Indian J. Phys. **47**, 598–613 (1973)

Some studies on acoustic features of human speech in relation to Hindi speech sounds

DWIJESH DUTTA MAJUMDER, ASOKE KUMAR DUTTA AND NIHAR RANJAN GANGULI

Electronics and Communication Sciences Laboratory
Research & Training School, Indian Statistical Institute, Calcutta-35

(*Received* 16 *June* 1972)

(Plates— 3 – 9)

The present paper aims to present a brief discourse on acoustic features related to human speech from the view point of analysis. After giving a brief review of modern methods currently in use in speech communication research, acoustic characteristics and features of human speech sound on the basis of spectrographic analysis of a limited number of Hindi Speech Sound are presented. The acoustic phonetic and the acoustic prosodic parameters of human speech are briefly explained and the formant frequencies, and duration of Hindi vowels, concentration of acoustic energy for plosives and some affricates along with other related parameters of consonants are presented in tabular form and discussed.

1. INTRODUCTION

1.1. One of the most common mode of communication between human beings is speech. The intelligence transferred through speech is largely contained in the sound waves and therefore the extraction of this intelligence should theoretically be possible from these and the associated phenomena. But because of the highly complex character of the speech waves, inadequacy of the knowledge of how the intelligence is imbedded in the acoustic and other parameters and the statistical variations associated with biological processes involved, this apparently simple problem has engaged the attention of researchers for almost a century. The attempt to analyse speech sounds dates back to Helmholtz (1877) and the present state of knowledge in this field may be ascribed to the work of scientists like Stumf (1926), Paget (1930), Barczinski & Thienhaus (1935), Chiba & Kajiyama (1941) Smith (1947), Tarnoczy (1948), Halle *et al* (1952), Fant (1952), Cooper *et al* (1952), Peterson (1951), Jackobson & Halle (1956) and scores of others.

Without going into the historical perspective of the subject, a brief review of modern methods in speech communication research may be helpful before we discuss acoustic-phonetic features of human speech sound. One of the purposes of scientific analysis of a language is to discover those features which are the physical manifestations of the functionally distinct phonological entities within the language under investigation. Different languages may be using different features in making phonemic distinctions. In recent years investigations into the relationship between the physical aspects of speech and the functional aspects of speech have been greatly aided by the development of highly sophisticated research tools, such as the sound spectrograph, the pitch extractors and speech synthesizers for analysis of articulatory dynamics at the coustic level, and development of articulatory models (Dunn 1950, Stevens, Kasowski & Fant 1935, 1960) and cineradiographic techniques (Moll 1960, Declerk 1965, Subtelny 1967) for investigating the articulatory dynamics and the generative processes at the physiological level.

Though the studies at the physiological level are beyond the scope of the present paper, it should be emphasized that investigation on *articulatory models* of speech mechanism and vocal mechanism using *electrical analog* of the vocal tract, where the tract is idealized by a sufficiently large number of cylindrical tube sections in cascade, though not yet comprehensive enough has given much insight into the problems of speech production.

1.2 *Cineradiography*

Lateral cineradiography is a technique by which a sequence of lateral X-ray photographs of the articulatory mechanisms such as lip positions, labial movement, dynamics of tongue positions and activity of the velum can be taken 500 frames per second, though for a very insufficient amount of time (a few minutes only), providing some data on the neuromuscular system which seem to be responsible for speech production responding to a sequence of instructions.

At the acoustic level the analytical methods employed before the last world war certainly contributed to the store of information on the acoustic cues of speech but those analyses in some cases led to erroneous conclusions The liberating impact on speech research come with the development at the Bell Laboratories in 1945 of the *Sound Spectrograph* and the subsequent development at several Laboratories of *Speech Synthesizers* that uses various means to transform spectrographic patterns to produce intelligible speech.

1.3. *Spectrographic Analysis of Speech Sound*

Spectrographic instruments (Kay Sonagraph model 7029-A) usually permits three types of display. First type of display gives an overall, three dimensional picture of the signal being analyzed; frequency, amplitude and time (Figures 4a, b). The second type of display permits the individual intensity of each frequency component to be displayed at any preselected point in time, which is usually referred to as a section (Figure 4a). The third analysis that can be performed shows the average amplitude of all frequencies present relative to time.

With this pattern, the entire input signal can be examined for flatness, or any amplitude study relative to time (Figure 4b). The sound spectrograms are in effect the automation of Fourier analysis of speech spectra in physical dimensions. It immediately made evident acoustic factors of speech that had not been suspected, and helped to verify or invalidate various aspects of the theories that analytical methods had been gradually yielding.

1.4. *Speech Synthesizing Technique*

The development of Speech Synthesizer in the early 1950's at the Haskins Laboratory opened a new way for a programmatic method of exploration. Elements of synthetic spectrograms were successively suppressed, and the patterns thus amputated were run through the synthesizer. By listening to the result, it was possible to determine, step by step, which acoustic elements formed the acoustic cues for recognition. The work soon revealed the importance of first three formants in vowel perception. In the next phase a new synthesizer called Digital Spectrum Manipulator has been developed also at the Haskins Laboratory with which it is possible to make microsurgical modifications of speech spectrograms for studying stress, intonation and relative importance of individual cues for sounds when multiple cues exist in the natural speech

It should be pointed out that these instruments were used for acoustic level of study, and were largely motivated by the problems of speech recognition which is also the motivation of the present authors (Dutta Majumder & Dutta 1969, 1968). It can be remarked that these studies have provided weighty evidence that it is not feasible to build machines for speech recognition based on the acoustic level alone, and so new methods, new instruments, new experiments and new directions would be required for the purpose.

