PHONETIC ANALYSIS OF BREATHY (MURMURED) VOWELS IN GUJARATI *)

Eli Fischer-Jørgensen

1. INTRODUCTION.

The Indian language Gujarati, spoken in a district north of Bombay, has a contrast between clear and breathy vowels. This contrast has been treated both from a phonemic and a phonetic point of view by P.B. Pandit (1), who called the breathy vowels "murmured", a terminology which will be used in the following. We also follow his transcription according to which underlining indicates murmur.

Gujarati has eight clear vowels / i e ɛ a ɔ o u ə/ and the corresponding murmured vowels / i̲ e̲ ɛ̲ a̲ ɔ̲ o̲ u̲ ə̲/. The nasalized vowels / ĩ ɛ̃ ã ɔ̃ ũ ə̃ / and nasalized murmured vowels are not included in the present investigation with the exception of the example /pɔ̃ci - pɔ̲̃ci/. Nasalized murmured vowels are rare.

Historically the murmured vowels have developed from (i) vowel + h + vowel, (ii) h + vowel, and (iii) from a fusion of the aspiration of a final voiced aspirated stop with the preceding vowel (2).

Phonemically murmured vowels may be interpreted as vowel + h, but phonetically they form one segment. Pandit (1) gives the following description: "Murmur is voiced breath, low pitched and simultaneous with the vowel" (p. 169) - "sotto voce, with voicing and slight lowering of pitch" (p. 170). He also published a spectrogram and drew attention to "the presence of random distribution of energy, more noticeable at higher frequencies" (p. 172).

A spectrographic investigation of the formant frequencies of murmured vowels compared to clear vowels has been undertaken by Radhekant Dave (2).

The purpose of the present investigation is to give a more all round phonetic description of these vowels in respect of airflow, duration, fundamental frequency, overall intensity, formant frequen-

---

*) This is a summary of a paper to be published in Indian Linguistics 1968. The numbers of tables and figures have been retained, although some have been left out here.

cies, and distribution of spectral energy, and including a restricted number of auditory tests.**)

2. THE INFORMANTS.

Seven informants have been used: RD, RT, SK, DD, GU, PPB, PvB (PvB is female, the others male). DD had a high-pitched falsetto voice, and only a small part of his recordings have been used. All informants have murmured vowels in their natural speech. In RD's and PBP's speech murmur is optional. RT tended to pronounce an [h] when reading aloud, and a good deal of his examples had to be removed. RD, SK and RT have no consistent distinction between close /e o/ and open /ɛ ɔ/. Only [e o] have been used in the transcriptions of their speech.

3. THE TEXTS.

The texts consisted of series of isolated words and short sentences, spoken several times by each speaker. It appears from the tables and graphs which words were used and how often they were repeated. Series of words, particularly when repeated, may lead to a sort of rhythmical singsong which may reduce the differences between the two types of words. Nevertheless only averages of the word series are given in the tables and have been statistically treated. This procedure has been chosen, partly because more examples had been recorded of the word series, partly because the sentences were spoken in pairs by some of the speakers, and this gave rise to contrastive intonations. RT's tape recording forms an exception because he also spoke the sentences five times and not in pairs. In this case words and sentences have been combined in the averages. - For the other informants the sentence examples are used for comparison.

**) The investigations have been carried out in the laboratory of our institute, and I am grateful to several of the staff members and students for help.

The experiments with synthetic and filtered vowels have been carried out in cooperation with Jørgen Rischel, who has been responsible for the technical part of these experiments.

I am particularly indebted to Radhekant Dave, who has found most of the informants, set up the word lists, controlled the recordings, and listened to the tests.

4. INSTRUMENTAL SET-UP.

4.1. A tape recording of PBP was made in Dehradun in 1957, a recording of PvB in the Institute of Phonetics in Amsterdam in 1964, tape recordings of the other informants in Copenhagen 1966-67.

4.2. A selected number of words containing the vowels [i e ɛ a o ə] from the tape recordings of six of the informants (DD was left out) have been used for spectrographic analysis by means of the Kay Electric Sonagraph. Narrow band spectrograms and narrow band sections (both with "High Shaping" and "Flat 2") were taken of all the words selected, and some spectrograms were taken with wide band (see Figs. 1-3). Moreover some spectrograms were taken with high compression, and some with half speed of the tape in order to examine the presence of noise in the higher frequency regions.

The spectrograms were used for measurements of formant frequency, formant amplitude and for the measurement of the amplitude of F_o and of the second and third harmonics H_2 and H_3. (Generally F_3 and F_4 could not be measured for [o].) Amplitude measurements were made on the basis of the sections. For the formants above c. 1ooo cps the high shaping section had to be used, but on the basis of the high shaping curve of the instrument the measurements were corrected by subtracting the appropriate number of decibels. As a measurement of the amplitude the root mean square value has been chosen, which means that the harmonics within the range of the formant have been summed up according to a pre-established table of dB-values. Two close formants make difficulties, but nevertheless this method was found to give more reliable results than the measurement of the envelope peak, which can be difficult to place exactly. (3).

Some reservations must be taken as to the exactitude of the measurements, because the new model of the Kay Electric Sonagraph (A 6o61) has been found to produce difference tones by intermodulation. However, the formant peaks of normal vowel spectrograms do not seem to be distorted (spectrograms taken with high shaping, with "Flat 1", "Flat 2" and with low pass filtering did not show any difference in amplitude relations), but the valleys in between the formants are hardly reliable.

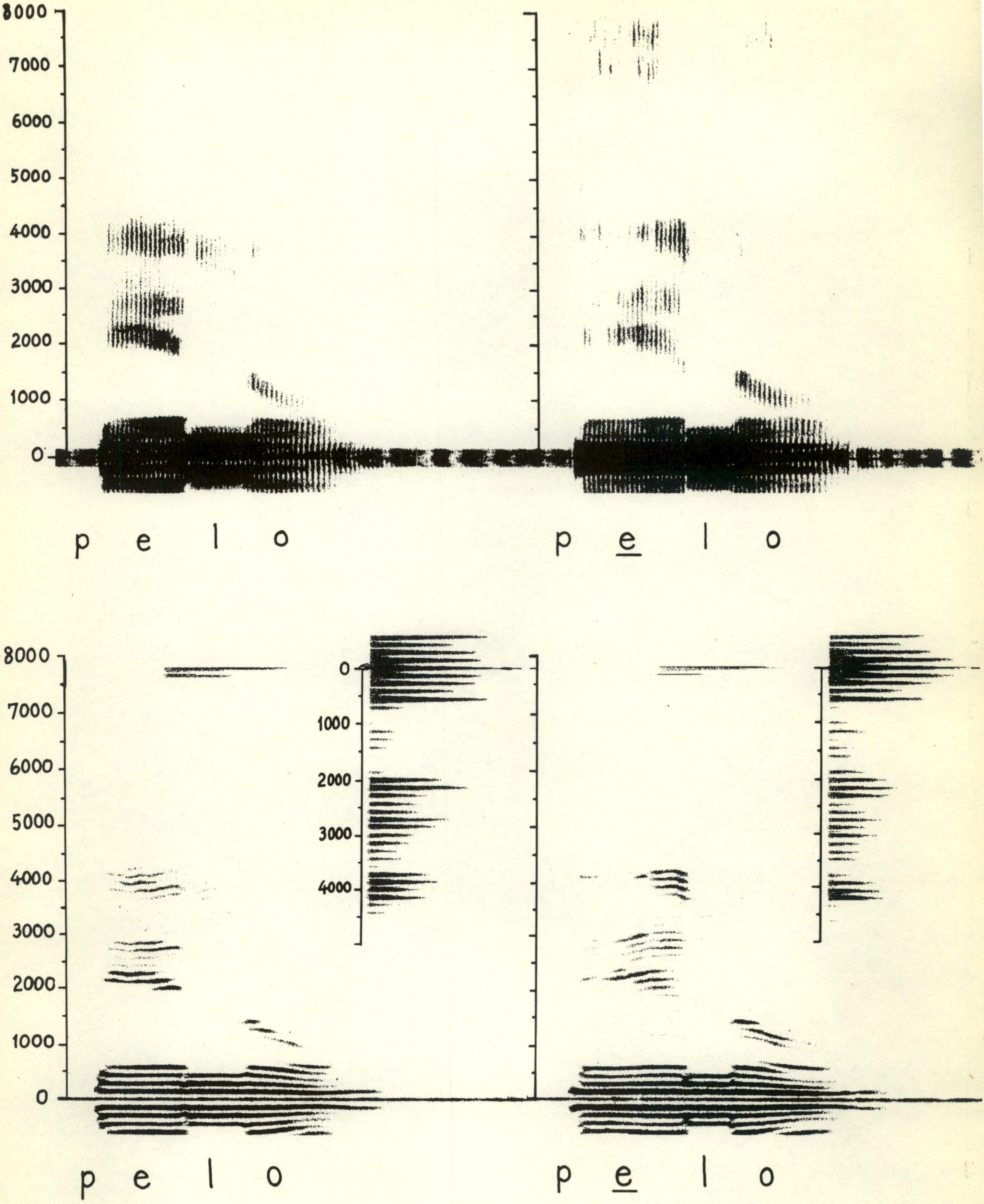


Fig. 1. Spectrograms of pelo and pe̲lo (RD).
Wide band and narrow band spectrograms and narrow band section.

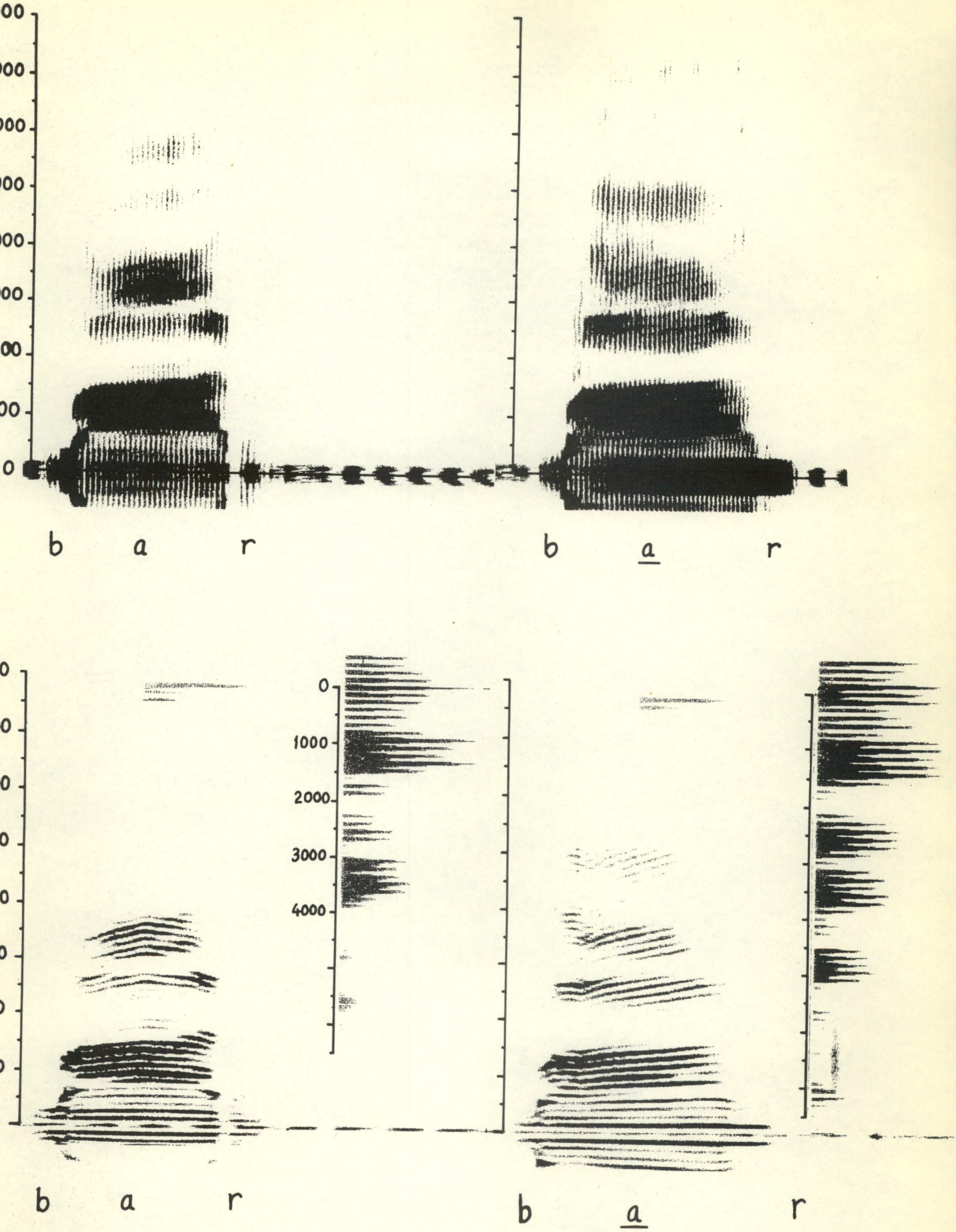


Fig. 2. Spectrograms of bar and ba̲r (GU).
Wide band and narrow band spectrograms and narrow band section.

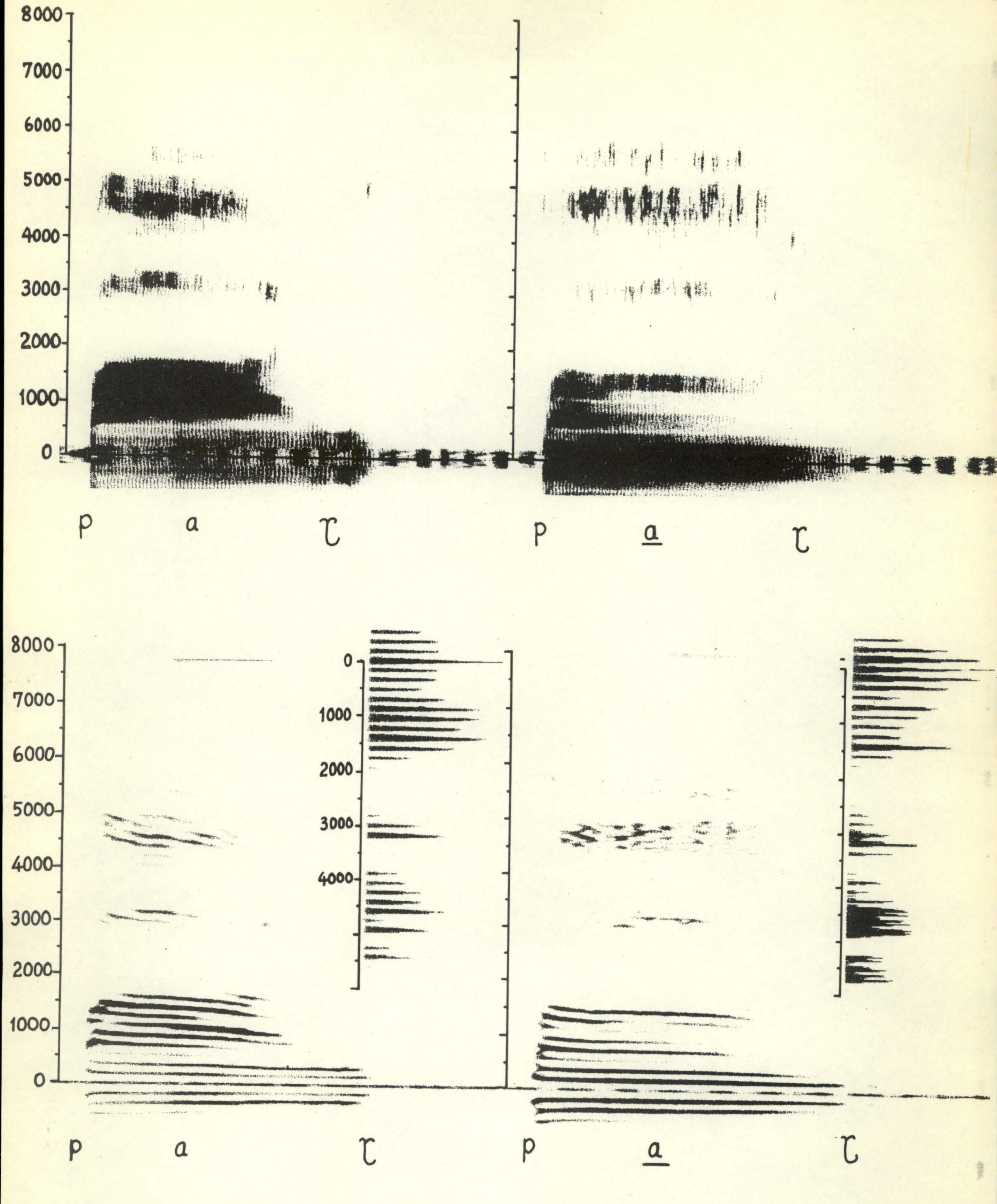


Fig. 3. Spectrograms of paɽ and pa̲ɽ (PvB).
Wide band and narrow band spectrograms and narrow band section.

4.3. All of the tape recordings were used for analysis of fundamental frequency and overall intensity be means of the trans-pitchmeter and intensity-meter connected with a four channel Elema Mingograph. The following curves were taken: (1) a duplex oscillogram, (2) a sharply highpass-filtered logarithmic intensity curve with an integration time of 5 ms (2.5 ms for the female subject), (3) an intensity curve with flat frequency response and an integration time of 1o ms (5 for the female speaker), and (4) a fundamental frequency curve. The highpass-filtering was made by means of an external passive filter (insertion loss filter), with a nominal cutoff frequency of 315 cps. The 3 dB point was at 34o cps, and the attenuation was more than 45 dB at 27o cps (see Fig. 4).

