. For a few subjects, additional instructions to speak slightly louder or softer were given. Excess variations in the loudness of their voice was avoided by asking the subjects to observe the VU meter deflection on the tape-recorder.

Analysis

The speech samples recorded on separate tracks in the recorder were then re-recorded into the computer. Editing of the recorded sample was done using the program Creative Mixer, Sound Blaster 16, so as to maintain the continuity of the sample. After editing, the samples were scaled using the Audiolab software program with settings 16 bit mono, sampling frequency of 22050 Hz at a rate of 43 kB/sec. This was done to minimize the intensity variations between samples. The LTASS program of Audiolab software was used to obtain the speech spectrum across the frequency range from 43 Hz-10,000 Hz. A graph was obtained with the frequency across the x-axis and the intensity across y-axis. The intensity levels at octave and mid-octave frequencies across the range of 250-8 kHz were noted. The 10 frequencies which were considered were 250 Hz, 500 Hz, 750 Hz, 1000 Hz, 1500 Hz, 2000 Hz, 3000 Hz, 4000 Hz, 6000 Hz and 800 Hz.

As the program did not permit to obtain the values at the exact frequencies mentioned above the intensity levels were obtained at the following frequencies. 246 Hz, 496 Hz, 746 Hz, 996 Hz, 1496 Hz, 1996 Hz, 2996 Hz, 3995

Hz, 5994 Hz, 7994 Hz. The resulted in variation from the frequencies by just 1.6% to 0.075%. Since the variation from the target frequency that was to be measured was so minimal, for practical purpose it would be considered equivalent to target frequency.

All the individual spectra in each group in adult females, adult males, boys and girls were accumulated by adding the speech files so as to get an average speech spectrum for each group. The accumulated speech spectra for each group was again added so as to obtain a single speech spectrum which is a representation of all the samples taken up for the study.

Statistical Analysis

The data thus obtained was subjected to statistical analysis. The mean, standard, deviation and 't' test were calculated at each frequency for adults and children; for adult males and adult females separately and for the unified spectrum.

RESULTS AND DISCUSSION

The objective of the present study was to obtain the long-term average speech spectrum in Malayalam. Three groups comprising of 15 males (Group I), 15 females (Group 2) and 15 children were taken for the study.

The data collected from the three groups were subjected to statistical analysis. The mean, standard deviation and 't' values were calculated for the data that was obtained. The analysis was done for the following :

I Comparison of the spectra of:

1. Males Vs. Females
2. Adults Vs.Children
3. Adult females VsXhildren
4. Adult males Vs.Children

II. Comparison of the composite or unified spectra of the three groups obtained in the present study with that of studies reported in literature.

I. Comparison of the Spectra

1. Males Vs. Females

While comparing the spectra obtained in adult males and females, it was found that the overall energy concentration was the same with more energy at the lower frequencies and progressive reduction towards the higher frequencies (Graph 1). The significant difference between the mean reveals that at lower frequencies the two groups did not differ statistically. However a significant different at the 0.05 level was observed above 2 kHz with females having higher energy (Table 1). The energy was higher by 2-6 dB with the maximum difference being present at 3 kHz. This finding is supported by previous research conducted by Byrne et al. 1994, where LTASS of 15 languages were compared. Their results indicated that in three language groups (Australian English, English as spoken in Memphis, USA and Mandarin) the energy was higher by 3-4 dB above 2 kHz. In general, they too, as in the present study, found the energy to be more in the higher frequencies. Fletcher (1953) suggested that this could be because female talkers emphasis sibilant sounds more than male talkers do.

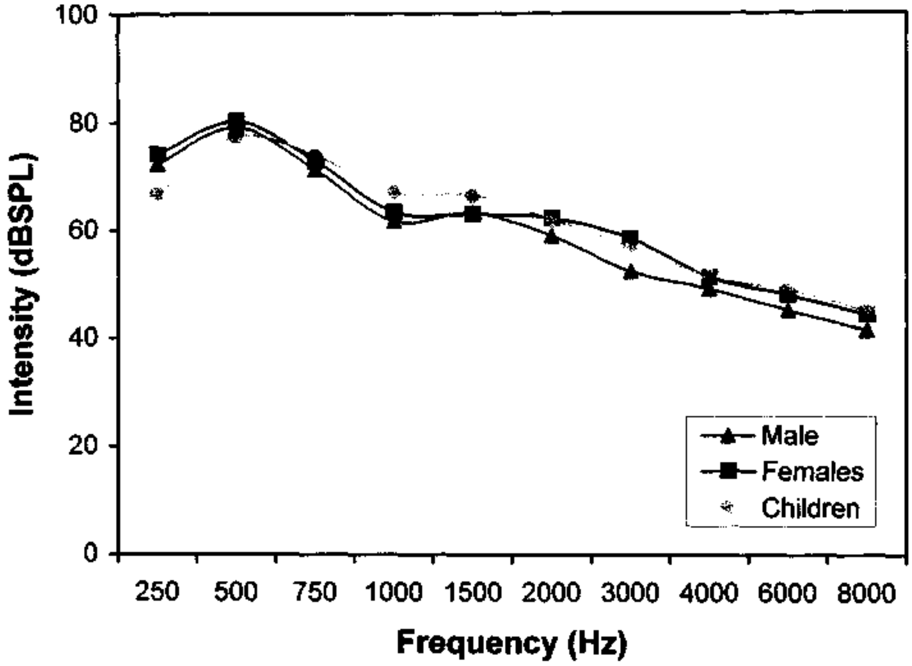
The present study revealed that the mean overall level for the group of males was 2 dB greater than the mean overall level for the group of adult females. This finding is consistent with previous investigations (Cox and Moore, 1988).