1.5 *Articulatory Analog Technique*

Recently, the development of an Articulatory Analog of the human vocal tract by Stevens at the MIT provides a new philosophical and experimental outlook to the problem The equipment is a synthesizer incorporating the acoustic information, along with the generative features of language, and may be thought of as forming a bridge between the research at the acoustic level and linguistic level.

However, there is another research tool which in itself is a subject of tremendous importance, the computer, which promises to open new objectives of speech research. Its power as tools of analysis, synthesis or simulation, as digitisers and sorters, will lead to profound changes in experimental phonetics and application of statistical methods in speech research

The present paper, which is part of an investigation being carried out in this Laboratory, motivated by problems associated with speech pattern recognition and coding for computer use, aims at making spectrographic analysis of Indian languages, and their statistical study with the help of computers.

2.0 General terms used in acoustic–phonetics and their measurements

It is wellknown that sound energy is propagated in the form of longitudinal pressure waves which can be described in terms of amplitude, frequency, phase and a mathematical relation between them generally known as the wave-equation. The simplest type of sound wave is a simple harmonic wave (figure 1) usually known as *pure tone, e g.* the tone produced by a vibrating tuning fork. A complex periodic sound wave can be resolved into a set of pure tones with appropriate frequencies and intensities (figure 2). The method of such analysis is known as Fourier analysis, though the speech wave rarely meets with the mathematical restrictions of Fourier analysis.

Figure 1 Pure Tone

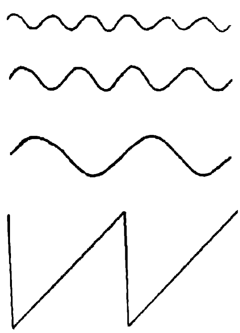

Figure 2. Saw-tooth wave form, the Fundamental and two Higher Harmonics

2 1 *Sound Intensity*

Oscillographic and spectrographic measurements of the speech wave involve the specification of *amplitude* or *intensity* measures. The term *amplitude* refers to the instantaneous or time average value of sound pressure, volume velocity, or particle velocity at a particular point in the sound field, or to the corresponding voltages or currents as delivered by a microphone.

Sound *intensity* is the energy per unit time transmitted through a unit area. In a plane or spherical free progressive sound wave the intensity in the direction of propagation is

$$I = P^2/\rho c \text{ erg sec}^{-1}/cm^2 \quad \dots (1)$$

Where P is the rms sound pressure in dynes/cm², ρ is the density of the medium in gm/cm², and c is the velocity of propagation in cm/sec. The product ρc is known as the *specific acoustical resistance* which is 41.4 dynes sec/cm³ at 20°C and 40 0 at 35°C (the latter value is appropriate for the wave propagation within the vocal cavities).

Sound pressure or intensity data are generally specified on a logarithmic scale with the unit deci Bel (dB) relative to a fixed reference The standardised sound pressure reference is $P_0 = 0\ 0002$ dynes/cm² which corresponds to a reference intensity

$$I_0 - 10^{-16} \text{ watt/cm}^2 \quad \dots (2)$$

The sound level is defined as the intensity ratio

$$L = \log_{10} \frac{I}{I_0} \text{ (Bel)} = 10 \log_{10} \frac{I}{I_0} \text{ deci Bel (dB)}$$

which is identical with sound pressure level ·

$$L = 20 \log_{10} \frac{P}{P_0} \text{ dB}, \quad \dots (3)$$

provided P_0 is related to I_0 by eqn. 1.

2.2 *Frequency, Voice Fundamental and Formants*

The number of complete oscillations executed by every particle of the medium through which the sound wave passes in one second is called the frequency of sound.

In a general periodic continuous wave the frequency of the periodic wave is called the *fundamental frequency*, and the simple harmonic waves of higher multiple frequencies are generally known as harmonics (figures 2).

The complex sound wave originating in the larynx passes through the vocal tract and its cavities before coming out in the open air. The sound wave which ultimately radiates into air has only those frequencies which are accentuated by the vocal tract and the cavities. Therefore in the plot of frequency *vs* energy, known as frequency spectrum or energy spectrum, there are peaks corresponding to the resonant frequencies, that appear in band like structure. These bands are known as formants (Figure 4a).

The basic property of a vocal chord sound source is its periodicity expressed by the duration T_0 of a complete voice period or by its inverse value, the voice fundamental frequency given by $F_0 = 1/T_0$. One often comes accross the following notations while making spectrographic studies of speech waves, which are also being used in the present paper

F_n = frequency of formant number n in c/s,

B_n = banwidth of formant nuber n in c/s,

L_n = level of formant number n in dB,

F_n = formant number n,

F_0 — frequency of the voice fundamental with or without dimension

3.0 Acoustical characteristic of speech

The acoustical characteristic of human speech is of a much more complex nature than the sound produced normally through mechanical means, such as musical instruments. In fact, human speech wave are never either fully periodic or fully noise like random phenomena. We can classify all speech waves into six basic categories namely, (1) quiescent, (2) burst, (3) quasi-random, (4) quasi-periodic, (5) double periodic and (6) irregular periodic (figure 3).