4.4. The selected texts used for spectrographic recording were also used for an oscillogram recorded on the Mingograph with double speed of the paper (2o cs) and with the tape played back at one quarter of the speed. In this way the upper frequency limit of the mingograph recordings was raised from 8oo to about 32oo cps, and the distance from one period to the next was eight times the distance of the other mingograms taken with a paper speed of 1o cm per second. The purpose was to see details of the temporal change of the vowel.

4.5. Three of the subjects (RD, RT and GU) spoke the list of words and sentences into the Aerometer used for airflow measurements. Curves of fundamental frequency and overall intensity were taken simultaneously from a throat microphone. Only one intensity curve was taken (in some cases with and in some cases without highpass filtering). The words were spoken in groups of four and five with varying order (see Figs. 6, 7 and 8).

4.6. A few recordings comprising five isolated vowels and five words, spoken each six times by RD, were made by means of the Fabre glottograph, combined with curves of fundamental frequency and airflow (see Fig. 5).

4.7. Attempts were made to synthesize sounds with the spectral characteristics of clear and murmured vowels. The synthesizer has been described by Jørgen Rischel (4).

4.8. Filtering. Some filtering experiments were made with highpass, lowpass and a combination of highpass and lowpass filtering of clear and murmured vowels with the purpose of observing the consequences for the perception of murmured vowels. For the highpass

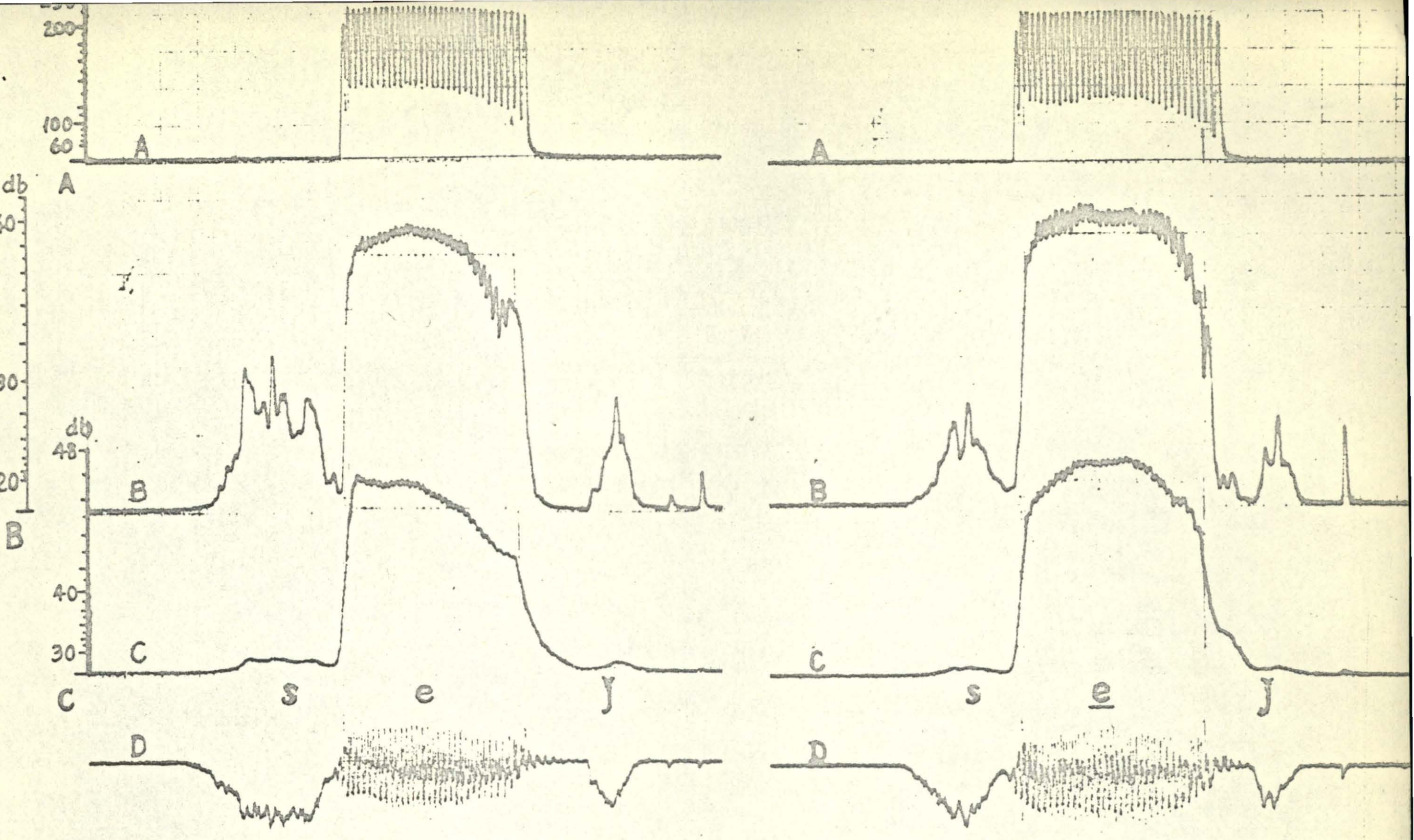


Fig. 4. Mingogram of seǰ and se̲ǰ (RD).

A. Fundamental frequency curve.

B. Highpass filtered logarithmic intensity curve.

C. Linear intensity curve without filtering.

D. Duplex oscillogram.

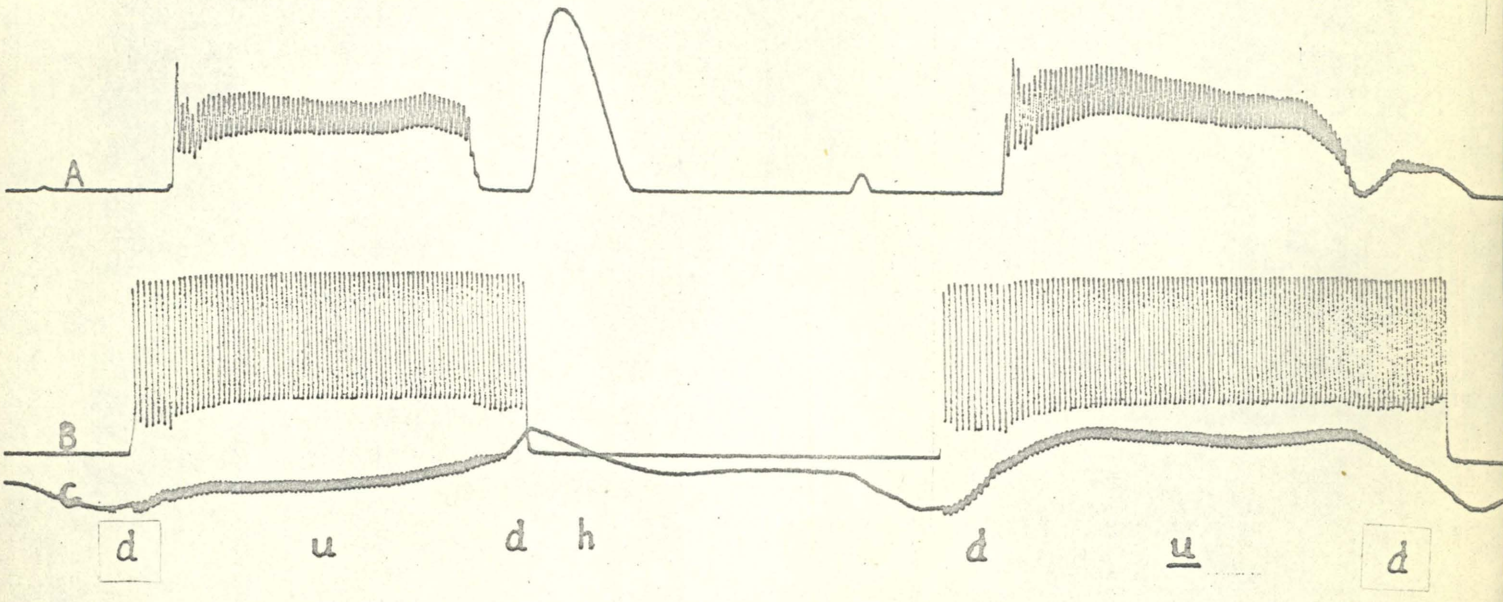


Fig. 5. Mingogram of dudh and du̲d (RD).

A. Airflow curve.

B. Fundamental frequency curve.

C. Glottogram.

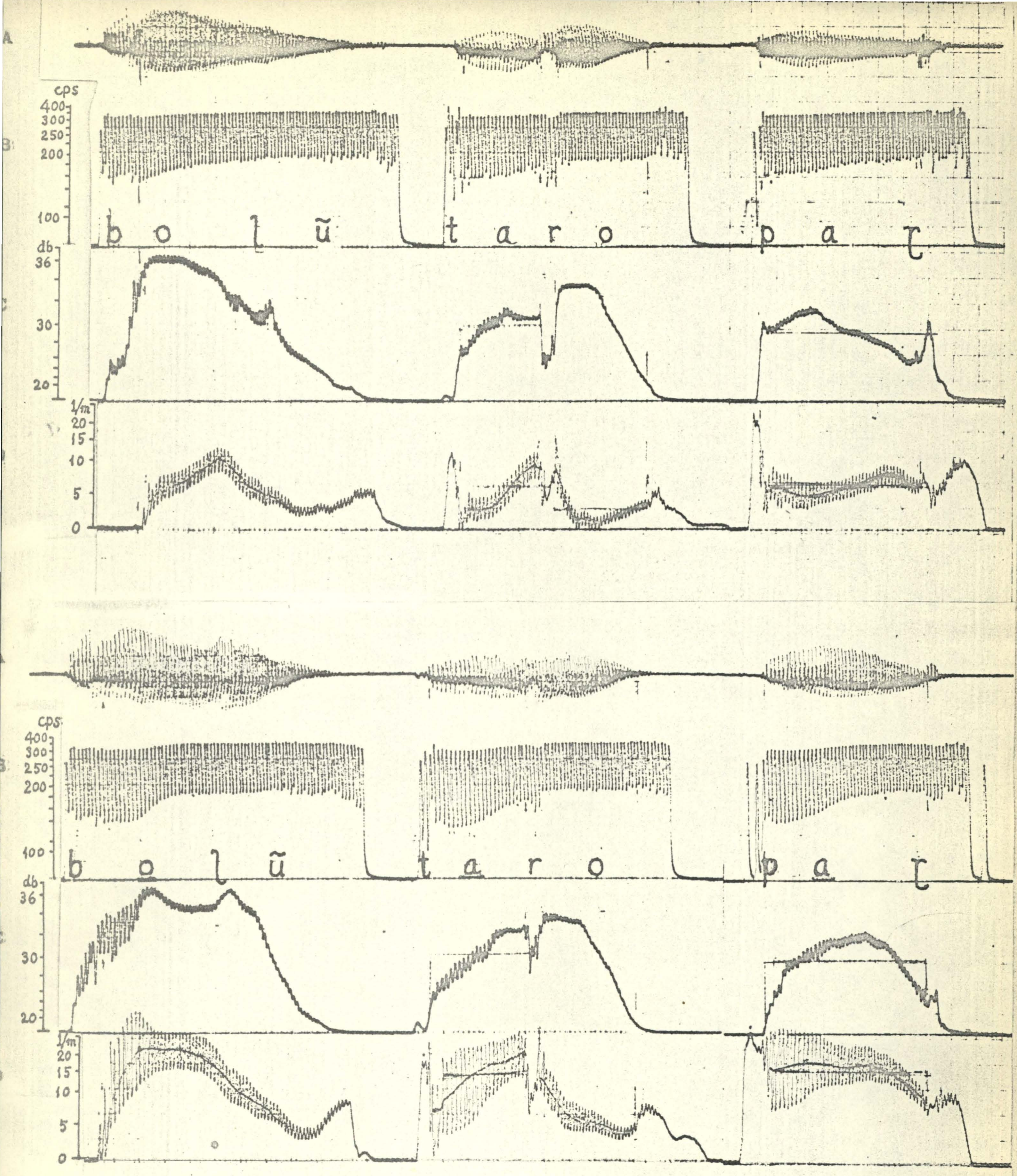


Fig. 6. Mingograms of boɭũ, taro, paɽ and bo̲ɭũ, ta̲ro, pa̲ɽ (RD).

A. Oscillogram.

B. Fundamental frequency curve.

C. Logarithmic intensity curve, highpass filtered 315 cps.

D. Airflow curve.

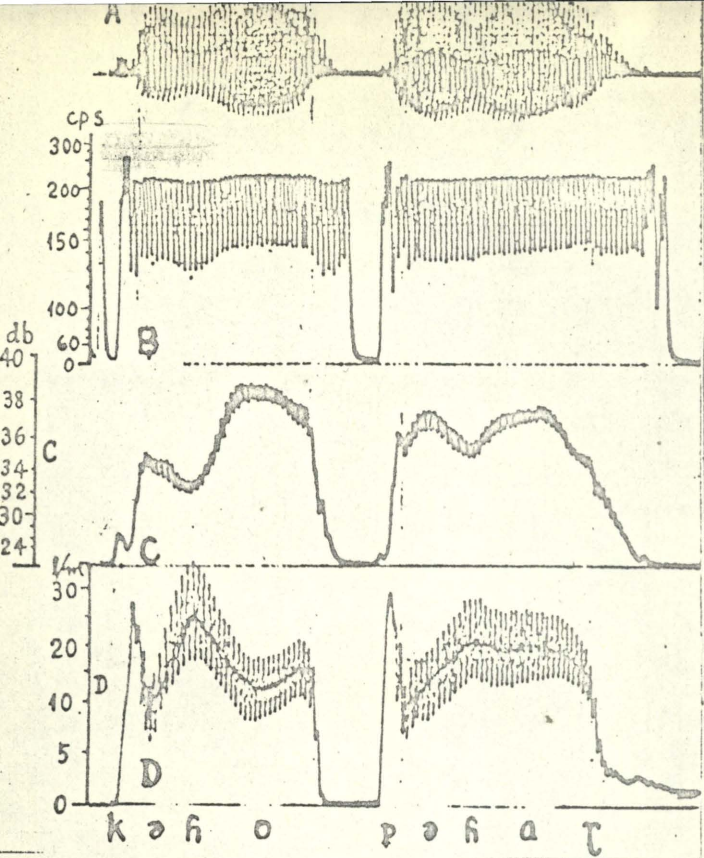


Fig. 7a.

Mingogram of intervocalic h in ko̲, pa̲ʈ (RT).

A. Oscillogram.
B. Fundamental frequency curve.
C. Intensity curve.
D. Airflow curve.

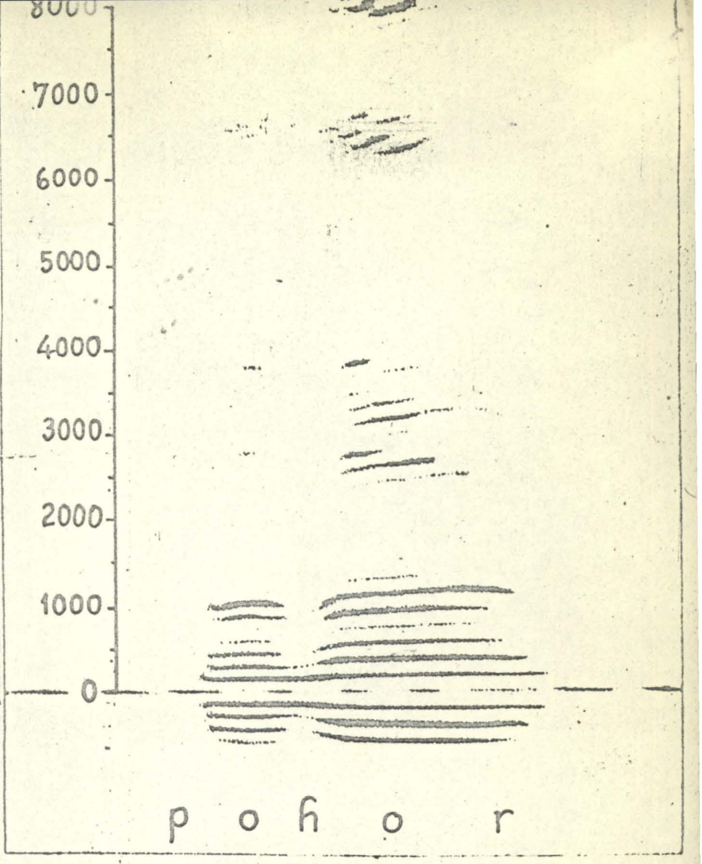


Fig. 7b.

Narrow band spectrogram of intervocalic h in po̲r (RT).

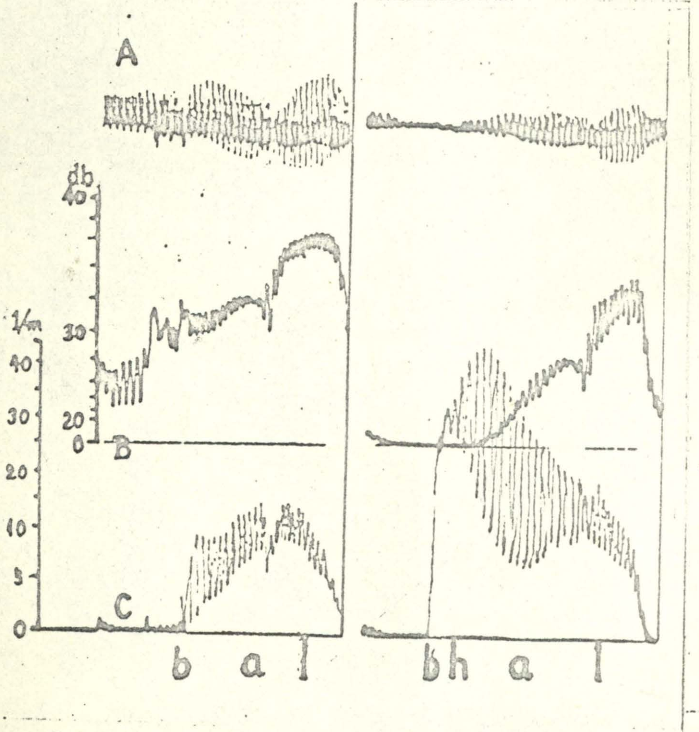


Fig. 8a. Mingogram of b and bh (RD).