TABLE 1 : Comparison of Male Vs. Female Speech Spectra

Frequencies	Males		Females		T value	Significance
	Mean	S.D.	Mean	S.D.		
250 Hz	72.34	3.52	74.11	4.56	1.194	.243
500 Hz	79.28	2.62	80.31	2.66	1.072	.293
750 Hz	71.35	2.10	72.85	3.76	1.346	.189
1000 Hz	61.78	4.07	63.33	4.09	2.025	.069
1500 Hz	63.21	3.66	62.88	3.17	0.262	.795
2000 Hz	59.03	3.09	62.16	3.15	3.342	.032*
3000 Hz	52.37	4.90	58.42	3.36	3.948	.001**
4000 Hz	49.19	2.92	51.19	4.12	3.926	.004 **
6000 Hz	45.23	2.79	47.97	4.19	2.104	.044*
8000 Hz	41.58	3.42	44.35	4.40	2.249	.049*

* P<0.05 ** PO.01

Graph-1 : Comparison of the Spectra of males, females and children in Malayalam



2. Adult Vs. children

The LTASS obtained for adult males and females were combined and compared with that of the children. The significance of difference between the means for the LTASS of adults and children revealed that the energy concentration at 250 Hz and 500 Hz was more in adults while at the frequencies 750 Hz and 1000 Hz children had more energy (Table 2, Graph 1). Though the difference was significant at the 0.01 level at 250 Hz and 1 kHz, it was significant at the 0.05 level in two in-between frequencies. The higher energy for adults at the lower frequencies could be attributed to the concentration of speech power which is more in adults compared to children (Kent, 1993; Peterson and Barney, 1952 cited in Kent and Read, 1992).

The intermediate frequencies of the spectrum is determined by the configuration of the vocal tract (Flanagan, 1958; Mendoza et al. 1996). Since children have shorter vocal tract, the energy concentration at these frequencies will be higher for them. In addition, the 1^{st} formant of children is in the range of 600-1000 Hz (Kent, 1993). The 1^{st} formant contributes maximally to the energy in speech and this could be another reason as to why children had more energy at 750 Hz and 1000 Hz.

TABLE 2 : Comparison of Adults Vs. Children Speech Spectra

Frequencies	Adults		Children		't' values	Significance
	Mean	S.D.	Mean	S.D.		
250 Hz	73.23	4.10	66.86	6.05	3.675	.001**
500 Hz	79.86	2.67	77.50	5.04	2.067	.045*
750 Hz	71.77	3.60	74.03	2.65	2.151	.037*
1000 Hz	62.55	4.40	67.08	3.08	3.562	.001**
1500 Hz	63.05	3.37	66.46	3.40	0.286	.704
2000 Hz	59.76	3.61	61.96	3.23	1.013	.312
3000 Hz	55.06	5.58	57.03	2.75	1.282	.207
4000 Hz	50.70	3.95	51.33	3.35	0.525	.602
6000 Hz	47.51	4.29	49.12	2.26	1.356	.182
8000 Hz	43.97	4.12	45.09	3.14	1.920	0.078

* $P < 0.05$ ** $P < 0.01$

3. Adult Females Vs. Children

There was no significant difference between the mean energy concentration of adult females and children except at 250 Hz where females had higher intensity level (Table 3 & Graph 1) and 1 kHz where the children had more energy. In the former frequency the significance of difference was at the 0.01 level while in the latter it was at the 0.05 level. The absence of a significant difference between the groups at higher frequencies, except 1000 Hz, could be because the formant frequencies of both the groups are close and hence have a similar energy concentration.

The fundamental frequency of females being around 250 Hz probably resulted in the power of speech being more in the group at this frequency. This finding has been supported by Kent and Read, 1992 and Pickett, 1980.

TABLE 3: Comparison of the Speech Spectra of Females Vs. Children

Frequencies	Females		Children		't' values	Significance
	Mean	S.D.	Mean	S.D.		
250 Hz	74.25	4.61	66.86	6.05	3.962	.001**
500 Hz	80.44	2.69	77.50	5.04	1.926	.059
750 Hz	72.85	3.76	74.03	2.65	1.002	.330
1000 Hz	69.33	4.09	67.08	3.09	2.012	.047*
1500 Hz	62.88	3.17	63.87	2.58	0.901	.356
2000 Hz	62.17	3.09	61.96	3.24	1.013	.312
3000 Hz	51.71	58.08	3.48	2.75	0.941	.365
4000 Hz	51.59	4.12	51.33	3.35	0.482	.852
6000 Hz	47.97	4.19	49.12	2.26	1.926	.359
8000 Hz	44.35	4.40	46.09	3.14	0.998	.392

* P < 0.05 ** P < 0.01

4. Adults males Vs children

Testing for the significance of difference between the means of adult males and children, it was found that at 250 Hz males had more energy concentration possibly again because of the power of speech being close to the fundamental frequency (Table 4) (Kent and Read, 1992; Pickett, 1980). However at higher frequencies i.e., 1000 Hz, and above 3000 Hz children had significantly had more energy. Among these frequencies the maximum difference was seen at 1 kHz and the least at 4 kHz. This higher energy concentration in children could be attributed to the fact that formant structure of children being similar to that of adult females, the energy is emphasized at the higher frequencies. Another possible reason could be that, as in the case of adult females, children also emphasize the sibilants and fricatives which contribute to the noise components in the source (Fletcher, 1953).

TABLE 4 : Comparison of the Speech Spectra of Males Vs. Children

Frequencies	Males		Children		't' values	Significance
	Mean	S.D.	Mean	S.D.		
250 Hz	72.34	3.52	66.86	6.05	3.031	.005 **
500 Hz	79.28	2.62	77.50	5.04	1.216	.234
750 Hz	70.68	3.20	74.03	2.65	3.116	.004**
1000 Hz	60.78	4.07	67.08	3.09	4.772	.000**
1500 Hz	63.21	3.67	63.87	2.58	0.492	.851
2000 Hz	59.03	3.14	61.96	3.24	1.523	.345
3000 Hz	51.71	5.40	57.09	2.75	3.398	.002**
4000 Hz	49.10	2.92	51.33	3.35	2.042	.048*
6000 Hz	45.85	3.36	49.12	2.26	3.763	.004**
8000 Hz	41.57	3.42	46.09	3.14	3.767	.001**

* P < 0.05   ** P<0.01

Comparison of the LTASS of males, females and children

In general it can be observed that in Malayalam (Graph 1), the talkers produced very similar average spectra, characterized by regular peaks occuring at multiples of fundamental frequency which decreased with increasing frequencies.