3 1 A *quiescent* speech wave is such a speech wave for which the instantaneous amplitude of the wave generated by vocal mechanism is approximately zero This type of speech wave is produced during quiet respiration or during certain closure of the vocal tract (Figure 3a).

3.2. A speech *burst* has the form of impulse and is produced by the release of a closure in the vocal tract (Figure 3b).

3 3 A *quasi-random* speech wave is produced by the vocal mechanism such that successive amplitudes at a regular interval of time are approximately uncorrelated and random (Figure 3c).

3.4 A *quasi-periodic* speech wave is produced by a vocal mechanism which is recurrent and in which the wave forms for successive periods are approximately the same (Figure 3d).

3 5 A *double periodic* sound wave is produced by a vocal mechanism which is recurrent and in which the wave forms of alternate periods are more similar than those of successive periods (Figure 3e).

3 6. An *irregular periodic* speech wave is produced by a vocal mechanism which is recurrent and in which the wave forms of successive periods vary in an irregular manner (Figure 3f).

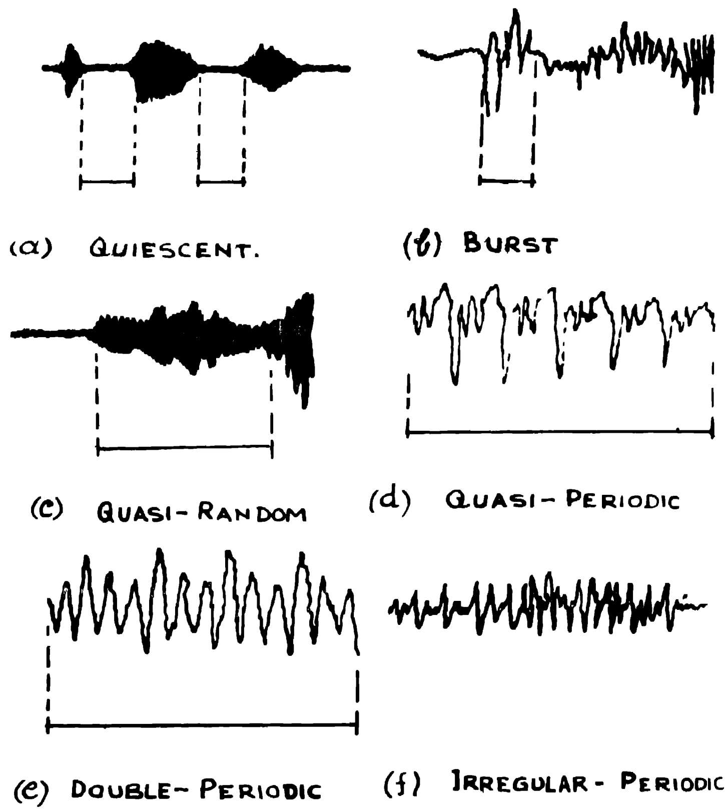

Figure 3. Basic speech waves

4.0 Acoustic parameters of human speech

The acoustical speech parameters are divided into two classes : (1) *acoustic phonetic parameters* and (2) *acoustic prosodic parameters*.Peterson & Shoup (1966) enumerated following six acoustic phonetic parameters, which we shall briefly explain along with their spectrographic illustrations from Hindi Speech Sounds

4.1. A *gap* is an interval during which the vocal mechanism does not produce sound, which is preceeded and/or followed by an interval during which the vocal mechanism does produce a sound, and during which a driving pressure is applied to a closure of vocal tract (Figure 4b, plate 4).

4.2. *The frequency of a vowel or constant formant* is the frequency of the decaying sinusoidal response of a resonance in the vocal tract (figure 4a, plate 3)

The natural range of variation of the voice fundamental frequency and of formant frequencies for non-nasal voiced sounds uttered by average male subjects according to Fant (1959) are :

F_0	60— 240	c/s
F_1	150— 850	c/s
F_2	500—2500	c/s
F_3	1500—3500	c/s
F_4	2500—4500	c/s

Females have one octabo higher fundamental pitch but 17% higher formant frequencies, children (upto 10 yrs age) have on the average 25%higher than males and $f_0 = 300$ c/s

It may be accepted that the first three formant frequencies are sufficient for the phonetic identity of a vowel. The individual formant frequencies show considerable overlapping for all the vowels. Even the first two formant frequencies taken together do not resolve the overlapping. The third formant frequency has to be considered to achieve proper resolution. Peterson (1951) suggested constant ratios for the formant frequencies whereas Fant (1959) suggested substitution of second formant frequency by an weighted average of second and third formant frequencies. However, we have presented from our spectrographic studies the formant frequencies of Hindi vowels in tabular form (table 1), which does not show much of a difference from that of Fant, and have used the method of difference for phonetic identification of vowels which show complete separation (figure 8).

4 3. *The amplitude of vowel or consonant formant* is the maximum amplitude of the decaying sinusoidal response of a resonance in the vocal tract (figure 4a, plate 3).

4.4. A *voice-bar* is a low frequency resonant response of the vocal tract that is excited by successive impulses of air emitted from true vocal folds while the voice tract is closed by one or more supra-glottal articulators (figure 5a, plate 6 & b, plate 7).