A. Oscillogram.
B. Fundamental frequency curve.
C. Airflow curve.

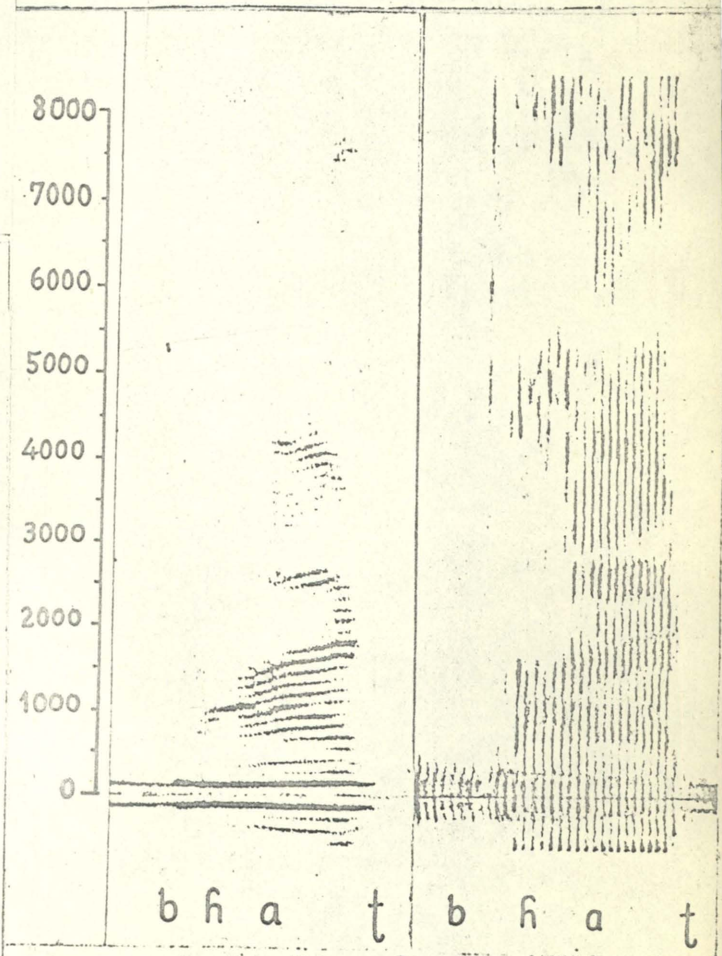


Fig. 8b.

Narrow and wide band spectrograms of bh in bɦaʈ (RD).

and lowpass filtering the heterodyne filter was used (5), and the cutoff frequencies (here defined as the points of 3 dB attenuation) were set to 23o cps and 31oo (in one case 32oo) cps respectively. This filter was also used, in connection with another filter, to exclude frequencies between 2oo and 5oo cps. Two passive filters were used for a similar purpose with the frequency limits 16o and 5oo cps.

The filters were very steep and the attenuation should be more than 4o dB some 5o cps from the cutoff frequency, except for the lowpass filtering at 31oo and 32oo cps, which was performed with a somewhat less sharp cutoff. Control measurements of the tape recordings showed, however, that the attenuation of harmonics outside the passband(s) amounted to only some 25 dB. This is clearly a distortion effect, which may be ascribable to the signal level being too high during the recording. *)

5. RESULTS.

5.1. Physiological Analysis.

5.1.1. Airflow. According to the subjective impression, particularly the impression of the speaker, the most obvious characteristic of murmured vowels is an increase of airflow.

The airflow has been measured by means of the aerometer (see section 4.5), and the curves indicate the amount of air passing per time unit. The measurements comprise the maximum airflow, the average airflow (the scale l/m indicates litre per minute), and the distance (in cs) from the beginning of the vowel to the maximum. The time constant of the instrument is sufficiently small to show also the voice ripple. A mid-line was drawn through the ripple by hand. The maximum was measured as the highest point on this line. The average airflow of the vowel has been found by a graphical approximation by means of a horizontal line drawn by hand. This could be done with an exactitude of somewhat below 1 l/m.

The airflow has been measured for three subjects (RD, RT and GU), the total material comprising 1442 words. Examples of the curves are given in Fig. 6.

---

*) In a later filtering experiment with an improved setup we succeeded in keeping the out-of-passband distortion products (measured **at the loudspeaker) some** 4o dB below the level of the strongest harmonics. However, a listening session comprising a restricted number of words made under these **improved conditions** gave only slightly deviating results from those presented in § 8.3.3 below.

The main results are given in Table 1.*)

Table 1 AIRFLOW

	Maximum Airflow			Average Airflow		
Speaker	RD	RT	GU	RD	RT	GU
Number of words	552	676	214	5o4	644	214
1/m murm	18.6	17.2	36.6	15.1	12.1	3o.5
1/m clear	8.9	11.4	25.7	7.2	8.3	19.3
difference	9.7***	5.8***	1o.9***	7.9***	3.8***	11.2***
increase in %	1o9	55	46	11o	46	6o

RD and RT have also spoken the word pairs in sentences. For RD the difference between clear and murmured vowels was slightly larger in the sentences, for RT considerably larger.

The averages of the different word pairs in the series of isolated words are given in graphical form in Fig. 9. **) The stability of the difference appears clearly from the graph. There is no single exception. For single word pairs (words spoken under the same conditions in the same reading) the difference is also 1oo per cent stable in the speech of RD and GU, 96-97 per cent in the speech of RT.

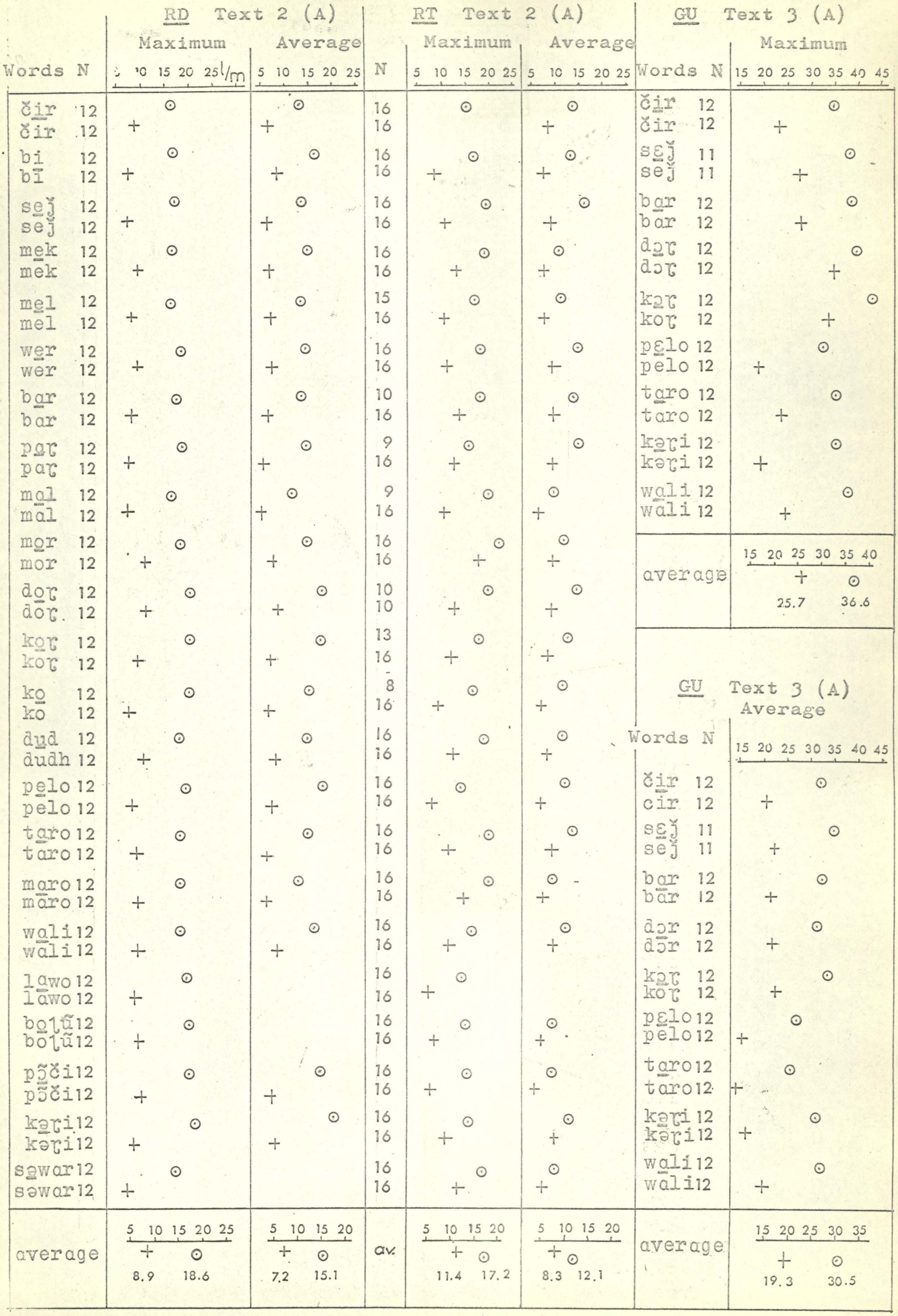
---

*) In this and the following tables the level of significance is indicated by means of asterisks, one asterisk indicating a significance level of 5 per cent, two asterisks a level of 1 per cent, and three a level of o.1 per cent. The significance has been calculated by means of a pair test (see Croxton, Elementary Statistics, Dover 1959, p. 241) generally based on the word averages or in the cases where the number of different pairs were too small (e.g. PvB, DD and sometimes GU), on a comparison between single pairs within the same reading.

The difference always indicates murmured vowel minus clear vowel, i.e. the number is positive if the murmured vowel has a higher value than the corresponding clear vowel.

**) In this and the following graphs clear vowels are indicated by crosses, murmured vowels by rings.

Fig. 9. AIRFLOW (⊙ = murmured + = clear)



5.1.2. In dissyllabic words the characteristic difference in airflow is spread out over both syllables. Only in RD's speech is the difference in the second vowel considerably smaller than the difference in the first vowel (see Fig. 6). RT has even a greater difference in the second vowel. On the other hand RD and, partly GU, have more airflow in an initial unvoiced consonant before murmured vowel than before clear vowel (see Fig. 6).

5.1.3. The airflow in the beginning and end of the vowels is strongly influenced by the surrounding consonants. It is often particularly low after voiced stops, which, in Gujarati, may be more or less implosive, particularly often so in PvB's and GU's speech.

5.1.4. A very short series of whispered words were spoken by RD, There is also in whispered speech an obvious difference in airflow between murmured and clear vowels, but the difference is smaller than in normal speech.

5.2. Vocal Cord Opening

The increased airflow of murmured vowels can be explained either by a wider glottis, or by an increased activity of the expiratory muscles or by both.

It has only been possible to investigate the glottal opening of one subject (RD). By means of a normal laryngoscope combined with stroboscopic light his glottis has been observed by a medical doctor and two phoneticians while he pronounced clear and murmured [a] and [ɛ]. All have noticed a wider opening in the rear part of the glottis for murmured vowels.

Moreover, recordings were made with the Fabre glottograph (see section 4.6) of a text (containing 4 word pairs and four pairs of isolated vowels) spoken 6 times by RD. - No calibration was possible, and the zero-line was not stable, but the average level of the curve of murmured vowels was in all cases higher than that of neighbouring clear vowels. The differences were most pronounced for [u-u̲] (see Fig. 5). This higher level indicates a higher resistance, which may be due to the opening of the glottis. An attempt was made to take glottograms of clear and murmured vowels with the new photo-electric glottograph of the Institute of Phonetics, but without success. It proved very difficult to place the tube with the photocell in the appropriate position because of Dave's narrow throat and strong reflexes.

5.3. Hypothesis about Activity of Expiratory Muscles.

Electromyographic investigations have not been undertaken. But as the measurements of overall intensity do not show any consistent difference between clear and murmured vowels, it is highly probable that the loss of intensity which one might have expected in murmured vowels because of the leaking glottis, is compensated for by a stronger activity of the expiratory muscles, keeping the subglottal pressure on the same level. This assumption is supported by Dave's subjective impression of a stronger "stress" in murmured vowels.

It is thus probable that the stronger airflow of murmured vowels is due both to glottal and expiratory conditions.

6. RESULTS OF THE ACOUSTIC ANALYSIS

6.1. General Remarks.

The measurements of duration, fundamental frequency and intensity have been based partly on the tape recordings (below indicated by T), partly on the aerometer recordings (below indicated by A). The spectral analysis has been based on tape recordings only.

6.2. Duration

6.2.1. The material included in the general averages (i.e. all word lists and RT's sentences)comprises 2oo7 words in all. The main results are given in Table 2, and the averages of single words in Figs. 1o and 11. The duration is measured in centiseconds, and the measurement was made with an exactitude of one cs.

It appears from the table that murmured vowels are longer than clear vowels and that the difference is significant for all speakers. But both the absolute and the relative difference is small. The sentences show somewhat larger differences (15-47 %), particularly for RT (45 %).

6.2.2. On the whole dissyllables show a somewhat greater relative difference than monosyllables, which increases the difference between the words with murmured and clear vowels(see Fig. 1).

6.2.3. The well known general tendency to lengthen the vowel before [r] and to shorten the vowel before stops is also seen in this

Table 2 DURATION (Word Series)

Monosyllables

Speaker	RD(A)	RD(T)	RT(A)	RT(T)	GU(A)	GU(T)	SK	PBP	PvB	DD
Number of Words	336	132	308	72	120	100	142	60	46	8
cs. murm.	29.3	27.9	22.1	23.4	32.4	24.3	18.0	32.6	35.3	19.0
cs. clear	28.1	24.0	19.3	21.2	30.2	21.4	16.7	30.5	33.2	15.8
Difference	1.2	3.9	2.8	2.2	2.2	2.9	1.3	2.1	2.1	2.2
Increase in %	4	17	15	11	7	14	9	10	6	14

Dissyllables

Number of Words	120	48	160	87	96	60	65	12	8	34
cs. murm.	17.9	16.9	14.5	16.8	15.3	13.7	12.1	16.5	19.5	16.0
cs. clear	15.5	13.7	12.1	12.5	14.3	12.2	10.5	14.4	16.8	13.5
Difference	2.4	3.2	2.4	4.3	1.0	1.5	1.6	2.1	2.7	2.5
Increase in %	16	24	22	35	8	15	17	18	16	23

Monosyllables + dissyllables

Number of Words	456	180	468	159	216	160	207	72	54	42
cs. murm.	26.3	25.0	20.1	19.7	24.8	20.3	16.3	29.9	33.1	16.6
cs. clear	24.8	21.3	17.4	16.4	23.2	17.9	14.9	27.8	30.9	13.9
Difference	1.5***	3.7***	2.9***	3.3***	1.6***	2.4***	1.4***	2.1**	2.2**	2.5***
Increase in %	6	17	18	24	8	14	11	10	7	21

Fig. 1o. DURATION (⊙ = murmured + = clear)