In all the three groups most of the energy was contained in the region around 500 Hz. This is the region of the lowest frequency resonance of the vocal tract (also known as the first vocal tract formant). Above and below 500 Hz the intensity measured in 1/3 octave bands fall on the average by 6 dB per octave. These findings are in consensus with those obtained by Srinivas, 1987; Pearson, Bennet and Fidel, 1942, (cited in Studebaker and Hockbert, 1993). The mean difference of energy across the frequency range was 35.19 dB for all the three groups. With an increase in frequency, there was a decrease in the intensity of the signal. The SD was also not markedly different for the three groups. It tended to be more in children in the lower frequency. Studebaker and Hockberg (1993) reported that at constant effort and in stressed syllables, the more intense vowels such as |x| and |u| are about 30 dB more than the weak consonants such as |f| |. In running speech, where there are also unstressed syllables, this could increase the dynamic range to 40 dB and account to some extend the general decrease in the intensity with rising frequency (Levitt, 1978; Kitzing, 1986).

II Comparison of the unified LTASS obtained in Malayalam with the spectra obtained in earlier studies

The mean values of the unified LTASS for Malayalam (i.e., a combination of adult males, adult females and children LTASS) was compared with the spectra reported by Byrne et al. (1994) and Cox and Moore, (1988). As the SD of the studies done earlier were not available, a statistical comparison between their findings and that of the present study could not be carried out. The spectrum obtained in Malayalam was placed alongside in a graph with the spectra obtained by Byrne et al. (1994) and Cox and Moore (1988) so that a visual inspection could be made of the variations in the energy concentration across the frequency range (Graph 2, Table 5).

Graph - 2 : Comparison of the unified speech spectrum in Malayalam with Earlier Studies

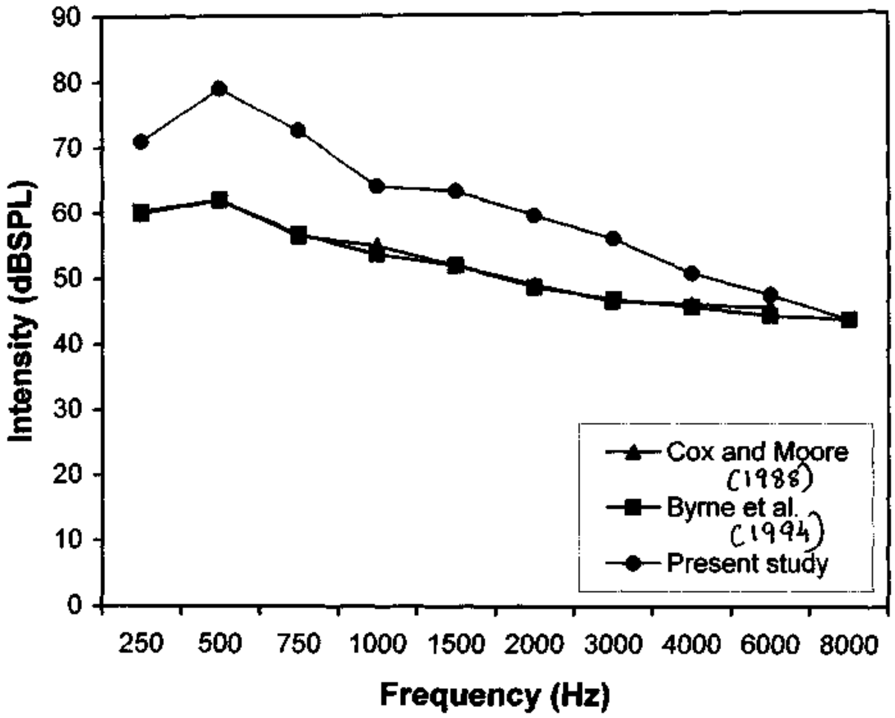


TABLE 5 : Comparison of the Unified Speech Spectrum in Malayalam with Earlier Studies

Frequencies	Cox and Moore (1988)	Byrne et al. (1994)	Present study	
			Mean	S.D.
250 Hz	60.0	60.3	71.10	5.59
500 Hz	62.0	62.1	79.07	3.71
750 Hz	56.5	56.8	72.74	3.03
1000 Hz	55.0	53.7	64.06	4.47
1500 Hz	52.0	52.0	63.32	4.23
2000 Hz	49.0	48.7	59.56	4.16
3000 Hz	46.5	46.8	55.93	4.46
4000 Hz	46.0	45.6	50.67	3.56
6000 Hz	45.5	44.3	47.42	4.31
8000 Hz	-	43.7	43.78	3.66

It can be seen that in the spectrum obtained in Malayalam, the mean intensity was highest at 500 Hz with it being 79 dB SPL while corresponding levels in the mentioned studies was 62 dB SPL. At the frequencies 250 Hz, 1 kHz, 1.5 kHz, 2 kHz and 3 kHz, the variation in the intensity was lesser. On the average, the difference between the LTASS of Malayalam and that of the studies mentioned in the literature was 9 dB, in these frequency. At 4 kHz and above, this difference was lesser. However, in general in Malayalam and in the earlier studies, the energy drops off gradually towards the higher frequencies.

There are several possible reasons as to why these variations between the LTASS of Malayalam and those reported in the literature occurred. They include:

(a) Language and phonemic content

There is clear evidence to suggest that the frequency related functions can be altered by the phonemic content of speech (Fletcher, 1953; Lehiste and Peterson, 1959, Miller and Nicely, 1955 cited in Studebaker and Hockberg, 1993). When comparing Malayalam language with English (Cox and Moore, 1987) and other languages like Spanish, Russian, Mandarin, French (Byrne et al. 1994), the energy concentration could be more in the lower frequencies owing to the greater usage of nasals, long vowels and voiced stops in spoken Malayalam.