4.5. The *frequency of a consonant antiresonance* is the frequency at which a minimum occurs in the magnitude of the envelope of the energy spectrum, if the minimum is in a region where the magnitude of the envelope of the spectrum is substantially reduced relative to the magnitude that the envelope would have if the antiresonance were not present (figure 5a, b, c, plates 6, 7 & 8).

The frequency of antiresonance is an important cue for distinguishing nasal consonants.

4.6. *A broad-band continuous specrum* is a band of energy which is continuous over a relatively wide range of frequencies in the speech spectrum which has higher total energy than any other band of frequencies that occur simultaneously in the spectrum and within which the peak energy divided by the energy of any other frequency within the band does not exceed 30dB (figure 6a, b, c, plates 8 & 9). The fricative consonants are characterised by the presence of broad band continuous spectrum (because of the associated turbulance noise).

4.7. In addition to these another acoustic phonetic parameter, the *formant bandwidth* is very useful and is defined to be the range of frequency of a formant measured upto 3dB down from the peak of the envelope.

5.0 The acoustic prosodic parameters

The acoustic prosodic parameters are 1) acoustic phonetic duration, 2) average fundamental frequency and 3) average speech power as explained below.

5 1. *Acoustic Phonetic Duration* : This is the time interval of the duration of a phone plus the duration of those phonetic transitions which are associated with a phone (table 4). Usually this parameter is associated with studies of stress and intonation. However it will be shown in a later section (6.0) that this parameter may also profitably be used for the distinction between aspirated and unaspirated consonants.

5 2. *Average fundamental frequency* is the frequency of the first harmonic of a periodic speech wave which is produced by the true vocal folds and which is averaged by a weighting function over an interval of time which we have shown earlier (figure 4a, plate 3)

5.3. *Average speech power* is the product of pressure and volume velocity in a speech wave produced by the vocal mechanism averaged by a weighting function over an interval of time (figure 4b.).

Plosives can be distinguished by the sharp impulsive nature of the speech power (figure 4b, plate 4) The laterals are characterised by rapid changes in speech power

6. Discussion on experimental results on acoustic parameter studies

The formant frequencies and the phonetic duration of eight Hindi vowels are presented in table 1. The vowels exhibit their characteristic formant frequencies. The low value of the standard deviations, particularly for Hindi अ (a) and आ (a) indicates acoustic stability of these phonemes.

It was observed that the spread for 1st formant is from 295 c/s (for Hindi इ, i) h, to 707 c/s (for Hindi आ) a and that for 2nd formant is from 675 c/s (Hindi U, उ) to 2400 c/s (for Hindi i). This compares well with the results given by Fant

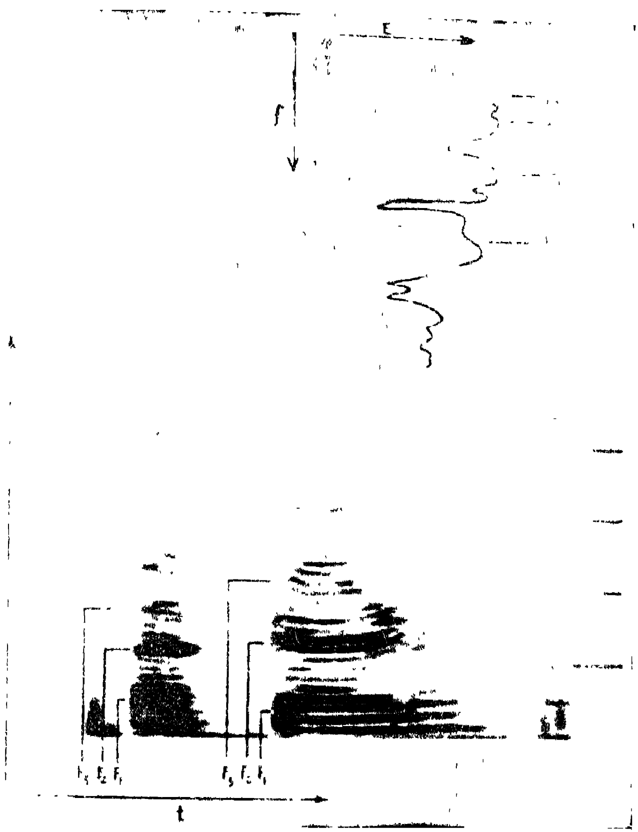

Figure 4a Narrow band spectrograph of word (पकर)

Upper part shows energy Vs. Freq. Plot Lower part shows energy dist in time domain
F_0 Fundamental Freq F_1, F_2, F_3 -- Formants