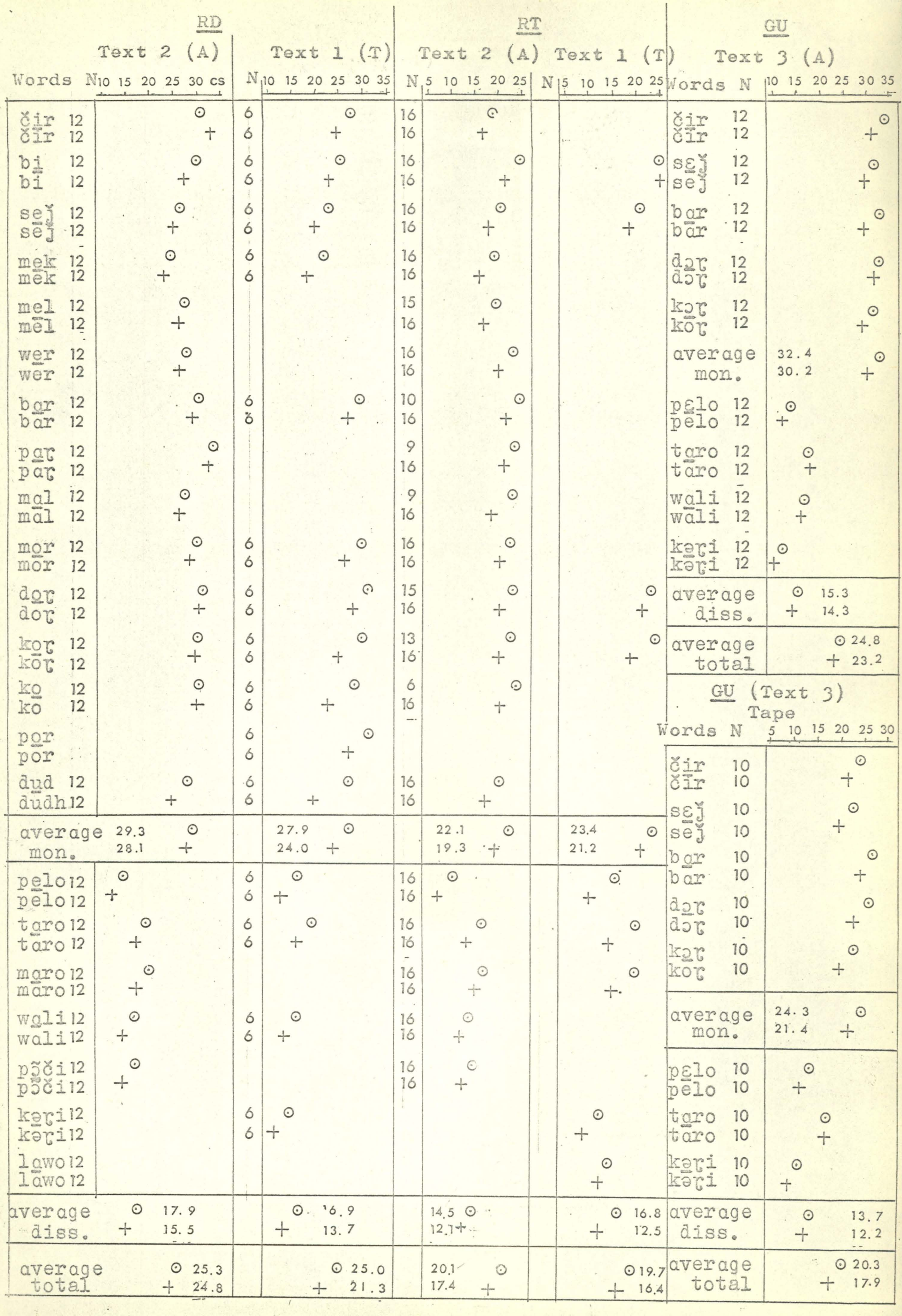
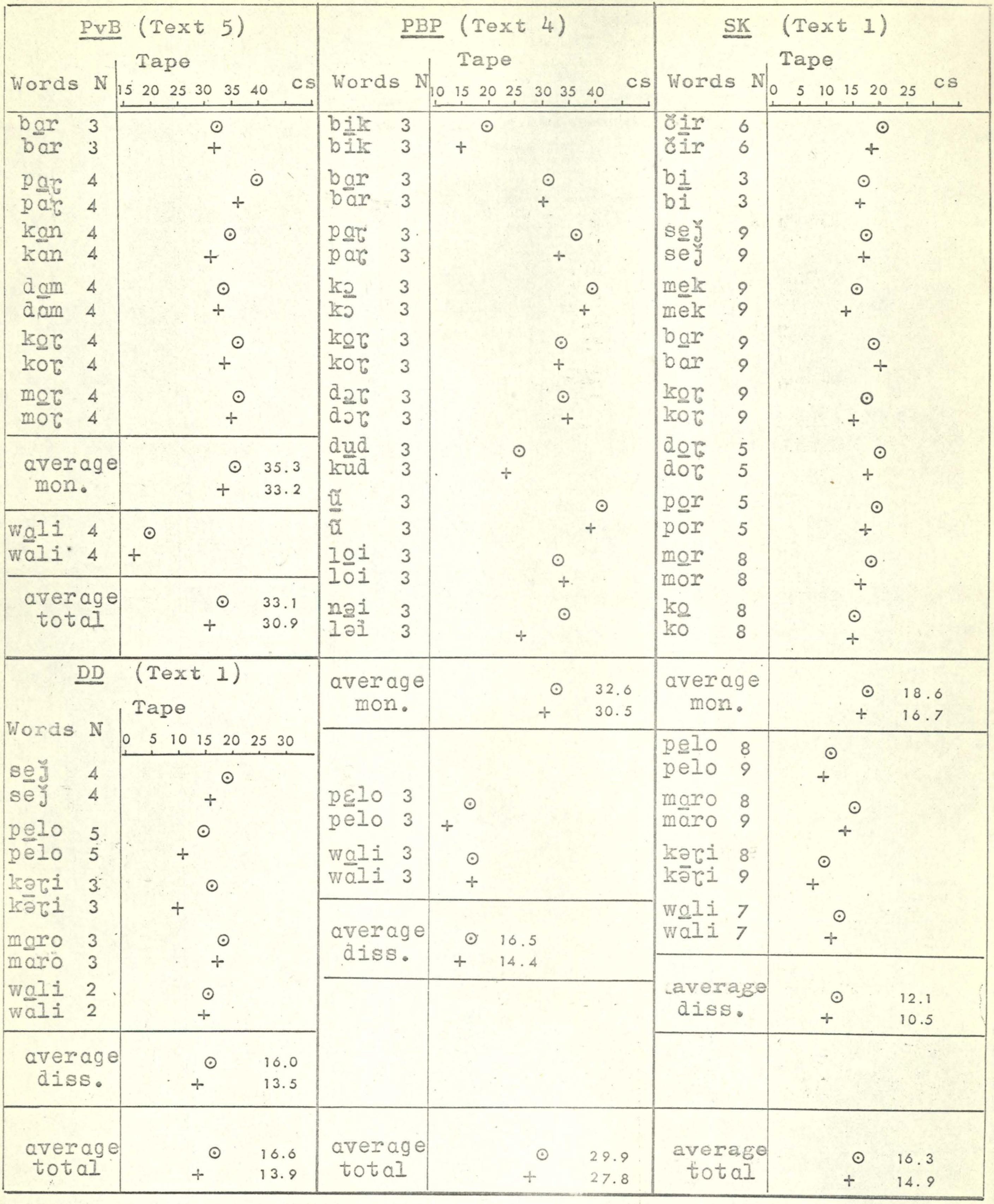


Fig. 11. DURATION (⊙ = murmured + = clear)



material.

6.2.4. RD has also spoken five examples of long isolated [a̲] and four of [a]. In this case the relation between murmured and clear vowels was reversed ([a̲] 54 cs, [a] 82 cs) without any overlapping. The reason is probably that in such long vowels the great amount of air required for the murmured vowels sets relatively narrow limits to their duration.

The reason for the longer duration of murmur vowels in normal speech is thus to be sought in the historical origin of murmured vowels as a fusion between one or two vowels and /h/.

6.3. Fundamental Frequency.

6.3.1. According to P.B. Pandit's description (1), murmured vowels are characterized by a slight lowering of pitch. As the disappearance of /h/ before or after the vowels in other Indian languages has been accompanied by a development of tonal differences (6), it is of particular interest to investigate this feature.

6.3.2. As word tones are always modified by sentence intonation (and even isolated words or word series will have a definite sentence intonation), it is important to restrict the comparison to words found under the same rhythmical conditions. In RD's A-recording and in both RT's recordings the words are spoken with rising or rising-level intonation. In RD's T-recording some words were spoken in pairs, and in this case the first word had rising intonation, the second rising-falling, with dominating fall. In the words spoken in series he has rising intonation in one reading, and rising-falling in another. The words of his T-recording have therefore been distributed on two categories marked as / and ∧. GU has rising-falling intonation, PBP variable intonation (only the frequency at the start has been measured). SK's curves could not be measured because of hoarseness.

The sentence intonation varies in the same way for clear and murmured vowels, so that the general trend of the movement is the same, and differences must be looked for in smaller details.

6.3.3. The frequency was measured at some selected points of the frequency curve, namely the beginning and end of the vowel, as well as maxima and minima.It was measured with an exactitude of five cps. Moreover the distance from the beginning of the vowel to the frequency maximum was measured (in cs).

As it appears from Table 4, comprising the averages for each recording, the words with murmured vowel have a lower minimum than the words with clear vowel (for the words with rising-falling tone this minimum is the first one). Although the difference is very small, it is relatively constant, and is statistically significant for all recordings except RT (A). PvB's curves could not be measured because the base line of the calibration curve was uncertain, but she shows the same tendency to have a lower start of murmured vowels.

6.3.4. A significant difference has also been found for the distance from the beginning of the vowel to the frequency peak (Table 5), but when the difference is measured in percentage of the vowel, it is extremely small or non-existent.

6.3.5. Figs. 13 and 16 show the differences in minimum, maximum, rise and distance from beginning to maximum for the averages of each word pair for some of the subjects.

6.3.6. The graphical display of Figs. 13 and 16 is simplified in the sense that the minimum has been identified with the start of the vowel. In reality the vowel may start with a small fall, and this is particularly often the case in murmured vowels (especially for RD and GU). The form of the curve is therefore often more complex in murmured vowels () than in clear vowels (), (see Fig. 6).

6.3.7. The sentences show the same features, but with some greater variation.

6.4. Intensity.

6.4.1. As already mentioned (5.3.) there is no consistent difference in overall intensity between clear and murmured vowels. The averages for the different speakers and recordings are given in graphical display in Fig. 17a.

6.4.2. More consistency is found in the shape of the curve. All speakers have, on the average, a longer distance from the beginning of the vowel to the peak of the curve in murmured vowels than in clear vowels, but the difference is not significant in GU's and PvB's recordings. The general averages are given in Table 7, and the averages of the individual words are shown graphically in Fig. 18 and Fig. 19 (see also Fig. 6).

Table 4. FUNDAMENTAL FREQUENCY

Minimum (Beginning)(cps).

Speakers	RD(A)	RD(T)		RT(A)	RT(T)	GU(A)	GU(T)	PBP
Number of Words	552	$\overset{/}{89}$	$\overset{\wedge}{91}$	348	159	1oo	147	72
cps murm.	137	114	12o	133	133	131	11o	129
cps clear	148	119	126	134	136	137	117	14o
Difference	÷11***	÷5***	÷6***	÷1	÷3**	-6***	-7***	÷11***
Lowering in %	8	4	4	1	2	5	6	7

Maximum (cps).

Difference	o	o	o	+2	+3	-4	-4
Rise							
cps murm.	36	21	12	18	38	11	13
cps clear	25	16	6	16	32	9	1o
Difference	11***	5***	6***	2**	6***	2***	3***

TABLE 5. FUNDAMENTAL FREQUENCY

Distance beginning-maximum (cs)

Monosyllables

Speaker	RD(A)	RD(T)		RT(A)	RT(T)	GU(A)	GU(T)
		/	Λ				
Number of Words	336	66	66	2o6	72	78	96
cs murm.	19.8	23.4	12.o	18.7	2o.4	19.6	16.5
clear	19.2	2o.3	8.7	16.9	18.o	15.6	12.2
Difference	o.6	3.1	3.3	1.8	2.4	4.o	4.3
% of duration murm.	68	85	42	85	87	6o	68
clear	68	84	36	88	85	52	57
Difference	o	1	6	÷3	2	8	11

Dissyllables

Number of Words	12o	24	24	143	67	22	32
cs murm.	16.8	16.5	16.6	12.7	17.3	14.8	13.7
clear	14.2	13.9	13.7	9.5	12.4	13.2	12.2
Difference	2.6	2.6	3.o	3.2	4.9	1.6	1.6
% of duration murm.	94	95	1oo	88	1oo	97	1oo
clear	92	1oo	1oo	78	1oo	92	1oo
Difference	2	-5	o	1o	o	5	o

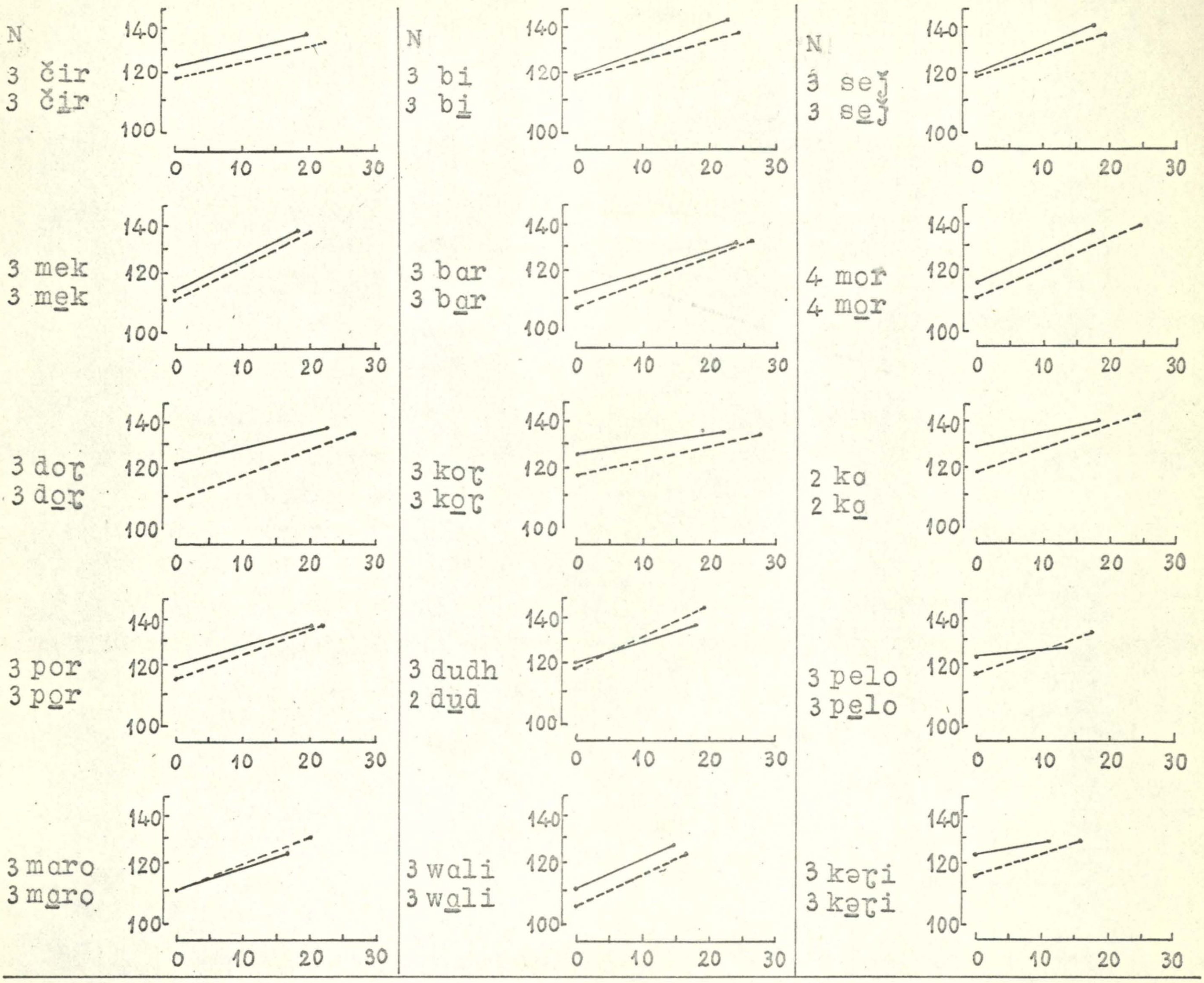
Monosyllables + Dissyllables

Number of Words	456	9o	9o	348	139	1oo	148
cs murm.	19.o	21.5	13.3	16.8	18.8	18.2	15.4
clear	17.9	18.6	1o.1	14.5	15.2	15.o	12.2
Difference	1.1***	2.9***	3.2***	2.3***	3.6***	3.1***	3.2***
% of duration murm.	72	87	52	84	95	76	76
clear	72	87	46	83	93	65	68
Difference	o	o	6	1	2	11	8

Fig. 13. FUNDAMENTAL FREQUENCY

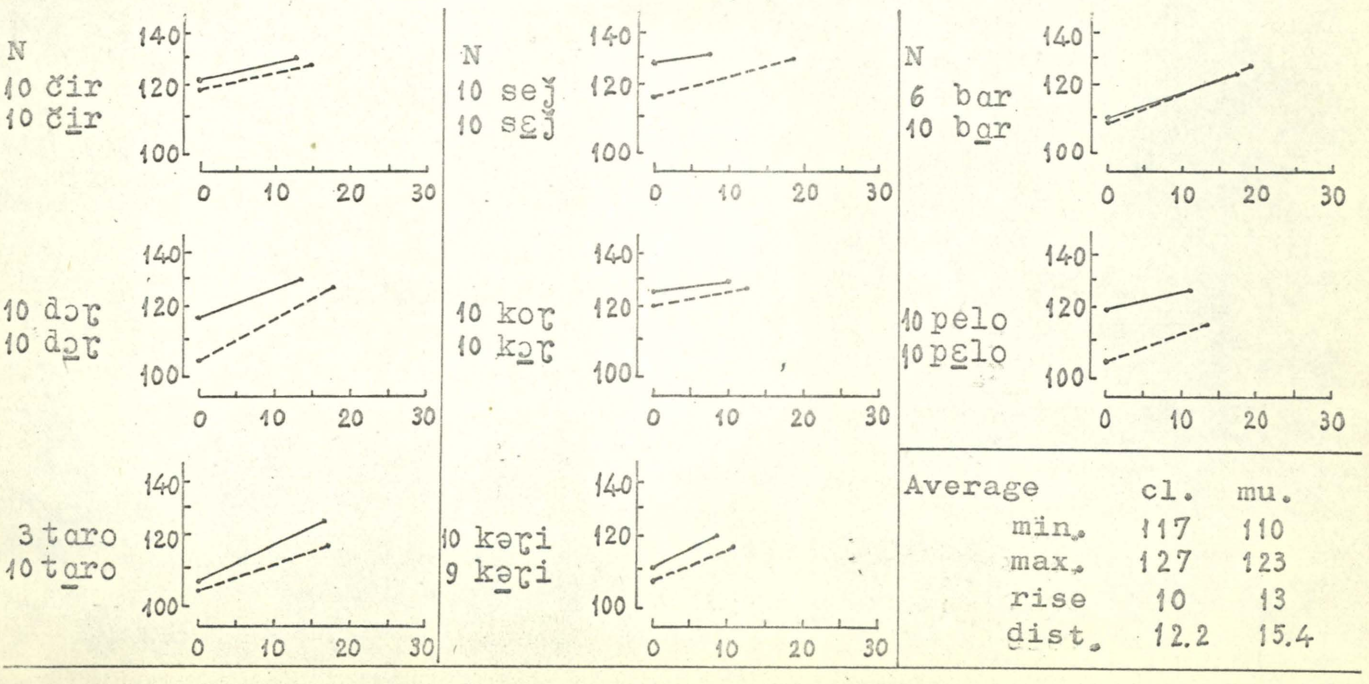
MINIMUM (cps), MAXIMUM (cps), DISTANCE MIN.-MAX. (cs)