(b) Context

Literature suggests that an increase in context causes the frequency related functions to shift downward in frequency and broaden in shape (Hirsh, Reynolds and Joseph, 1954, cited in Studebaker and Hockberg, 1985). The speech material used in Malayalam was a passage of 2 min. 20 sec while in the studies done by Cox and Moore (1987) and Byrne et al. (1994) was only about 2 min. Though it is not much of a difference in terms of time, it may have contributed to some of the increase in energy in the lower frequency in Malayalam.

(c) Vocal effort

There is a general tendency for individuals speaking Malayalam to speak rather loudly. This could be another reason as why the LTASS of Malayalm had greater energy than that of the studies reported in the literature.

(d) Methodological Variations

The passage chosen for the studies done by Byrne et al. (1994) and Cox and Moore (1988) were on the basis of being relatively easy to read. They have mentioned whether all the phonemes of the language are being represented. However for the present study material was taken based on the frequency of occurrence of the phonemes in the language as well as ease of reading.

Hence, the variations seen in the present study and that reported in the literature, probably are due to differences in the vocal efforts of the speakers, methodological variations as well as the concentration of low frequency speech sounds in the language.

Based on the findings of the present study, when the LTASS of Malayalam is applied in hearing aid prescription overall less gain should be provided, particularly at the frequencies 500 Hz and 750 Hz.

From the findings of the present study it can be concluded that:

a. The spectra of males and females are almost identical except at frequencies above 2 kHz where females have more energy concentration.
b. When the spectra of adults are compared to that of children except at 250 Hz and 500 Hz where adults have more energy and 750 Hz, 1000 Hz where children have more energy the spectra are similar.

c. When compared to the previous investigations, there is more energy concentration at lower frequencies especially 500 Hz in Malayalam, which could be attributed either to the language and phonemic content or the differences in the vocal effort.

SUMMARY AND CONCLUSION

The present study was carried out to obtain the long term average speech spectrum of adult males, adult females and children. This was done to check the following:

1. Whether a single spectrum can be used for the groups.
2. Whether there is any significant difference between the spectrum obtained in Malayalam and that computed in the earlier studies.

The subjects taken up for the study were 30 adults (15 males and 15 females) in the age range of 18-40 years (mean 28 years) and 15 children in the age range of 8-10 years (mean 9 years). The speech material of about 2 min 20 sec taken for the study was such that it represented all the phonemes in the frequency of occurrence in the language. The task involved reading the material at a comfortable pitch and loudness.

The recording was done such that the noise levels were well below that of the speech levels. The long term average speech spectrum was computed using the Audiolab software. A graph was obtained with the frequency across the x-axis and intensity across y-axis. The intensity levels at octave and mid-octave frequencies across the range of 250-8 kHz were noted separately for adult males, females and children. A unified speech spectrum including all the three groups was also computed and compared with the spectra obtained in earlier studies. The results indicated a small but significant difference between the mean of adult males and females with females having more energy above 3 kHz. While comparing adults and children it was observed that adults had more energy concentration at the lower frequencies corresponding to their fundamental frequency.

In general, the speech power in all the three groups was greater between 100-600 Hz where the energy of fundamental frequency of voice and the first formant overlaps. The energy decreases with increasing frequency such that the difference between the peaks at lower frequency and at 8 kHz is approximately 35 - 40 dB. The similarity of the LTASS across the groups demonstrate that it is reasonable to propose a unified LTASS for males, females and children. Thus a single LTASS in Malayalam can be used for various applications including hearing aid prescription and calculation of articulation index.

While comparing the unified LTASS in Malayalam with that obtained in earlier studies (Cox and Moore, 1988; Byrne, et al. 1994) it can be seen that the energy concentration in the LTASS obtained in Malayalam language is much more than the earlier studies at lower frequencies especially 500 Hz. Hence less amount of gain needs to be provided in this frequency band, when fitting hearing aids.

Implication

1. Speech spectrum plotted on an audiogram format would be clinically useful when discussing hearing impairment with patients, their families and other professional workers (Olsen, Hawkins & Van Tasell, 1987).
2. The speech spectrum would be useful in hearing aid selection and gain prescription, so that adjustments can be made to shape the amplified spectrum and the speech is maintained within comfortable listening region (Tyler, 1979; Sanders, 1982).
3. It would indirectly help in the calculation of articulation index which would contribute to increase in speech intelligibility (Pavlovic, 1989).

Suggestions for Future Research

1. Similar long term average spectrum could be derived in the other Indian languages also.
2. A comparison of the speech spectra in all the Indian languages could be carried out to see whether one unified speech spectra could be used to represent the average speech levels for calculating hearing aid gain and articulation index.
3. A number of other parameters could also be studied including the formant peaks and the spectral tilt.

REFERENCES

Angello, J. & Wagner, W. (1974). Speech spectra and perception in the design of hearing aids. In K. Donelly (Ed.), *Interpreting hearing aid technology* (pp. 85-106). Springfield, Illinois: C.C. Thomas.

Angew, J. (1999). Variations in speech signals processed by hearing aids. *The hearing Journal,* 52(11), 68-70.

Boothroyd, A. & Medwetsky, L. (1992). Spectral distribution of / s/ and the frequency response of hearing aids. *Ear and Hearing,* 13,150-157.

Brandt, J., Ruder, K., & Shipp, T. (1969). Vocal loudness and effort in continous speech. *Journal of Acoustical Society of America,* 46,1543-1548.

Byrne, D. (1977). The Speech spectrum. Some aspects of its significance for hearing aid selection and evaluation. *British Journal of Audiology,* 11, 40-46.

Byrne, D., & Dillon, H. (1986). The National Acoustics Laboratories (NAC) new procedure for selecting the gain and frequency response of a hearing aid. *Ear and Hearing,* 7, 257-265.

Byrne, D., & Fifiled, D. (1974). Evaluation of Hearing aid fittings for infants. *British Journal of Audiology,* 8, 47-54.