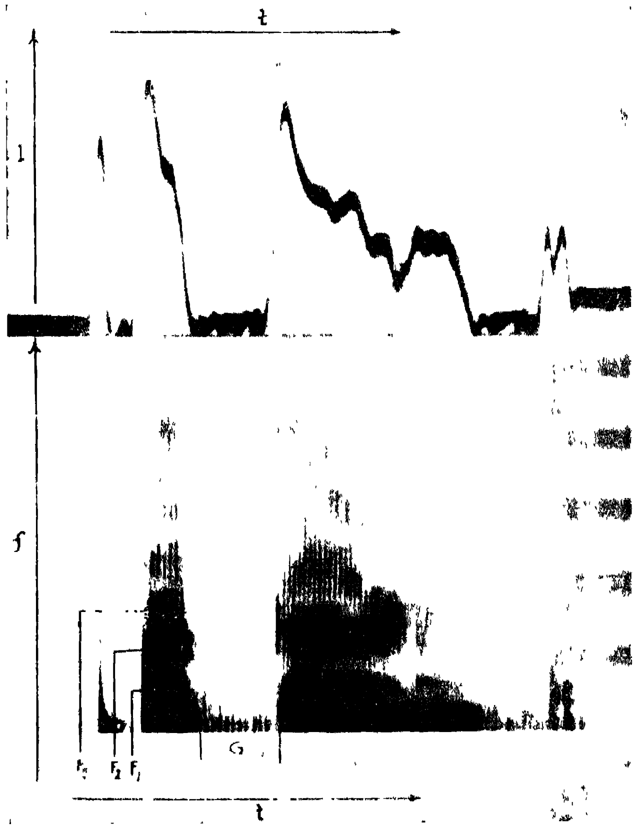

Figure 4b Wide band spectrograph of word (पकर)

Upper part shows intensity Vs. time plot. Lower part shows energy Dist. in time do
F_1, F_2, F_3 - Formant frequencies G - Gap

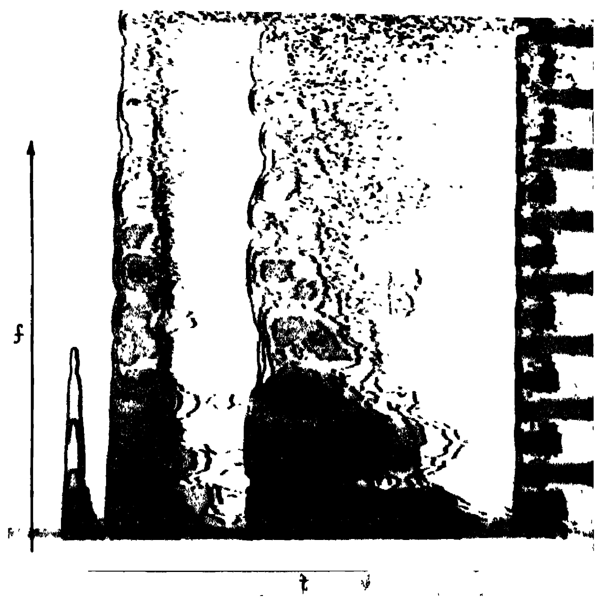

Figure 4e Contour display of the word (पकर)

Figure 5a. Wide band spectrograph of the word (मिठी)
B = Voice Bar
Fa = Anti Resonance

Indian J Phys. Vol 47, No. 10, 1973

Plate 7

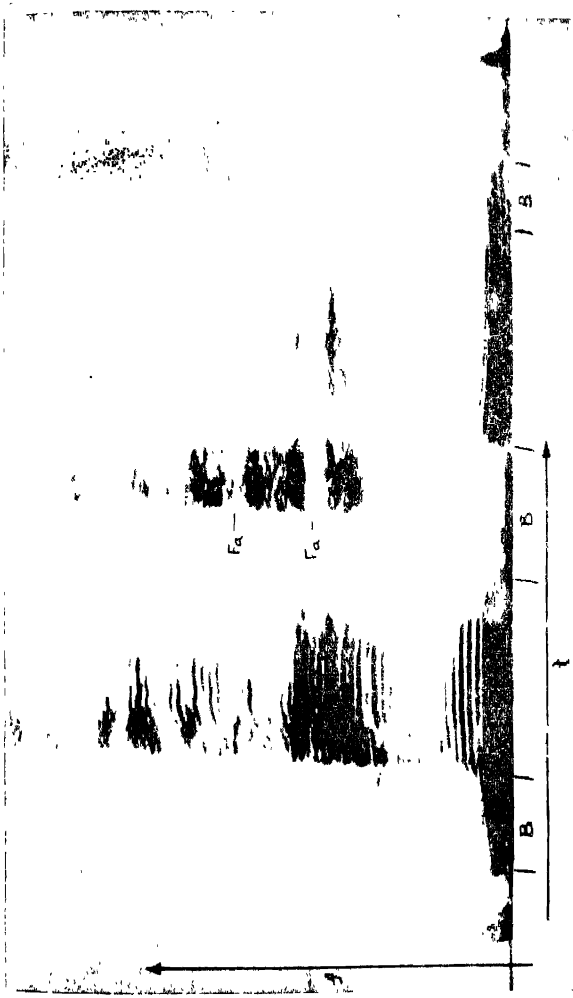

Figure 5b Narrow band spectrograph of the word (मिठा)

B = Voice Bar

F_a = Anti Resonance

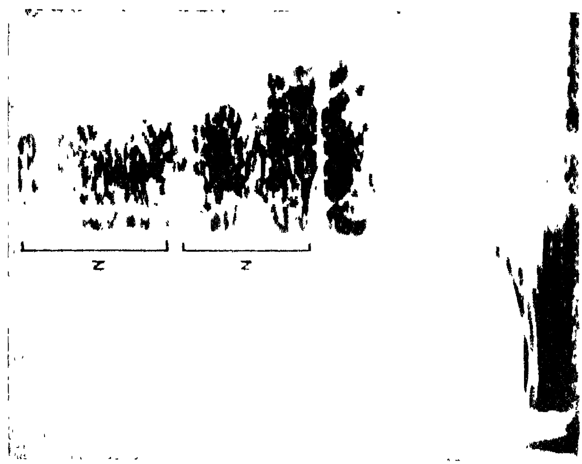

Figure 6a. Narrow Band spectrograph of the word (औंस)
N = Wide Band continuous spectrum