RD Text 1 (T) (/) — =clear ---- =murmured



Average		min.	max.	rise	dist.
	cl.	119	135	16	18.6
	mu.	114	135	21	21.5

GU Text 3 (T)

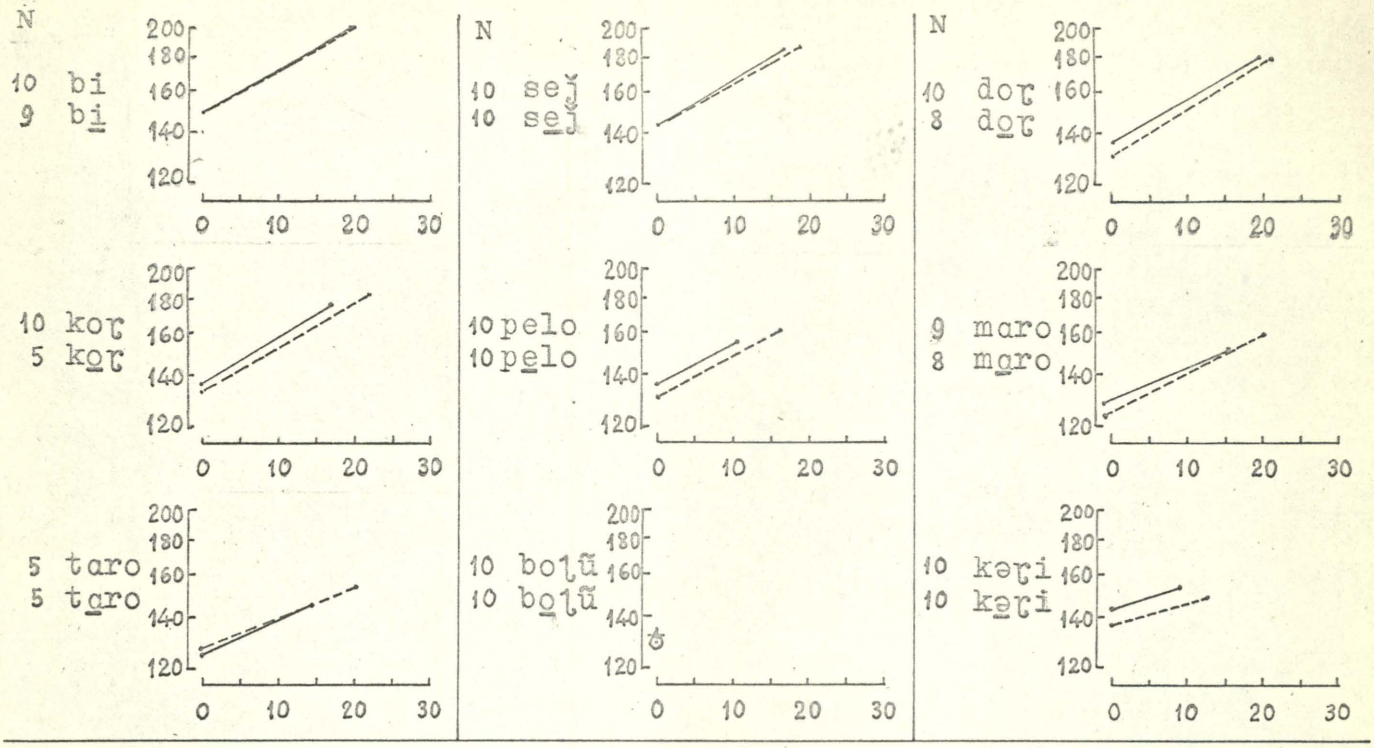


Average	cl.	mu.
min.	117	110
max.	127	123
rise	10	13
dist.	12.2	15.4

Fig. 16. FUNDAMENTAL FREQUENCY

MINIMUM (cps), MAXIMUM (cps), DISTANCE MIN.-MAX. (cs)

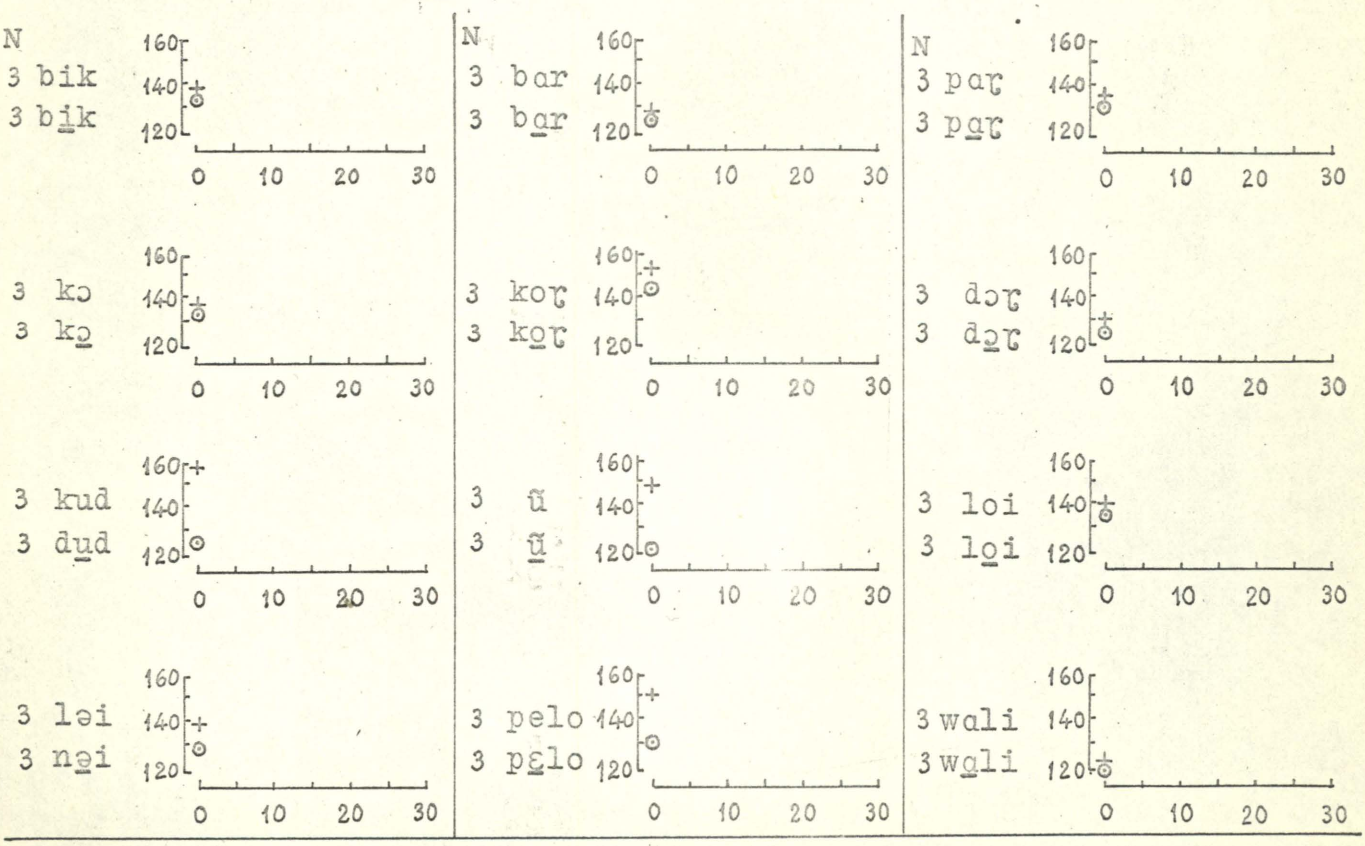
RT Text 2 (T)



Average

	min.	max.	rise	dist.
cl.	136	168	32	15.2
mu.	133	171	38	18.8

PBP Text 4



Average

	min.
cl.	140
mu.	129

Fig. 17. INTENSITY

ATTENUATION IN HIGHPASS FILTERED CURVE (⊙ = murmured + = clear)

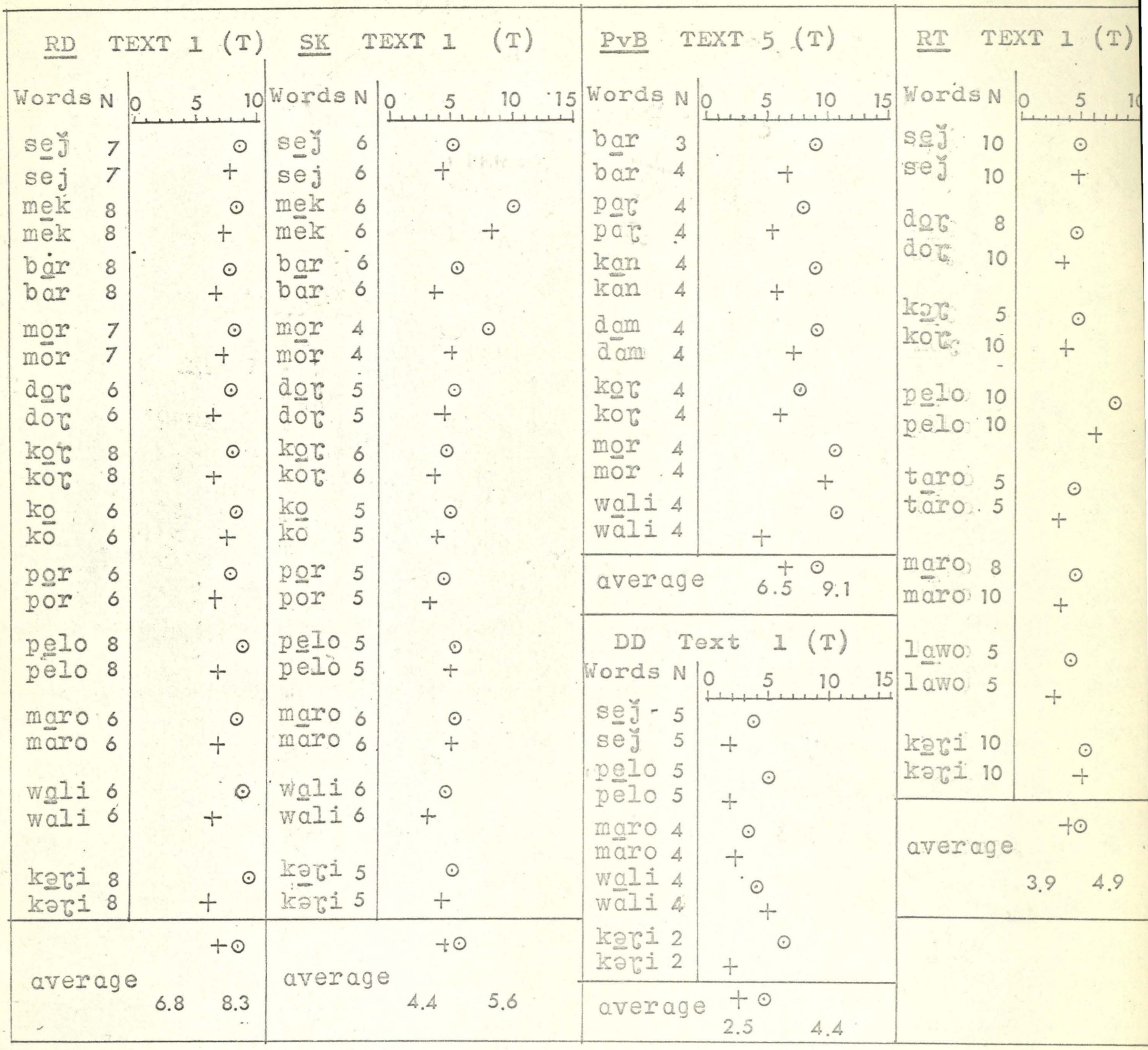


Fig. 17 a. AVERAGE INTENSITY

DIFFERENCE BETWEEN MURMURED AND CLEAR VOWELS

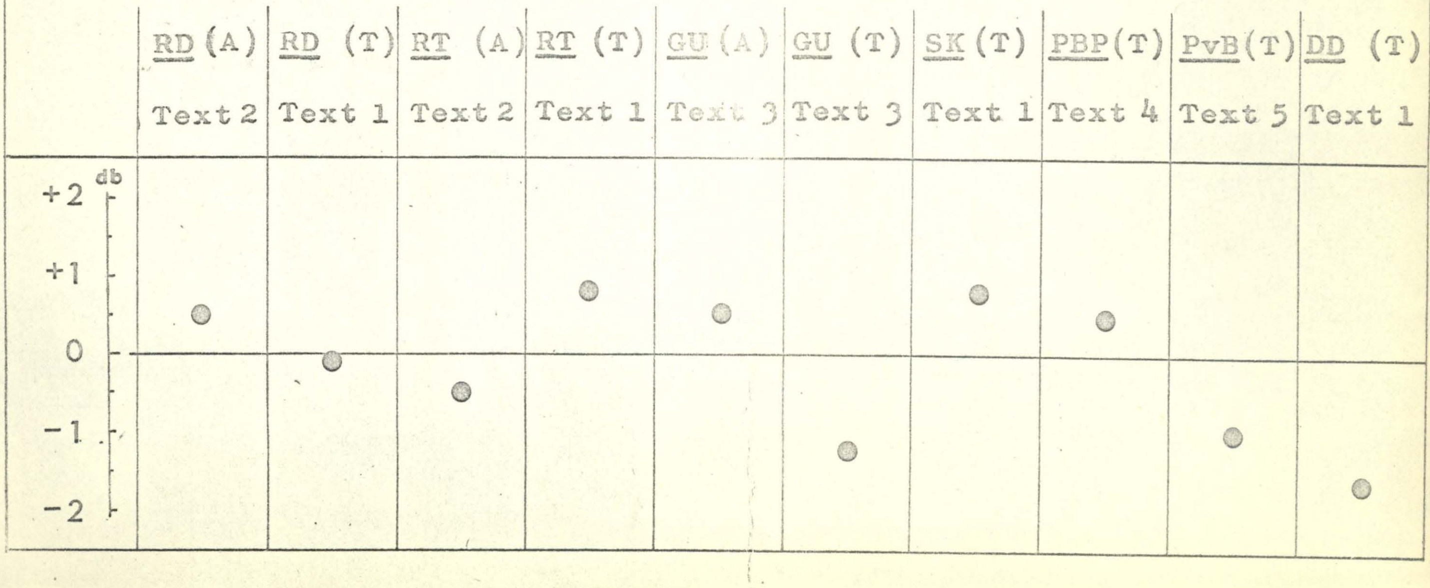


Fig. 18. INTENSITY

DISTANCE FROM BEGINNING TO MAXIMUM (CS) (⊙ = murmured + = clear)

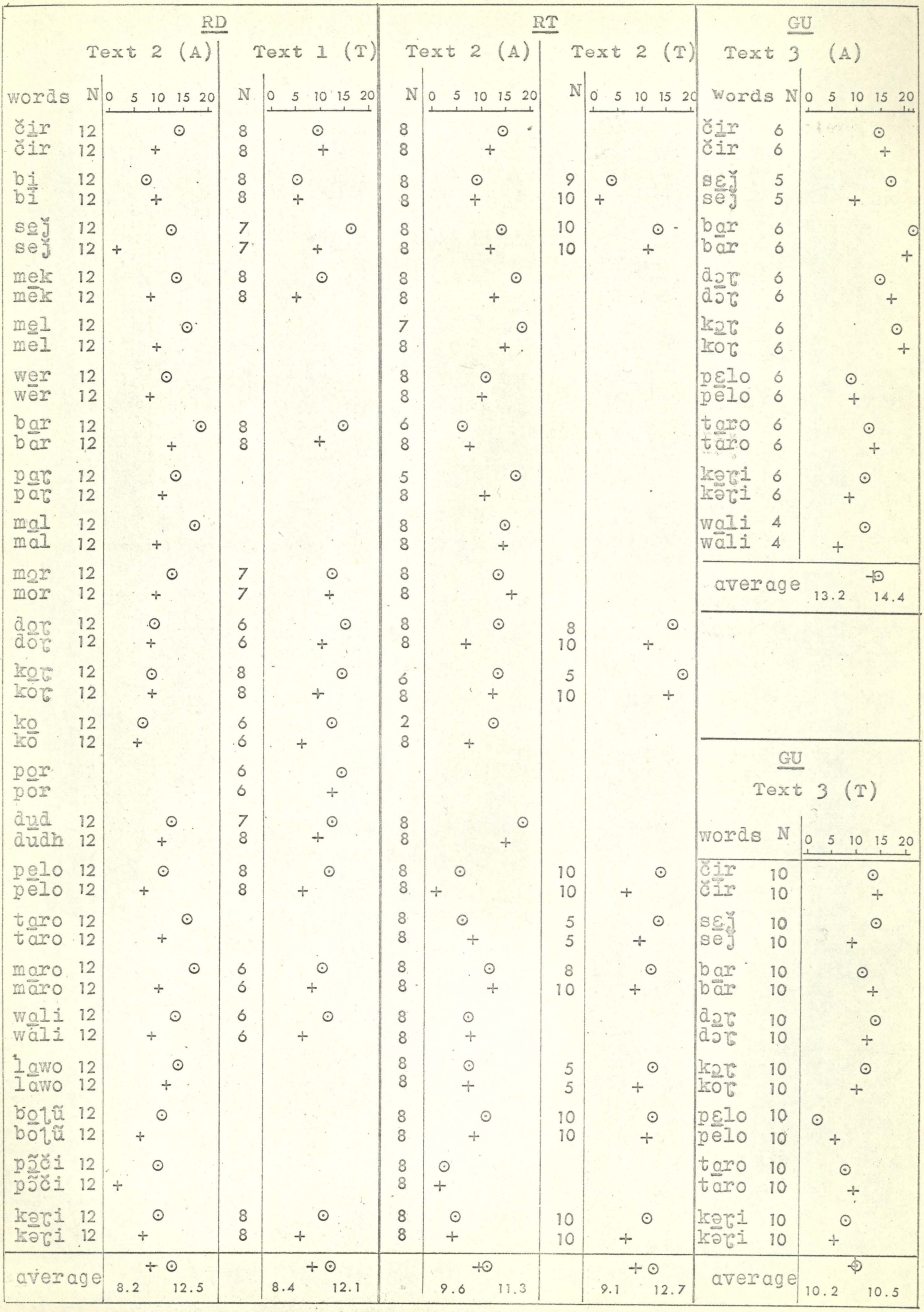


Fig. 19. INTENSITY

DISTANCE FROM BEGINNING TO MAXIMUM (CS) (⊙ = murmured + = clear)