Byrne, D., & Tonisson, W. (1976). Selecting the gain of hearing aids for persons with sensorineural impairments. *Scandinavian Audiology,* 5, 51-59.

Byrne, D., Dillon, H., Tran, K., Arlinger, S., Wilbraham, K., Cox, R., Hagerman, B., Hetu, R., Kei, J., Lui, C, Kiessling, J., Kotby, M.N., Nasser, N.H.A., Kholy, W.A.H., Nakanishi, Y., Oyer, H., Powell, R., Stephens, D., Meredith, R., Sirimanna, T., Tavartkiladze, G., Frolenkov, G.I., Westerman, S., Ludivgsen, C. (1994). An international comparison of long term average speech spectra. *Journal of the Acoustical Society of America,* 94(4), 2108-2120.

Cleveland, T.F., Sundberg, J., & Stone, H.E. (2001). Long term average spectrum characteristics of country singers during speech and singing. *Journal of Voice,* 15, 54-60.

Comelisse, L.E., Gagne, J.P., & Seewald, R.C. (1991). Ear level recordings of the long term average spectrum of speech. *Ear and Hearing,* 12, 47-54.

Cox, R.M. (1983). Using ULCL measures to find frequency / gain and SSPL90. *Hearing Instruments,* 34,17-22.

Cox, R.M., & Moore, J.N. (1988). Composite speech spectrum for hearing aid gain prescription. *Journal of Speech and Hearing Research,* 31, 102-107.

Cox, R.M., Matesich, J.S., & Moore, J.N. (1988). Distribution of short-term rms levels in conversational speech. *Journal of the Acoustical Society of America.* 84(3), 1100-1104.

Davis, J.M., & Hardick, E.J. (1981). *Rehabilitative Audiology for children and adults.* John Wiley and sons.

Dillon, H. (1988). Comparison in hearing aids. In R.E. Sandlin (Ed.), *Handbook of hearing aid amplification,* Volume 1 (p. 121-145). Boston, MA: College Hill Press.

Doherty, T. (1976). An evaluation of selected acoustic parameters for use in speaker identification. *Journal of Phonetics,,* 4, 321-326.

Doherty, T., & Hollien, H. (1978). Multiple factor speaker identification of normal and distorted speech, *Journal of Phonetics,* 6, 1-8.

Flanagan, J. (1958). Some properties of the glottal sound source. *Journal of Speech and Hearing Research,* 1, 99-116.

Fletcher, S.G. (1953). Acoustic Phonetics. In K. Berg, S.G. Fletcher, (Eds.). *The hard of hearing child* (pp 57-84). New York: Grune and Stratton.

Formby, C, & Monsen, R.B. (1982). Long term average speech spectra for normal and hearing impaired adolescents. *Journal of the Acoustical Society of America,* 71(1), 196-202.

French, N.R., & Steinberg, G.C. (1947). Factors Governing the intelligibility of speech sounds. In M.E. Howley (Ed.). *Speech Intelligibility and Speaker Recognition* (pp 128-152). Strovdsberg, Pennsylvania: Dowden, Hutchinson & Ross.

Goberman, A.M., & Robb, M.P. (1999). Acoustic Examination of preterm and full term cries. The long term average spectrum. *Journal of Speech,. Language and Hearing Research,* 42, 850-861.

Hodgson, W.R. (1977). Speech Acoustic and Intelligibility. In W.R. Hodgson & P.H. Skinner (Eds.). *Hearing Aid Assessment and use in Audiologic Habilitation,* (pp 115-119). Baltimore: Williams & Wilkins.

Hodgson, W.R. (1988). Discrimination of Acoustic Signals. In R.E. Sandlin (Ed.), *Handbook of Hearing Aid Amplification, Volume I: Theoretical and Technical Considerations,* (pp 290-295). San Diego, California: Singular Publishing Group.

Hollien, H., & Majeswki, W. (1977). Speaker identification by long-term spectra under normal and distorted speech conditions. *Journal of the Acoustical Society of America,* 62, 975-980.

Holte, L, & Margolis, R.H. (1987). The relative loudness of third-octave bands of speech. *Journal of the Acoustical Society of America,* 81(1), 186-190.

Kent, R.D. (1993). *The Speech Sciences.* The University of Wisconsin - Madison. San Diego: Singular Publishing Group.

Kent, R.D., & Read, W.C. (1992). *The acoustic analysis of speech.* San Diego: Singular Publishing Group.

Kitzing, P. (1986). LTAS criteria pertinent to the measurement of voice quality. *Journal of Phonetics.* 14, 477-482.

Kiukaanniemi, H., & Mattila, P. (1980). Long term speech spectra. A computerized method of measurement and a comparative study of Finnish and English. *Scandinavian Audiology,* 9, 67-72.

Kiukaanniemi, H., Soponen, P., & Mattila, P. (1982). Individual differences in the long - term speech spectrum. *Folia Phoniatrica,* 34, 21-28.

Lau, P.H.L. and Wong, L.L.N. (2000). The effect of type and level of noise on LTAS. Available : http://www.google.com.

Levitt, H. (1978). The Acoustic of speech production. In M. Ross and T.G. Giolos (Eds.). *Auditory Management of Hearing impaired children* (pp 45-48). Baltimore: University Park Press.

Li, K.P., & Hughes, G.W. (1974). Talker differences as they appear in correlation matrices of continous speech spectra. *Journal of the Acoustical Society of America,* 5S, 833-837.

Linville, S.E., & Rens, J. (2001). Vocal Tract Resonance Analysis of Aging voice using long term average spectra. *Journal of Voice,* 15(3), 323-330.

Lofquist, A., & Manderson, B. (1987). Long term average spectrum of speech voice analysis. *Folia Phoniatrica,* 39, 221-229.

Me Candless, P.T. (1985). Principles and clinical utility of hearing aid fitting formulas. In R.E. Sandlin (Ed.), *Text book of Hearing Aid Amplification Technical & clinical consideration,* San Diego, California: Singular publishing group.