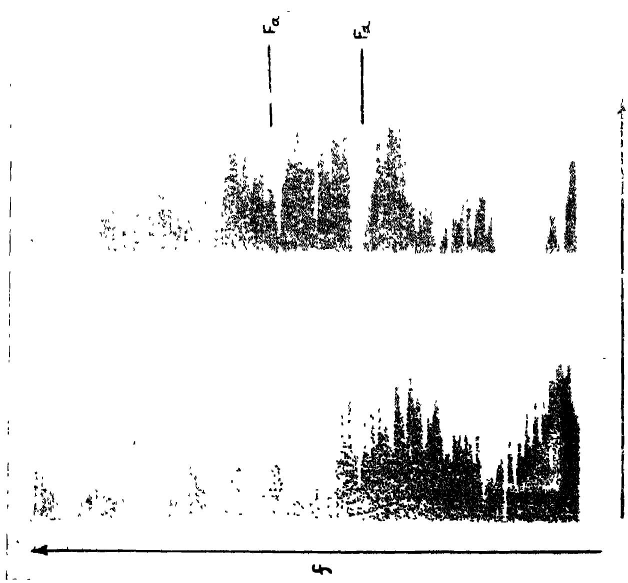

Figure 5c Frequencies vs energy plot showing consonantal anti-resonance

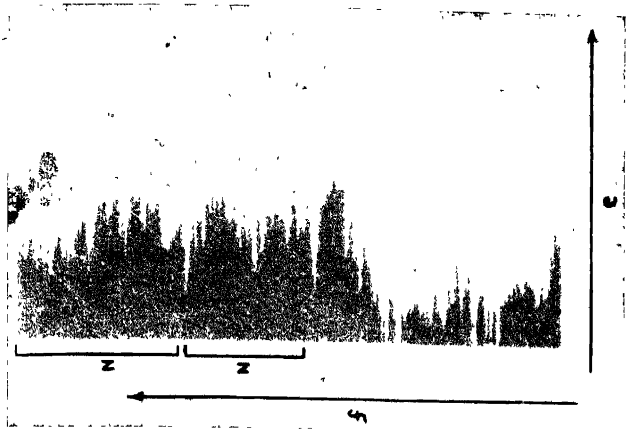

Figure 6c Freq vs energy plot showing wide band continuous spectrum N

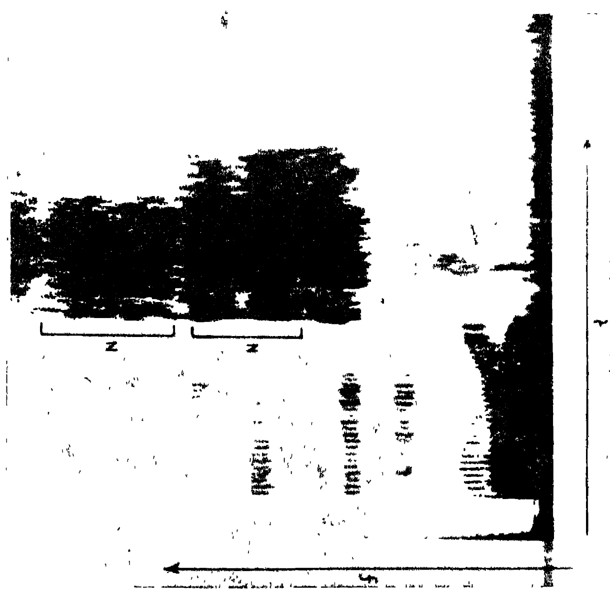

Figure 6b Wide band spectrograph of the word (अँख)
N = Wide band continuous spectrum

Table 1 Formant frequencies and phonetic durations of Hindi vowels

Vowel	no of obs.	F_0 c/s Mean	F_0 c/s S.D.	F_1 c/s Mean	F_1 c/s S.D.	F_2 c/s Mean	F_2 c/s S.D.	F_3 c/s Mean	F_3 c/s S.D.	F_4 c/s Mean	F_4 c/s S.D.	T m.s. Mean	T m.s. S.D.	Remarks
œ–A (अ)	10	107	16	550	35	1402	131	2247	210	3675	—	127	33	No. of observation for F_3 is nine and for F_4 one. SD is not calculated for no. of obs. less than 4
a (आ)	10	101	14	707	36	1220	76	2420	118	3550	—	284	65	No. of observation is 5 for F_3 and 2 for F_4
I (इ)	5	115	20	310	12	2180	218	3060	514	3550	—	76	12	No. of obs. is 4 for F_3 and 1 for F_4.
i (ई)	5	120	10	295	20	2400	300	3230	199	3970	—	324	53	No. of observation is 4 for F_2 and 2 for F_4
u (उ)	3	125	—	366	—	1000	—	1450	—	—	—	108	—	No. of observation is 1 for F_3 and nil. for F_4
U (ऊ)	1	150	—	350	—	675	—	1275	—	—	—	225	—	No trace of F_x
e (ए)	2	112	—	325	—	2250	—	2675	—	3650	—	292	—	—
O (ओ)	3	108	—	383	—	800	—	—	—	—	—	255	—	No trace of F_3 and F_x. No. of observation for F_2 is 2.