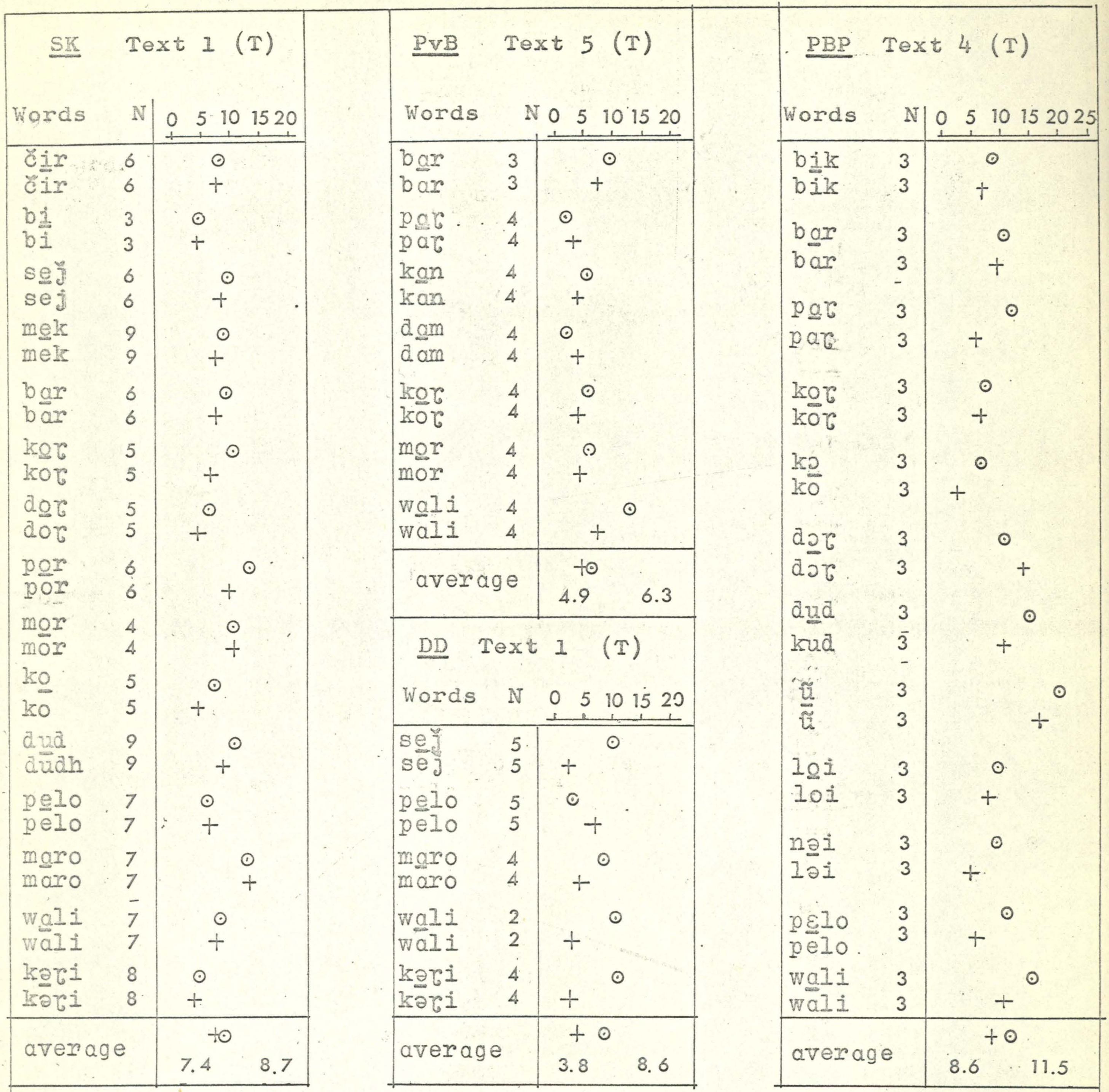


Table 7. INTENSITY

Distance from beginning to maximum (cs).

	RD(A)	RD(T)	RT(A)	RT(T)	GU(A)	GU(T)	SK	PBP	PvB	DD
Number of words	528	215	338	70	102	80	186	72	54	40
murm.	12.5	12.1	11.3	12.7	14.4	10.5	8.7	11.5	6.3	8.6
clear	8.2	8.4	9.6	9.1	13.2	10.2	7.4	8.6	4.9	3.8
diff.	4.3***	3.7***	1.7***	3.6***	1.2	0.3	1.3*	2.8**	1.4	4.8**

The words in sentences show dimilar differences. Both murmured and clear vowels have a somewhat longer distance to the intensity peak when the tone is rising than when it is rising-falling.

6.4.3. Some of the subjects (RD, PBP, partly RT) have not only a longer distance to the peak, but also a lower start and a greater rise of the intensity curve (see Fig. 6 [taro] and [paɽ]).

6.5 Spectral Structure.

6.5.1. General Remarks.

The analysis of spectral structure is based on a restricted number of words (324 in all) comprising the vowels [a a̱ i i̱ e e̱ ɛ̱ o o̱ ɔ ɔ̱]. The only example of [u -u̱] (dudh-du̱d) was left out as somewhat dubious, and the examples with ə-ə̱ had very short vowels and were difficult to measure.

The results are given in graphical form in Figs. 20-24. In these charts frequency is given horizontally and amplitude vertically. The scales are logarithmic. Besides the formants (F_1, F_2, F_3, and F_4) the fundamental (F_o) is given and also the second and third harmonics (H_2 and H_3) in the cases where they do not enter into F_1. Two separate peaks in the region of F_4 have been indicated as F_{4a} and F_{4b}. N indicates the number of word pairs, (in contradistinction to the other graphs, where it indicates the number of single words). The transcription with [ɛ] and [ɔ] is only used where the speaker distinguishes between the phonemes /e/, /o/ and the phonemes /ɛ/, /ɔ/. Thus /ɔ/ need not indicate a more open vowel than [o]. Close and open e and o-sounds are treated together in the following sections as [e] and [o]. Words containing the same vowels have been combined in one figure, if they did not show any appreciable difference in formant structure.

Fig. 2o. SPECTRAL STRUCTURE

Frequency and Amplitude of the Fundamental (F_o), the Second and Third Harmonics (H_2 and H_3), and F_1 - F_4

Frequency ⟶ (cps), Amplitude ↑ (db). a - a.

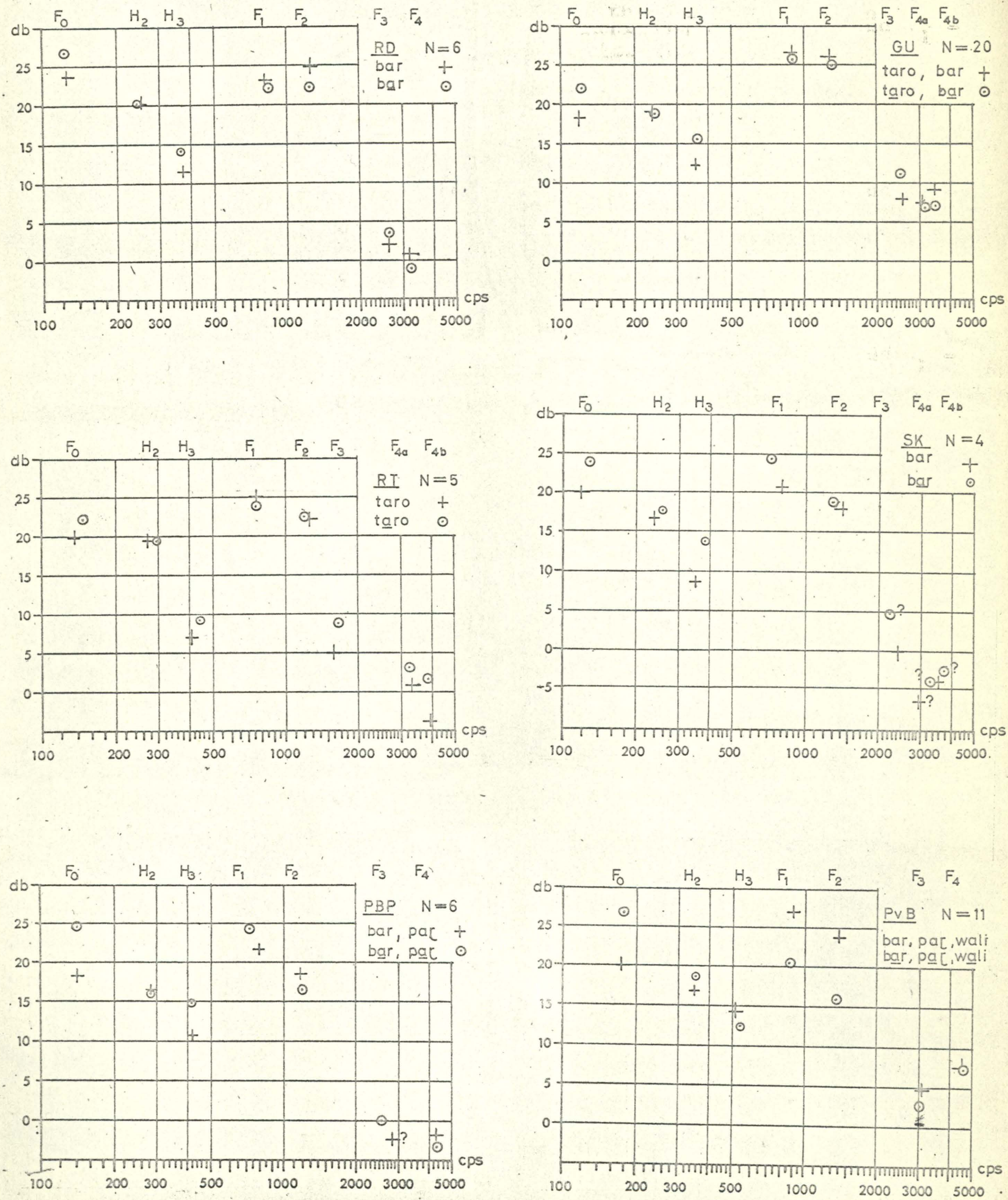


Fig. 21. SPECTRAL STRUCTURE

Frequency and Amplitude of the Fundamental (F_o) and F_1 - F_4

Frequency → (cps), ↑ Amplitude (db). i - i̲ .

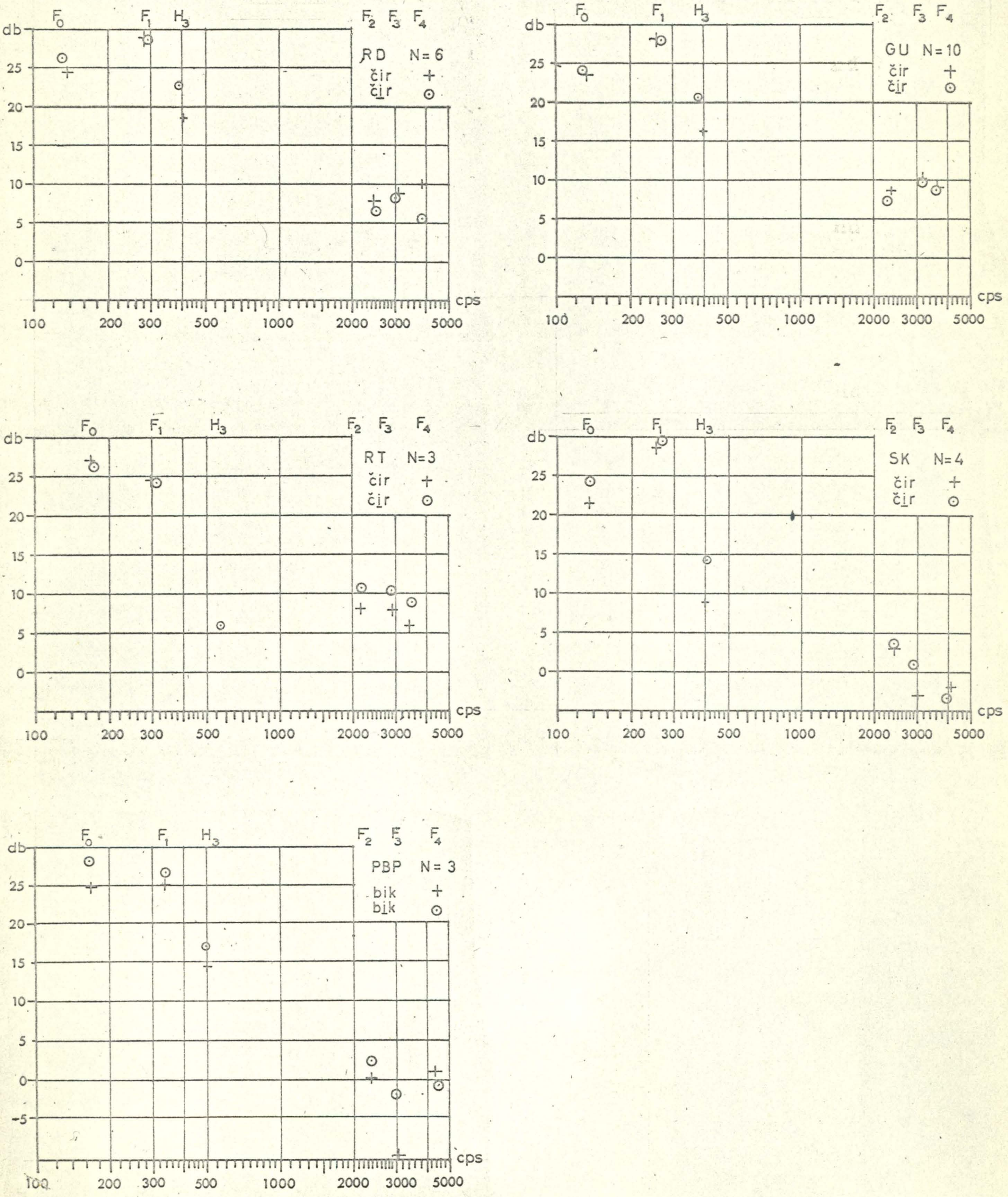


Fig. 22. SPECTRAL STRUCTURE

Frequency and Amplitude of the Fundamental (F_o), the Second Harmonic (H_2), and F_1 - F_4

Frequency ⟶ (cps), ↑ Amplitude (db) . e - e̲.