Mc Candless, G.A., & Lyregaard, P. (1983). Prescription of gain / output (POGO) for hearing aids. *Hearing Instruments,* 34,16-21.

Mendoza, E., Munoz, J., & Narango, N.V. (1996). The long-term average spectrum as a measure of voice stability. *Folia Phoniatrica,* 48, 57-64.

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

Mrinal, J. (1998) Long-term average speech spectrum in Hindi and its implication in hearing aid selection and evaluation. An Unpublished Master's Dissertation, University of Mysore, Mysore.

Niemoeller, A., Mc Cormick, L, & Miller, J. (1974). On the spectrum of spoken English. *Journal of the Acoustical Society of America,* 55, 461.

Nittrouers, S., Mc Gowan, R.S., Milenkovic, P.A., & Bechler, D.A. (1990). Acoustic measurements of men & women's voices. A study context effects and covariation. *Journal of Speech and Hearing Research,* 33, 761-775.

Olsen, W.O., Hawkins, D.B., & Van Tasell, D.J. (1987). Representations of the long term spectra of speech. *Ear and Hearing* (Suppl.), 8, 100-108.

Owens, E. (1961). Intelligibility of words, Varying in familiarity. *Journal of Speech and Hearing Research,* 4, 113.

Pascoe, DP. (1978). An approach to hearing aid selection. *Hearing Instruments,* 29, 12-16, 36.

Pavlovic, C.V. (1984). Derivation of primary parameters and procedures for use in speech intelligibility predictions. *Journal of the Acoustical Society of America,* 82, 413-422.

Pavlovic, C.V. (1989). Speech spectrum considerations and speech intelligibility predictions in Hearing aid evaluations. *Journal of Speech and Hearing Disorders,* 54, 3-8.

Pavlovic, C.V., & Studebaker, G.A. (1984). An evaluation of some assumptions underlying the articulation index. *Journal of the Acoustical Society of America,* 75,1606-1612.

Pavlovic, C.V., Studebaker, G.A. & Sherbecoe, R.L. (1985). "An articulation index procedure for predicting the speech recognition performance a hearing-impaired individuals". *Journal of the Acoustical Society of America,* 80, 50-57.

Pickett, J.M. (1980) *The sounds of speech communication.* Baltimore : University Park Press.

Popelka, G.R., & Engebretson, A. M. (1983). A computer based system for hearing aid assessment. *Hearing Instruments,.* 7,6-7, 9, 44.

Pradhan, B. (1994). Acoustic characteristics of Nepali Language (Long term average speech spectrum). An Unpublished Master's Independent Project, University of Mysore, Mysore.

Ramakrishna ,B.S., Nair, K.K., Chiplunkar, V.N., Atal, B.S., Ramachandran, V., Subramanian, R. (1962). *Some aspects of the relative efficiencies of*
^ *Indian languages.* Ranchi: The Catholic Press.

Ross, M. (1975). Hearing Aid Selection for the Preverbal Hearing impaired child. In M.C. Pollack, (Ed.). *Amplification of the hearing impaired* (pp. 223-227). New York: Grune & Stratton.

Sadaoki-Furvi (1986). On the role of spectral transition for speech perception. *Journal of the Acoustical Society of America,* 80,1016.

*Rashmi, M(1985). Acoustic Aspects of the speech of children
An unpublished master's dissertation project,
university of Mysore , Mysore.

Sanders, D.A. (1982). *Aural Rehabilitation* (A management model), Englewood cliffs, New Jersy: Prentice-Hall.

Schwartz, M. (1970). Power spectral density measurements of oral and whispered. *Journal of Speech & Hearing Disorders,* 13, 445-446.

Seewald, R.C., Ross, M. & Shapiro, M.K. (1985). Selecting amplification characteristics for young hearing impaired children. *Ear and Hearing.* 6, 48-53.

Shapiro, I. (1976). Hearing aid fitting by prescription. *Audiology,* 15, 163-173.

Skinner, M., & Miller, J. (1983). Amplification bandwidth and intelligibility of speech in quiet and in noise for listeners with sensorineural hearing loss. *Audiology,* 22, 253-279.

Skinner, M.W. (1984). Effects of Frequency Response, Bandwidth and overall gain of linear amplification systems on performance of adults with sensorineural hearing loss. In G.A. Studebaker & I. Hochberg (Eds.). *Acoustical factors affecting hearing and performance* (pp. 135-149). Massachusets: Allyn and Bacon.

Skinner, M.W. (1988). *Hearing aid Evaluation.* Englewood Cliffs. New Jersey: Prentice Hall.

Srinivas, N.C. (1987). The power density spectrum of Telugu speech. *The Journal of AIISH,* 18,74.

Stelamachowicz, G., Mace, L.A. , & Kopur, J & Carney (1993). Long term and short term characteristics of speech. Implications for hearing aid selection for young children. *Journal of Speech and Hearing Research,* 36, 609-620.

Studebaker, G.A. & Hochberg, I. (1985). *Acoustic factors affecting hearing aid performance.* Massachusets: Allyn and Bacon.

Studebaker, G.A. (1985). Directivity of the human vocal source in the horizontal plane. *Ear and Hearing,* 6, 315-319.

Studebaker, G.A., & Sherbecoe, R.L (1985). Frequency importance functions for speech. In G.A. Studebaker & I. Hochberg (Eds.). *Acoustical factors affecting hearing aid performance* (pp 183-189). Massachusets: Allyn and Bacon.

Tyler, R. (1979). Measuring hearing loss in the future. *British Journal of Audiology, (suppl 2),* 29-40.

Wendler, J., Doherty, E., & Hollian, H. (1980). Voice classification by means of long term spectra. *Folia Phoniatrica,* 32, 51-60.

Yanagihara, N. (1967). Significance of harmonic changes and noise components in hoarseness. *Journal of Speech & Hearing Research,* 10, 531-541.

Zeng, F.G., & Turner, C.W. (1990). Recognition of voiceless fricatives by normal and hearing impaired subjects. *Journal of Speech and Hearing Research,* 33, 440-449.