F_0 = Average Fundamental Frequency. F_n = N-th Formant Frequency; T = Phonetic duration in milliseconds
S.D. = Standard Deviation from the mean.

(vide section 4.2) for vowels. The phonetic identity of vowels seems to depend not on the absolute values of the formant frequencies but on the relative overall formant structure of the speaker. Third formant frequency seems to serve as an indicator to the formant structure of an individual speaker. The Hindi vowels show considerable overlapping in the F_1-F_2 plane. Even for a single speaker U (ऊ) overlaps with O (ओ) and i (ई) with e (ए) (figure 7). This overlapping does not exist in $F_3^1-F_3^2$ plane (where $F_3^1 = F_3-F_1$ and $F_3^2 = F_3-F_2$) (figure 8). So a clear separation of different vowels and clustering of the same vowels may be obtained giving proper weight to F_3. In coarticulation with consonants the vowel formants show marked shift during transitory parts. The rate and direction of the formant movements are important cues for identification of consonants.

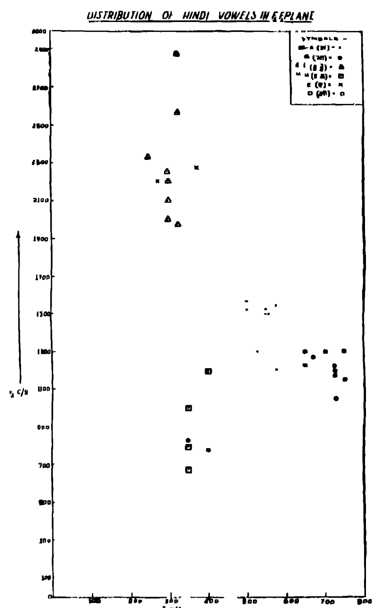

Figure 7. Distribution of Hindi Vowels in F_1-F_2 Plane

We shall now discuss some interesting features of the concentration of energy for consonants. Table 2, presents the concentration of acoustic energy for plosives and some affricates. Though the number of observations are not large enough to be conclusive but they indicate certain features of considerable significance from the acoustic phonetic point of view. Sometimes the energy concentrations are so weak that they are untraceable in the spectrographs.

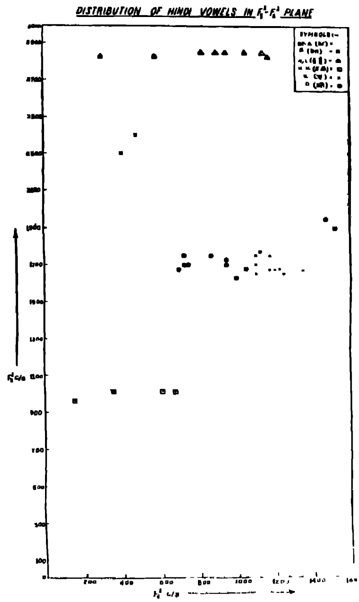

Figure 8. Distribution of Hindi Vowels in $F_3{}^1 - F_3{}^2$ Plane

The first concentration of energy for aveolars and labials is below 1000 c/s, whereas that of volars and dentals, are above 1000 c/s. Affricates exhibit the first concentration at very high frequencies 2500 c/s and above It is often very difficult to separate plosive or fricative part from the aspirated part of an aspirated consonant directly from a study of a normal spectrogram. Of course, the duration of plosion can be determined from the frequency *vs* time plot. However the onset of the vowel transition is usully clearly marked and represents the end of the preceeding consonant. It will therefore, be easy to determine the total duration of the aspirated consonants. The detection of occlusive portion of a consonant, where it exists presents no difficulty.

Table 2 Concentration of acoustic energies of Hindi consonants (plosives & affricates)

		no of obs	Plosion c_1	Plosion c_2	Plosion c_3		no of obs	Plosion c_1	Plosion c_2	Plosion c_3	Aspiration c_1	Aspiration c_2	Aspiration c_3		no of obs	Plosion c_1	Plosion c_2	Plosion c_3		no of obs	Plosion c_1	Plosion c_2	Plosion c_3	Aspiration c_1	Aspiration c_2	Aspiration c_3
Velar	क (k)	8	1300 to 1500	3600 to 3850	4600 to 4800	ख (kh)	2	1150 to 1400	—	—	1150 to 1500	—	—	ग (g)	5	1250 to 1700	2100 to 2450	—			—	—	—	—	—	—
Affricates	च (c)	4	2600 to 2800	3600 to 7000	—	छ (ch)	2	2700 to 3200	4000 to 5500	×	2000 to 2700	3100 to 7500	×	ज (j)	7	2400 to 2800	3500 to 4500	5000 to 5700	झ (jh)	3	2600 to 2800	3500 to 3800	4200 to 4500	150 to 400	1400 to 2000	2100 to 7000
Avelar	ट (t)	3	400 to 600	1500 to 2000	2500 to 3700	ठ (th)	3	400 to 500 (no trace)	1500 to 1700	2800 to 5000	1300 to 1500	2500 to 3100	3400 to 6000	ड (d)	1	500 to 600	1300 to 1700	3500 to 3700			—	—	—	—	—	—
Dental	त (t)	4	1400 to 2000	3500 to 4000	—									द (d)	3	1600 to 1800	3500 to 3700	—	ध (dh)	3	1100 to 2000	3000 to 4300	—	1300 to 1800	3600 to 4500	5100 to 5500
Labial	प (p)	4	200 to 600	1100 to 1300	2100 to 2200	फ (ph)	3	400 to 600	1600 to 1800	—	—	—	—	ब (b)	5	300 to 500	1150 to 1500	1600 to 2500	भ (bh)	1	—	—	—	1600 to 1800	2200 to 2800	3200 to 4500