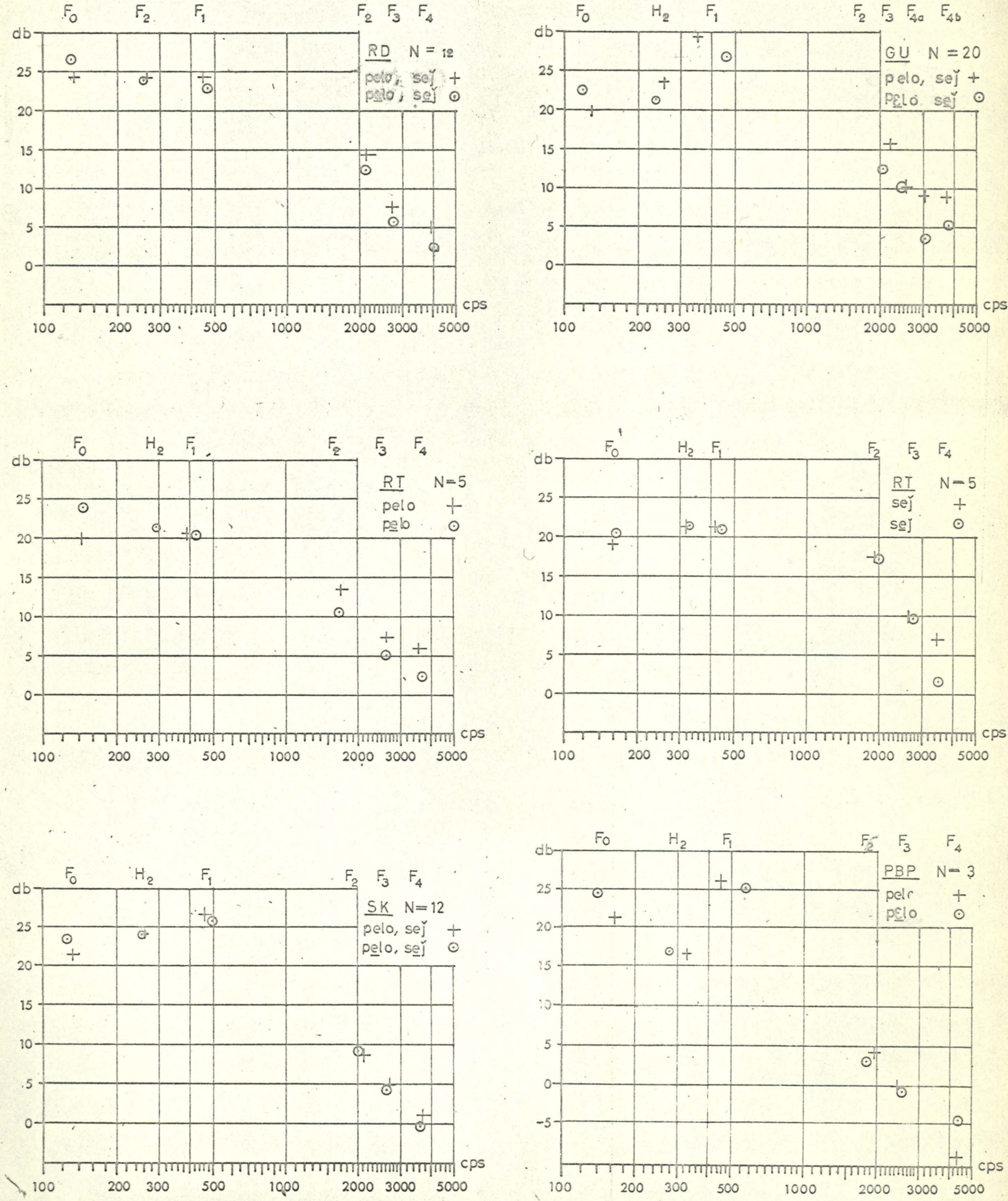


Fig. 23. SPECTRAL STRUCTURE

Frequency and Amplitude of the Fundamental (F_o), the Second Harmonic (H_2), and F_1 - F_2 (- F_4)

Frequency ⟶ (cps), ↑ Amplitude (db). o -o̲ .

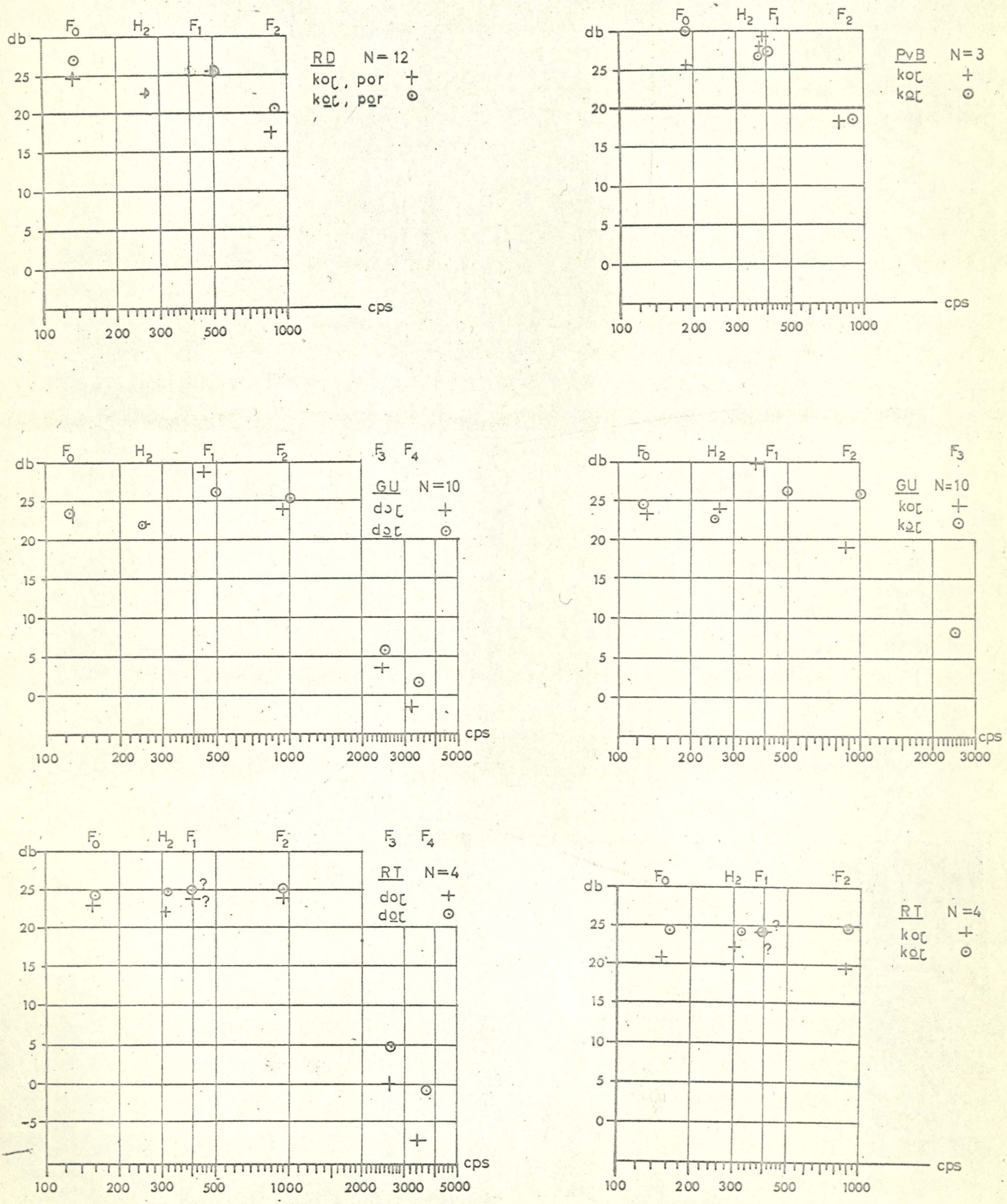
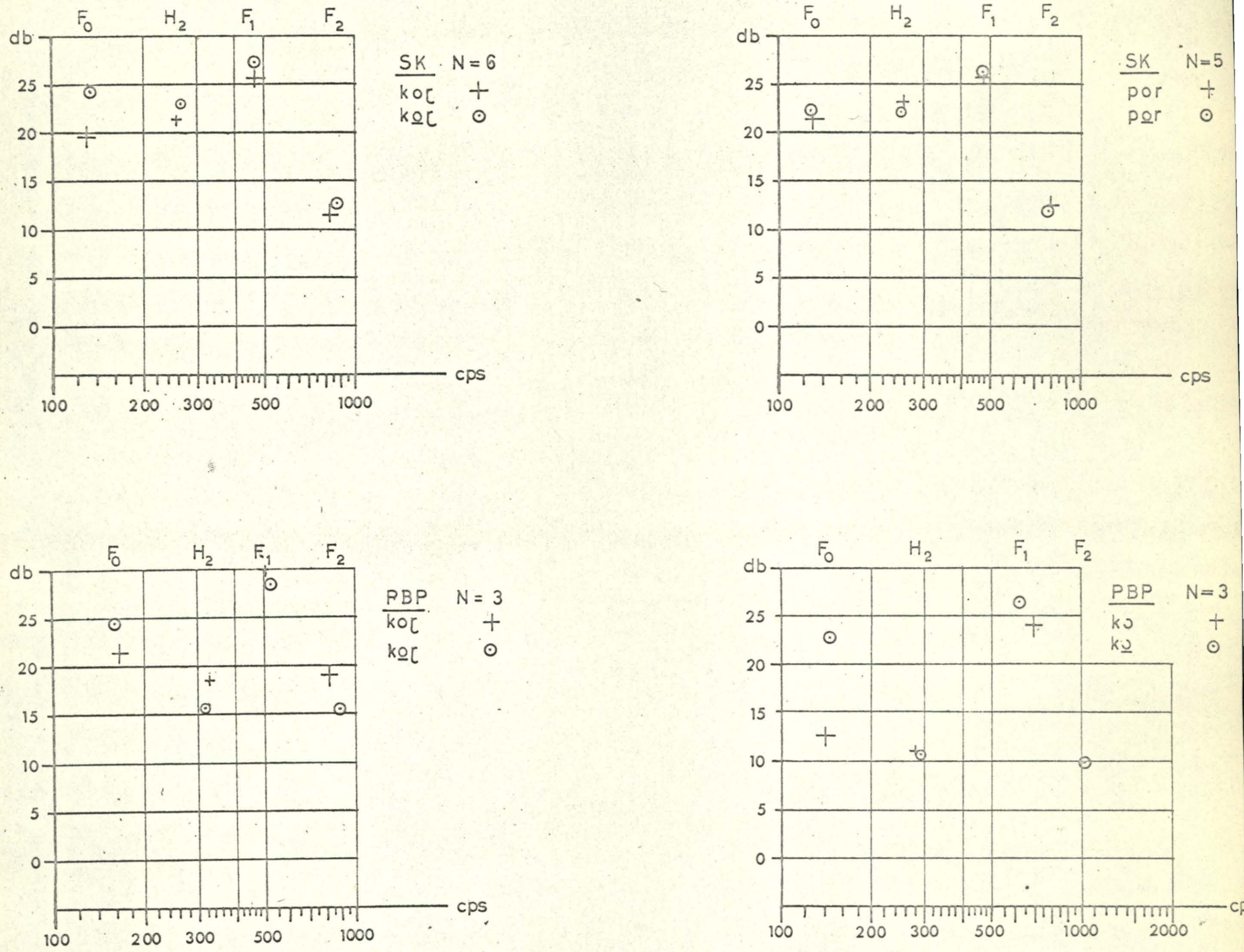


Fig. 24. SPECTRAL STRUCTURE

Frequency and Amplitude of the Fundamental (F_o), the Second Harmonic (H_2), and F_1 - F_2

Frequency ⟶ (cps), ↑ Amplitude (db). o - $\underline{o}$.



6.5.2. Formant Frequency.

Dave did not find any consistent differences between the formant frequencies of clear and murmured vowels (2). My measurements have confirmed this result. The only difference common to all speakers is a somewhat higher F1 in [e̲] than in [e]. Four of the six speakers have also a higher F1 in [o̲], four have a lower F2 in [e̲] and five a higher F2 in [o̲]. This means that there is a tendency for [e̲] and [o̲] to be somewhat more open than clear [e] and [o]. The tendency to a shift up of F1 is what should be expected as a result of a larger coupling to the trachea (7), but it must be noted that the tendency is not found in [a-a̲], that it is very slight in [i-i̲] and also very slight for RD, SK and RT in [e-e̲] and [o-o̲]. The speakers GU, PBP, and PvB have a much more pronounced difference, but these speakers distinguish between close and open /e/ and /o/, and in most cases it would be more correct to say that they have the phonemes/e/ and /o/ in words with clear vowel and the phonemes /ɛ/ and /ɔ/ in words with murmured vowel. The examples /dɔɽ/ (GU) and /ko̲ɽ/ (PBP) show that the degree of opening is not a mechanical consequence of the murmur. In the cases where the difference is pronounced it should thus not be considered as a synchronic phonetic difference, accompanying the difference between clear and murmured vowels, but as a result of a diachronic development. It is, however, interesting that a former intervocalic /h/ (which is the historical source of the murmur) has also been considered as one of the sourcesof the open /ɛ/, e.g. in /pɛ̲lo/ (8).

6.5.3. Distribution of Spectral Energy.

The most obvious spectral characteristic of murmured vowels is the relatively high level of the fundamental compared to F_1-F_4.
6.5.3.a. A rough quantitative picture of this difference can be obtained from a comparison between the highpass filtered intensity curve, used for delimitation of the mingograms, and the intensity curve without highpass filtering.

By subtracting the attenuation caused by highpass filtering of oral vowels from that of murmured vowels, one gets a measure of the greater predominance of the lowest harmonics in the latter. Table 8 gives the general averages, and Fig. 17 the averages of individual word pairs in graphical form.

TABLE 8. REDUCTION IN LEVEL BY HIGHPASS FILTERING

Difference between murmured and clear vowels. (db).

Speakers	RD(T)	RT(T)	SK	PvB	DD
Number of words	174	131	130	55	40
Difference murm.-clear	1.5***	1.0***	1.2***	2.6***	1.9***

Only the T-recordings, which comprise the two types of intensity curves for the same words, have been utilized. The close vowels /i-i/ and /u-u/ have been left out, because in these vowels most of the first formant was removed too. GU's and PBP's highpass filtered curves could not be used because of a technical mistake.

6.5.3.b. The spectrograms have been analysed in more detail. The results are given in graphical form in Fig. 20-24 (vertical dimension), and Table 9 gives the differences between murmured and clear vowels as regards the absolute level of F_o (L_o) and the relative level of F_o compared to the levels of the second harmonic and the level of F_1-F_4, i.e. L_o-$L(H_2)$ and L_o-L_1, L_o-L_2, L_o-L_3,and L_o-L_4.

It appears from Table 9 and from the graphs Figs. 20-24 (see also 1-3) that the absolute level of F_o is higher in murmured than in clear vowels, and that L_o-L_1, L_o-L_2, and L_o-L_4 is higher in the murmured vowels [a, e] and [i] than in the corresponding clear vowels, whereas for [o] only L_o-L_1 is higher. L_o-L_3 is variable except for [e]. Some older spectrograms of [u-u] show the same tendency as found here for [o-o].

The relative drop of intensity above F_o is evident already in the second harmonic; L^o-$L(H_2)$ is higher in murmured vowels than in clear vowels. The spectra of [a-a] and [i-i] show that this is not true of the third harmonic (in e- and o-sounds the problem is more complicated because H_3 is part of F_1).

This fact may be interpreted in different ways. (i) One might imagine a larynx spectrum with an increase of intensity in the odd harmonics. This assumption, however, was not corroborated by an inspection of the sections. (ii) It might be due to a broadening of the first formant of i and of the subformant of a, and the increase of F_o might be part of this broadening. But a number of extra recordings of bar-bar spoken by RD and GU, partly with rising-falling fundamental, partly on a very low fundamental showed that at low

Table 9. SPECTRAL AMPLITUDE.

Difference between murmured and clear vowels.

(Murmured - clear)(dB)

[a̲-a]

Speakers	RD	RT	GU	SK	PBP	PvB
Words	ba̲r/bar	ta̲ro/taro	ba̲r/bar ta̲ro/taro	ba̲r/bar	ba̲r/bar pa̲ɽ/paɽ	ba̲r/bar pa̲ɽ/paɽ wa̲li/wali
N	12	10	40	8	12	22
L_o	3.0	2.5	3.7	4.1	6.2	6.7
$L_o-L(H_2)$	3.2	2.6	3.9	3.1	6.7	5.0
L_o-L_1	4.1	3.8	4.6	0.3	3.6	13.7
L_o-L_2	5.7	2.3	4.9	3.4	8.4	14.6
L_o-L_3	1.5	-1.3	0.5	-0.5	4.0	8.6
L_o-L_4	4.9	0.2	5.9	1.4	7.4	3.3

[i̲-i]

Speakers	RD	RT	GU	SK	PBP
Words	či̲r/cir	či̲r/cir	či̲r/cir	či̲r/cir	bi̲k/bik
N	12	6	20	8	6
L_o	1.9	-0.9	0.6	2.8	3.5
$L_o-L(H_2)$	3.4	-0.9	0.9	2.4	1.7
L_o-L_1	2.0	-0.5	0.4	1.9	1.8
L_o-L_2	3.1	-3.6	1.9	2.2	1.3
L_o-L_3	2.5	-3.4	1.0	-1.2	-4.3
L_o-L_4	6.3	-3.8	1.3	4.3	5.3

[e̲-e]

Speakers	RD	RT		GU	SK	PBP
Words	se̲ǰ/seǰ pe̲lo/pelo	se̲ǰ/seǰ	pe̲lo/pelo	se̲ǰ/seǰ pe̲lo/pelo	se̲ǰ/seǰ pe̲lo/pelo	pɛ̲lo/pelo
N	24	6	6	40	24	6
L_o	2.1	1.4	3.7	2.4	1.9	3.0
$L_o-L(H_2)$	2.4	1.2	3.5	4.8	1.8	3.0
L_o-L_1	3.5	1.6	3.9	5.1	2.7	3.9
L_o-L_2	3.9	1.7	6.8	5.7	1.4	4.1
L_o-L_3	3.3	1.9	5.9	2.6	2.4	3.7
L_o-L_4	4.7	6.8	7.4	8.1	3.5	-1.5

[o̲-o]

Speakers	RD	RT		GU		SK		PBP		PvB
Words	ko̲ɽ/koɽ po̲r/por	ko̲ɽ/ koɽ	do̲ɽ/ doɽ	kɔ̲ɽ/ koɽ	dɔ̲ɽ/ doɽ	ko̲ɽ/ koɽ	po̲r/ por	ko̲ɽ/ koɽ	kɔ̲/ kɔ	ko̲ɽ/ koɽ
N	24	8	8	20	20	12	10	6	6	8
L_o	2.4	3.4	1.3	1.1	0.2	4.6	0.9	3.0	10.0	4.3
$L_o-L(H_2)$	2.3	1.5	-1.2	2.5	0.5	3.0	2.0	6.0	10.3	5.5
L_o-L_1	2.5	3.2	0.2	4.7	2.8	3.0	0.3	4.5	7.5	5.5
L_o-L_2	-0.5	-1.8	0.1	-5.7	-1.1	3.5	1.5	7.5	?	4.1
L_o-L_3			-3.3	-8.8	-2.0					
L_o-L_4			-5.0	-8.7	-2.9					

fundamental frequencies (100-110 cps) there is a valley between the fundamental and the subformant at 250-300 cps, and the difference between clear and murmured vowels lies, in these cases, in the fundamental.