APPENDIX

ഗാന്ധിജി

ജനിക്കുമ്പോൾ ആരും മഹാൻമാരല്ല. സ്വാർഥ ചിന്തയില്ലാതെ മനുഷ്യരാശിക്കു നൻമ ചെയ്തുന്നവരാണു മഹാൻമാരായിത്തീരുന്നത്. അത്തരത്തിൽപ്പെട്ട അനേകം പേർ ലോകത്തിലുണ്ടായിട്ടുണ്ട്. ഇന്നും ഉണ്ടായിക്കൊണ്ടിരിക്കുന്നുമുണ്ട്. മരിച്ചാലും അവരുടെ മഹത്വം മങ്ങുന്നില്ല. ജനഹൃദയങ്ങളിൽ എക്കാലത്തും അവർ ജീവിക്കുന്നു. സർദാർ പട്ടേൽ, നേതാജി സുഭാഷ് ചന്ദ്രബോസ്, ലോകമാന്യതിലകൻ, രവീന്ദ്രനാഥടാഗോർ, മഹാത്മഗാന്ധി തുടങ്ങിയ ഭാരതീയ മഹാന്മാരും ഝാൻസിറാണി, സരോജിനിനായിഡു തുടങ്ങിയ ധീരവനിതകളും അക്കൂട്ടത്തിൽപ്പെടുന്നു. അവരിൽ ഏറ്റവും പ്രധാനിയാണ് മഹാത്മാഗാന്ധി. ഭാരതീയരായ നാം ഗാന്ധിജിയെ നമ്മുടെ രാഷ്ട്രപിതാവായി ആദരിക്കുന്നു. 1869 ഒക്ടോബർ 2 ന ഗുജറാത്തിലെ പോർബന്തർ എന്ന ഗ്രാമത്തിലാണ് അദ്ദേഹം പിറന്നത്. മോഹൻദാസ് കരം ചന്ദ് ഗാന്ധി എന്നായിരുന്നു അദ്ദേഹത്തിന്റെ മുഴുവൻ

പേര്. കുട്ടികൾ അദ്ദേഹത്തെ 'ബാപ്പുജി' എന്നാണ് വിളിച്ചിരുന്നത്.

സ്വന്തം ഗ്രാമത്തിലെ സ്കൂളിൽ നിന്ന് മെട്രിക്കുലേഷൻ പരീക്ഷ ജയിച്ചു. പിന്നീട് ഇംഗ്ലണ്ടിൽ പഠിച്ചാണ് ബാരിസ്റ്റർ ബിരുദം നേടിയത്. തെക്കെ ആഫ്രിക്കയിൽ കുറെനാൾ വക്കീൽപ്പണിയിൽ ഏർപ്പെട്ടു. അവിടുത്തെ ഇന്ത്യക്കാരുടെയും കറുത്തവർഗക്കാരുടെയും സ്വാതന്ത്ര്യത്തിനും. ഐക്യത്തിനും വേണ്ടി പ്രവർത്തിച്ചു. അങ്ങനെ അവർക്ക് ഒരു ഛായയായി തീർന്നു. ഒരു വിദ്യാലയവും അവിടെ അദ്ദേഹം നടത്തി വന്നിരുന്നു.

1915-ൽ ഗാന്ധിജി ഇന്ത്യയിലേക്ക് തിരിച്ചു വന്നു. പാവപ്പെട്ട കൃഷിവലൻമാരുടെ ഇടയിലാണ് അദ്ദേഹം പ്രവർത്തനം ആരംഭിച്ചത്. കൂടാതെ 'ഖാദി' ഉൽപാദനത്തിലൂടെ അനേകം പേർക്ക് ജോലി നൽകി. നിസ്വാർഥനും സത്യസന്ധനും അഹിംസയിൽ വിശ്വസിച്ചിരുന്നവനുമായ അദ്ദേഹത്തെ ജനങ്ങൾക്ക് ഇഷ്ടമായി. അവർ അദ്ദേഹത്തിന്റെ അനുയായികളായി. ബീഹാറിലെ ചമ്പാരൻ എന്ന സ്ഥലത്ത് തോട്ടം തൊഴിലാളി

കളെ ചൂഷണം ചെയ്തുകൊണ്ടിരുന്ന വെള്ളക്കാർക്കെതിരായി അദ്ദേഹം സമരം സംഘടിപ്പിച്ചു. പിന്നീടുള്ള ജീവിതകാലം മുഴുവനും സമരബഹുലമായിരുന്നു.

അക്രമത്തെ അക്രമം കൊണ്ടു നേരിടുകയല്ല ഗാന്ധിജി ചെയ്തത്. ക്രിസ്തുവിനെപ്പോലെ ക്ഷമിക്കാൻ അദ്ദേഹത്തിനു കഴിഞ്ഞു. അഹിംസയായിരുന്നു അദ്ദേഹത്തിന്റെ സമരമുറ. സത്യാഗ്രഹമായിരുന്നു അതിനുപയോഗിച്ച മാർഗം. സത്യമാണ് ഈശ്വരൻ എന്ന ദൃഢതയോടെ വിശ്വസിച്ച മഹാത്മാവാണ് ഗാന്ധിജി.

ഈശ്വരന്റെ മുമ്പിൽ എല്ലാവരും സമന്മാരാണ് ഇന്ത്യയിലെ അധഃകൃത സമുദായങ്ങളെ അദ്ദേഹം ഹരിജനങ്ങൾ (ഈശ്വരന്റെ മക്കൾ) എന്നു വിളിച്ചത് അതുകൊണ്ടാണ്. 'ഹരിജൻ' എന്ന പേരിൽ ഒരു പത്രവും അദ്ദേഹം നടത്തിവന്നിരുന്നു.