C_1 = First concentration of acoustic energy; C_2 = Second cncentration of acoustic energy;
C_3 = Third concentration of acoustic energy.

The table 3 represents the concentration of acoustic energy for three groups of consonants namely nasals, laterals and fricatives. The nasal and lateral consonants show concentration of acoustic energy at lower as well as higher ranges, whereas the fricatives show concentration of acoustic energy at higher ranges only.

Tables 4 and 5 represent time duration for occlusion, plosion, aspiration and total average time, corresponding to the consonants of table 2 and 3 It will appear from table 4 that the total duration of consonant minus the occlusion period could be effectively used as a distinctive feature for distinguishing aspirated from unaspirated counterparts, the aspirated consonants exhibiting much larger value of duration.

Table 3. Concentration of acoustic energies for Hindi consonants (nasal, lateral & fricatives)

		no. of obs.					no. of obs.	c_1	c_2	c_3	c_4
Nasal	न (*n*)		150 to 500	—	—	(m)		150 to 450	1150 to 1400	2200 to 2400	—
Lateral	र (r)	6	150 to 500	1200 to 1700	2000 ro 2300	ल (l)	7	200 to 450	1400 to 1700	2300 to 3000	3400 to 3800
Fractive	स (s)	7	1500 to 1800	3600 to 4500	5000 to 7000	श (s′)	3	1700 to 3000	3700 to 6000		
	ह (h)	5	150 to 700	1150 to 1800	3500 to 6000	—	—	—	—	—	—

C_1 = First concentration of acoustic energy; C_2 = Second concentration of acoustic energy
C_3 = Third concentration of acoustic energy C_4 = Fourth concentration of acoustic energy

Table 4. Period of occlusion, plosion and aspiration for Hindi consonants

consonant plosive	no. of obs.	t_o	t_p	consonant plosive	no. of obs.	t_o	t_{asp}
क (k)	8	98	24	ख (kh)	2	126	130
ग (g)	5	60	23	—	—	—	—
च (c)	4	67	38	छ (ch)	2	76	130
ज (j)	7	133	30	झ (jh)	7	103	93
ट (t)	3	94	22	ठ (th)	3	153	81
ड (d)	1	27	18	—	—	—	—
त (t)	4	99	18		—	—	
द (d)	3	63	18	थ (dh)	3	60	73
प (p)	4	76	22	फ (ph)	3	78	135
ब (b)	5	112	18	भ (bh)	1	99	63

t_o = Occlusion period
t_p = Plosion period
t_{asp} = Aspiration period

Table 5. Time duration of Hindi consonants

	no. of obs.	average total duration
न (n)	5	95
म (m)	6	96
ल (l)	7	132
र (r)	6	151
स (s)	7	194
श (s′)	3	168
ह (h)	5	80

Acknowledgements

The authors express their grateful thanks to Prof. Djorje Kostic, Director Institute of Experimental Phonetics, Beograd, Yugoslavia for the helpful discussion, and Prof. C. R. Rao, F.R S., Director, Research and Training School, ISI and Late Prof. P. C. Mahalanobis, F.R.S. for their kind interest in the work. The authors also thank Shri J Das, Shri B. Mukherjee and Sm. Krishna Dutta Roy for their help in the work. The authors express their gratitude to the authorities of All India Radio particularly the Hindi News announcers of AIR, New Delhi, for providing the Indian Statistical Institute with their recorded voices for a number of Hindi Speech Sound.

References

Barczinski L. & Thienhaus E. 1935 *Klangspectren und Laustarke deutscher Sprache, Arch. Neerl. de Phon. Exp.* II, 47.

Cooper F. S., Delattre P. C., Libberman A. M , Borstand J M. & Gerstman L. J. 1952 *J.A.S.A.* **24**, 597.

Chiba T. &Kajiyama, 1941 *The Vowel—its Nature and Structure*, Tokyo.

Dutta Majumder D. & Dutta A. K. 1969, *JITE*, **15**, 4.

Dutta Majumder D. & Dutta A. L. 1968, *Indian J. Phys.* **42**. 7.

Dutta Majumder D. & Dutta A. K. 1968 *Automazione E Strumentazione*, Milan.

Declerk J. L. 1965 *Proc. Fifth. International Congress in Acoustics*, Liege, **A26**.

Dunn H. K. 1950 *J. Acoust. Soc Amer.* **22**, 740.

Fant G. 1959 *Acoustic Analysis and Synthesis of Speech with Application to Swedish, Ericsson Technics* 11,.

Final Rept. AFCR—64-300. 1963 Research Laboratory of Electronics, MIL, Cambridge,