ഒരു തുള്ളി ചോര പോലും ചിന്താതെ സമാധാനപരമായ മാർഗത്തിലൂടെ അദ്ദേഹം വെള്ളക്കാരെ ഈ നാട്ടിൽ നിന്നു കെട്ടുകെട്ടിച്ചു. അങ്ങനെയാണ് 1947 ആഗസ്റ്റ് 15-നു ഭാരതം

Malayalam Passage (IPA Transcription)

dʒənikumbʰol a:ɾum mahanmaɾalla. Swaɾt̪ʰa tʃindaillade manuʃəɾa:ʃikʰə nanma tʃeyunnavaɾannu mahanmaɾ a:yi t̪i:ɹunad̪ə. at̪ʰaɾat̪ʰilpetta ane:kam pe:ɾ lokat̪ʰiluṇḍʰayiḍʰuṇḍə. innum uṇḍayikʰondɪrikunnumuṇḍə. maɾitʃa:lum aveɾude mahat̪vəm maŋunila. dʒnahɾədayaŋlil ekkalat̪ʰum avaɾ dʒi:vikʰun̪n̪u. Sardaɾ patte:l, neda:dʒi ʃubʰa:ʃ tʃəŋdrabos, lokamanya tikalan, ɾavi:ŋ dɾna:t̪ t̪agəɾ, mahat̪ma gand̪ʰi. t̪ʰuḍaŋiyə bʰaɾadi:yə mahanma:ɾum, jʰansi: ɾa:ni, Saɾo:dʒini na:yd̪u tuḍaŋi:yə dʰi:ɾə vanid̪akalum a:ku: tatʰil pedun̪n̪u. avaɾil e:t̪avum pɾadʰa:ni ya:nnə mahat̪ma gandʰi. βʰaɾat̪i:yaɾa:yə na:m gandʰidʒi:ye nammude ɾaʃtɾapit̪a:v a:yi a:d̪ʰəɾikʰunnu.

1869 ɔktɔ:bəɾ ɾandinə gudʒuɾa:-tʰile Poɾband̪eɾ enne gɾamatila:nə ad̪eham piɾan̪əd̪ʰə. mohanda:s kaɾamtʃənd gand̪ʰi ena:yiɾunnu adehat̪inde muɾuvan pe:ɾə. kuṭikəl ad̪ehat̪ʰe ba:pudʒi enna:nə vili-tʃiɾunnad̪ə.

സ്വതന്ത്രയായത്.

1948-ൽ ഒരു പ്രാർത്ഥനാ യോഗത്തിൽ പങ്കെടുത്തു കൊണ്ടിരിക്കെ, ആ മഹാത്മാവ് വെടിയേറ്റു മരിച്ചു. ലോകം കണ്ട ഏറ്റവും നല്ല മനുഷ്യരിൽ ഒരാളാണ് ഗാന്ധിജി.

Swantʰam gramatʰile Sku:ḷil ninne metriku-
leʃən pari:kʃa dʒəi:tʃə pin:de iŋgəlandil
patitʃa:nnə bo:risṭṭər bi:rudʰam ne:diyadʰə.
t̪ekʰe a:frikʰayil ku:reno:l vakʰi:lpaṇi:yil
ye:rpaṭṭu. avidet̪ʰe indyakʰa:ruḍeyum, karutə
vargəkʰa:ruḍeyum Swa:dəntrət̪i:num aykyat̪inum
pravart̪ʰitʃu. aŋŋne avarkʰə oru tʃa:yeyayi
t̪i:rənnu. Oru vid̪yaleyavum avide adʰeham
hadətʰi vanirunnu.

1915-nil gʰa:ŋdʰi:dʒi indyayile:kʰə
t̪iritʃu vannu. pa:vappeṭə kriʃivalənma:ruḍe
i:ḍeyila:nnə ad̪ʰeham pravart̪ʰanam a:rambʰitʃəo
ku:ḍad̪e "kʰa:d̪i" ulpa:d̪anət̪ʰilu:ḍe ane:kam
pe:rkʰə dʒə:li nalgi.

niswart̪ʰenum satyasand̪enum ahimsayil
viʃvasitʃirunnavanuma:yə ad̪ehəte dʒenaŋḷkʰə
iʃtəma:yi. avər ad̪ʰehat̪inḍe anuyayikaḷa:yi.
bi:harile tʃəmba:rən enne st̪alet̪u t̪ətem
t̪ət̪ila:ḷikele tʃu:ʃənam tʃəyd̪u kənḍiruna
velakʰa:rəkʰed̪ira:yi ad̪ʰeham samarem

səŋgʰəḍipitʃu. piniḍule dʒi:viḍəka:lam
muḻuvanum samarabʰehuləma:yirunnu.

akramətʰe akramam kəṇḍu
ne:riḍukeyaḷa gandʰidʒi tʃəyḍade. kristuvine
po:le kʃəmikan adehaṭinu kaḻiŋ . ahimse-
-yayirunnu adehaṭinde samaramurạ. satya:-
-grahəma:yirunnu aḍinupeyəgitʃe ma:rgam.
saṭyama:nne i:ʃəvaran ennu aṛiḍəḍeyəḍe
viʃvəsitʃə mahatəmava:nnə gandʰidʒi.

i:ʃvarənde munbʰil ela:varum
samanma:ra:nne. inḍyeyile adʰəmkriṭə samu-
-dʰayaŋle aḍʰeham haridʒanaŋḷ enne vilitʃəḍə
aḍukəṇḍa:nə. haridʒən enne pe:ril oru
patravum aḍʰeham naḍetʰivannirunnu.

ɔru ṭʰuḷi: tʃərapɔ:lum tʃiŋḍʰaḍe
sama:ḍa:nəparama:yə margetʰilu:de aḍʰeham
veḷḷəka:re e: na:ṭil ninnə keṭṭukeṭṭitʃu.
aŋenneya:nnə 1947 a:gastu 15. dinu
bʰa:radam swatəntrəyayəḍə.

1947 nil oru pra:rtənna yəgəṭi

paŋgeduṯukʰɔnḍiɹikæ aː mahatmaːvə.

vedɫyeːṯṯu maɹitʃu. lɔːgam kandə eːṯəvum

nallə manuʃəɹil ɔraːlaːnne ganḏʰidʒi.