(iii). It is thus more probable that the increase in the level of F_o is due to a change in the voice source spectrum, caused by a relaxation of the glottis, and that the increase in the level of H_3 compared to H_2 is due to an increase in the frequency of the subformant in[a̲] and of F_1 in[i̲](which is thus probably slightly higher than indicated in the table and graphs). This would also fit with the assumption of a rise in F_1 in breathy vowels (cp. 6.5.1)

It has only been possible to apply significance tests to a restricted part of the material because of the small number of pairs. But the differences have been found to be significant for the absolute level of F_o (L_o), for L_o-L(H_2), and L_o-L_1 for o̲-o and [e̲-e](GU and RD) and for[a̲-a](GU and PvB), and moreover for the difference L_o-L_2 and L_o-L_4 for [e̲-e](GU and RD) and [a̲-a](GU and PvB).

6.5.4. Other Spectral Characteristics.

(a) It would be natural to expect more noise in murmured vowels because of the stronger airstream. Pandit has also found some random distribution of energy particularly at higher frequencies in [pɔ̲r] compared to [pɔr] (1,p. 172). In the present material noise at higher frequencies was clearly seen in PvB's murmured vowels but in spite of the fact that special curves were taken with this purpose it was rarely found in PBP's and RD's vowels and hardly ever in those of the other informants.

(b) It has been assumed that breathy vowels should have broader formants (3). A certain broadening can sometimes be seen in the sections, but by no means always. The oscillograms taken with 1/4 speed of the tape (4.4) show more damping of the single periods in [a̲] than in [a] for four of the informants, but very little for other vowels.

(c) The oscillograms show a tendency to asymmetry in the murmured vowels of some informants, the amplitude of the first positive deflection of each period being stronger than the corresponding negative deflection. This may have something to do with the damping or perhaps, according to a suggestion made by Gunnar Fant, to a slower closing of the vocal cords.

(d) The fundamental frequency curve is often more smooth for murmured vowels (see Fig. 6). This is particularly evident for informants with hoarse voice, and is probably due to a more relaxed glottis.

(e) A later start or a momentary weakening of higher formants (F_3 and F_4, and sometimes F_2) is often found in RD's murmured vowels (see Fig. 1), and sometimes in the vowels of other speakers.

(f) A less sharp delimitation between vowel and [r ɽ l ɭ] pointing to a less precise articulation is a relatively frequent feature (see Fig. 6 taro-taro).

7. COMPARISON BETWEEN MURMURED VOWEL AND [h].

As murmured vowels have come into existence through a fusion of vowel and [h], a comparison between murmur and [h] may be of interest. Voiced [h] is found in Gujarati after voiced stops. Figs. 8a and 8a contain mingograms and spectrograms of [bh-]. Fig. 8a shows that [h] has a very strong F_o, weak higher formants and some noise at higher frequencies. Figs. 7a and b contain curves of intervocalic [h] spoken by RT instead of murmured vowels, the spectrogram (7b) shows the same strong intensity of F_o and weakness of higher formants as 8a, Fig. 7a an increase of airflow combined with a decrease of intensity and of fundamental frequency compared to the surrounding vowels. (The [h] of [kəho] is, in all respects, stronger than that of [pəhaɽ]). The strong airflow, low frequency, and the relatively strong intensity of F_o have all been found as characteristics of murmured vowels. The cases with a drop of frequency and low intensity in the beginning often found in RD's curves of murmured vowels are signs of an incomplete fusion of [h] with the vowel, so that the murmur-element is stronger in the beginning (see Fig. 6).

Ilse Lehiste (9) has measured the airflow of voiceless [h] and found it to be very strong (this has also been found for /h/ in other languages, it is e.g. obvious in Danish for both voiced and voiceless [h]). Moreover she found a pronounced weakening of F_1, and a rise in frequency, in the cases where F_1 was visible (this has also been seen in spectrograms of Danish [h]).

The difference between murmured and clear vowels in Gujarati is neutralized after aspirated consonants. The vowel found in this position is considered as clear, but curves of vowels preceded by aspirated consonants spoken by PBP, PvB, and RD show a certain assimilation of the beginning of the vowel to [h] : the fundamental is stronger, the airflow stronger, and there may be some noise at higher frequencies (the latter phenomenon is particularly obvious in PvB's curves). - For the stronger fundamental see Fig. 8b.

8. THE RELATIVE IMPORTANCE OF THE ACOUSTIC CUES.

8.1 Stability.

As one of the possible criteria of stability we have chosen the percentage of individual word pairs characterized by the difference in question; as pairs are considered the words of the same reading standing under the same conditions. The percentages are listed in Table 1o.

Table 1o. STABILITY

Percentage of word pairs characterized
by the different acoustic cues.

Speaker	RD(A)	RD(T)		RT(A)	RT(T)	GU(A)	GU(T)	SK	PBP	PvB	DD
1. Duration	68	9o		86	85	75	8o	74	63	77	(67)
2. Frequency		/	^								
a Minimum	89	71	67	35	54	81	76		71	48	
b Rise	83	69	71	51	72	48	56		73	67	
c Distance to peak	66	77	84	81	85	79	76				
Intensity											
Distance to peak	74	72		6o	81	75	63	66	66	55	(67)
Distribution of spectral energy											
L_o		92			77		79	78	94	1oo	
$L_o - L(H_2)$		89			64		83	74	89	93	
$L_o - L_1$		94			69		9o	78	78	1oo	
$L_o - L_2$		67			55		69	74	83	1oo	
		(87)			(57)		(91)	(69)	(75)	(1oo)	
$L_o - L_4$		91			65		82	78	83	91	

The number of words on which the percentages are based can be seen in tables 2-9. The numbers in parentheses under L_o - L_2 for RD and GU are the percentages for [a-a̲, e-e̲, i-i̲] without [o-o̲]). The percentages above 7o have been underlined. The percentage for L_o - L_4 is based on [a-a̲, e-e̲, i-i̲] alone.

Besides the cues listed RD and RT have a relatively high stability for the extent of the rise of the intensity curve, the rise being more extensive in murmured vowels. The percentages are: RD(A) 77, RD(T) 81, RT(T) 71, but RT(A) 41.

The sentences have somewhat higher percentages, particularly for RD(A) and RT(A).

8.2. Independency.

The question must be raised whether the acoustic cues found are mutually independent. If not, it is problematic to consider them as separate cues. However, no constant correlations have been found. It is only necessary to make some reservations for the distance to the frequency peak. It is in principle not dependent on duration, but the size and significance of the difference may be partly due to differences of duration, because the duration simply sets a limit to the place of the peak. This is certainly true of most of the dissyllables and also of the monosyllables with rising intonation in RD's T-recording and in both of RT's recordings.

8.3 Influence on Perception.

It has only been possible to investigate this very important aspect of the question in a very preliminary and insufficient way, in the first place because our synthesizer is under construction, and its possibilities are restricted, and secondly because we had only two Gujarati listeners (RD and RT) at our disposal.

8.3.1 Synthesis. The present state of our synthesizer allowed us to synthesize isolated vowels containing up to five formants which could be varied independently in frequency, bandwidth and amplitude, and a low frequency channel which permitted the variation of the amplitude of the fundamental and of the second harmonic. The vowels were composed of four normal formants and a subformant below F_1. The frequency contour of the fundamental could be given a fixed inflection (either level or slightly rising) or it could be produced manually, but of course not with precision. - Rischel succeeded without difficulty in synthesizing a clear [a] matching the spectral characteristics of Dave's [a] closely, which Dave recognized as a good Gujarati [a]. It sounded even very much like Dave's voice. But it proved impossible to produce a vowel which he would recognize as murmured, although it was possible to simulate all the amplitude relations found in the previous analysis. Only when these characteristics were exaggerated, he found that the vowel approached a murmured vowel very slightly. This seems to indicate that the specific spectral structure mentioned in 5.3.4 is not sufficient for identification. If F_1 was weakened too much, the vowel sounded nasal.

Addition of noise in higher formants helped a little, but the vowel was still not recognized as murmured. - It was also attempted to mix an [h] spoken by Dave (in [aha]) with the best of the synthesized vowels, but without result (noise and vowel did not fuse completely). The same was tried with a natural [a] , but with the same negative result. The synthetic vowels had slightly rising tone. Some rather primitive attempts at making a frequency or an intensity dip in the beginning did not give any improvement.

<u>8.3.2 Listening to unchanged words in random order.</u>

a. Each speaker had listened to his own tape recording and found it satisfactory. Moreover Dave had listened carefully to all recordings, and the words which he did not find quite murmured (this was a small number) or which he perceived as containing vowel + h + vowel (a good deal of RT's words) were discarded before the acoustic analysis. The material should thus be satisfactory. But the perceptual differences were small, even to a trained ear, and it was therefore decided to try whether the Gujarati speakers themselves could distinguish the words when they were given in random order. A test was prepared containing a number of words with different vowels from RD's, RT's, GU's, PBP's,and PvB's recordings. It was attempted to include words with different predominant cues. The test contained 45 different words spoken by RD, 36 by GU, 25 by RT, 16 by PBP, and 16 by PvB. The words spoken by different speakers were kept apart in separate groups. Each word was repeated twice with one second's interval and there was a pause of 5 seconds between the different words. A number was spoken by me before each word. The order of the words was quasi-randomized. Words belonging to the same word pair were never given in succession, and words with the same vowel were only rarely brought in succession. There was thus, generally, at least two other words between e.g. [bar] and [b̤a̤r]. RD (Dave) and RT acted as listeners. Dave was asked to indicate on a sheet whether the vowel of the word in question was murmured or clear, RT (who is not a linguist) was asked to identify the word as one of two words written in Gujarati orthography on the sheet after each number. He was asked to underline the word he heard.

b. Table 13 contains the number of correctly identified words.

Table 13. IDENTIFICATION OF UNCHANGED WORDS BY DIFFERENT SPEAKERS AND LISTENERS.

Speaker	RD			RT			GU			PBP			PvB		
	mu.	cl.	tot.	mu.	cl.	tot.	mu.	cl.	tot.	mu.	cl.	tot.	mu.	cl.	tot.
Number of words presented	25	20	45	14	11	25	21	15	36	8	8	16	9	7	16
Number correctly identified RD	24	18	42	6	9	15	12	14	26	4	8	12	8	7	15
Number correctly identified RT	17	16	33	8	10	18	14	11	25	6	8	14	6	7	13
Percentage correctly identified RD	96	90	93	43	82	60	57	93	72	50	100	75	89	100	94
Percentage correctly identified RT	68	80	75	57	91	72	67	73	69	75	100	88	67	100	81

The most astonishing result is the low percentage of murmured vowels identified. It is the more astonishing as the listening conditions should be good. High quality tape recorders and head phones were used, and the level was chosen so as to be convenient for the listeners, and RD knew all the speakers (except PvB) and had worked with his own recordings and those of RT. Moreover, the words not identified by RD and RT are often different, so that the percentage of murmured vowels identified by both is still lower. For the different speakers the percentage is: RD 65, RT 29, GU 33, PBP 25, PvB 67. One of the reasons is that the two listeners have difficulty with different vowels. In table 14 the results are grouped according to different vowels instead of different speakers.

Table 14. IDENTIFICATION OF UNCHANGED WORDS FOR DIFFERENT VOWELS AND LISTENERS.

Vowel		i̲	i	tot.	e̲	e	tot.	a̲	a	tot.	o̲	o	tot.
Number of words presented		18	12	30	14	13	27	25	19	44	20	17	37
Number correctly identified	RD	16	11	27	13	9	22	15	19	34	10	17	27
	RT	6	10	16	12	7	19	20	18	38	13	17	30
Percentage correctly identified	RD	89	92	90	93	69	81	60	100	77	50	100	73
	RT	33	84	53	86	54	73	80	95	86	65	100	81

It appears from the tables and graphs that clear vowels are generally much better identified than murmured vowels, i.e. they are rarely heard as murmured. The opposite difference in RD's perception of his own vowels is too small to be of any significance. The only real exception is that there is a tendency to perceive clear [e] as murmured (and murmured [e̲] is perceived correctly more often than other vowels.

Generally, however, it seems as if murmured vowels are conceived as the marked member of the opposition. One might have expected the opposite, at any rate in the case of RT, since many Gujarati speakers have free variation between clear and murmured vowel in e.g. /ba̲r/, but only clear vowel in /bar/, so that a perceived [ba̲r] can only be interpreted as /ba̲r/, whereas a perceived [bar] could be interpreted as both /bar/ and /ba̲r/. But both listeners seem to have identified a vowel as murmured only in the cases where they perceived some positive indication of murmur. This means probably that a significance test should not be based on the assumption that the two answers "clear" and "murmured" have the same probability. (No significance test has been applied to this material because of the relatively small number of answers.)

c. In view of the small number of words and listeners one should not draw too many conclusions from details of the test, and since the decision was often difficult, there are probably a good deal of answers which are due to chance.

In spite of this uncertainty a detailed comparison in respect of duration, fundamental frequency, intensity and spectrum of the words which were correctly identified with those which were misheard has given some results.

<u>Duration</u> seems to be of very little importance for the choice, but this does not exclude the possibility that duration may play a role in connected speech.

As to <u>fundamental frequency</u>, there is some evidence that words with purely falling tone (which also have a relatively high start) tend to be heard as clear. Only a few words with murmured vowel and falling tone have been identified correctly, and they have a particularly typical spectral structure. A low start in combination with a tone dip is in some cases seen to be more important than spectral structure and intensity.

As for the <u>distance</u> to intensity and frequency <u>peak</u> no conclusions can be drawn from the material.

The distribution of spectral energy on the other hand seems to be decisive in many cases, particularly when the spectral structure of the vowels of speaker and listener are compared.

In RD's perception of [a̲] there is only one case where the spectral structure seems to be irrelevant. Also RD's perception of [o̲] and [e̲] and RT's perception of [a̲, e̲] and [o̲] are influenced by spectral structure, but not to the same degree. RT has only identified six of eighteen examples of [i̲]. This may perhaps also be explained by spectral structure in the sense that he has himself no difference, and may therefore not be prepared to listen for this feature.

It may thus be concluded that in this listening test the most important cue seems to have been distribution of spectral energy, but that fundamental frequency is also of importance.

The listening test was carried out after the completion of the acoustic analysis. One might ask whether the stability of the cues would not have been better if only words identified correctly in the test had been used. The analysis of the mistakes show, however, that there would hardly have been much change, except for the distribution of spectral energy. Here the stability would probably have been raised, but not to a hundred percent.

<u>8.3.3.</u> <u>Listening to filtered vowels.</u>

a. Since the relative prominence of the fundamental had proved to be very constant in murmured vowels, it was found of interest to suppress it by filtering and to examine the result for perception.

The restricted text used for spectrographic analysis was used for this purpose; only SK was left out because of his hoarseness.

The texts were highpass filtered by means of a heterodyne filter (5) with a cutoff frequency of 230 cps (3 dB point), the words with [a-a̲] also with a cutoff frequency of 500 cps. **Also lowpass** filtering with a cutoff frequency of 3200 cps was undertaken. In a preliminary listening session Dave found the murmured vowels with highpass filtering slightly weakened, those with lowpass filtering unchanged except his own [o̲]. As highpass filtered murmured [a̲] was somewhat more difficult to perceive with a cutoff frequency of 500 cps than with a cutoff frequency of 230 cps, it was assumed that the subformant around 250-300 cps might be of some importance, and therefore a band stop filtering was made with the cutoff frequencies 200-500 (and for the speakers with low F_0 160-500) by which the subformant was weakened. But Dave did not find that the bandstop filtering had any influence. The loudness level seemed to be more important. Murmured vowels require a good deal of loudness for their identification.

A shorter test was then made with a selected number of words in random order. The test contained 91 highpass filtered and 101 band-stop filtered words spoken by RD, RT, GU, PBP, and PvB, and 17 lowpass filtered words spoken by RD. The recording and the listening conditions were as in the test with unchanged words. Only Dave acted as a listener.

There were 26 highpass filtered, 38 band-stop filtered and 4 lowpass filtered clear vowels. They were almost all heard correctly (96,95, and 100% respectively). -

The lowpass filtered murmured vowels were also heard correctly.

The results for the highpass filtered and band-stop filtered murmured vowels are given in Tables 15 and 16, grouped according to speakers and vowels respectively. Vowels heard correctly and incorrectly by Dave when unfiltered have been separated in the tables and designated by + and - respectively.

Table 15. IDENTIFICATION OF FILTERED MURMURED VOWELS GROUPED ACCORDING TO SPEAKER (LISTENER RD).

1. BS(160(200)-500)

Speaker	RD		RT		GU		PBP		PvB		tot.	
	unfilt.		unfilt.		unfilt.		unfilt.		unfilt.		unfilt.	
	+	-	+	-	+	-	+	-	+	÷	+	÷
Number presented	12	3	7	6	10	8	4	4	8	1	41	22
Number correctly identified	9	0	5	1	7	3	4	3	7	0	32	7
% correctly identified	75	0	71	17	70	38	100	75	88	0	78	32

2. HP 230

Speaker	RD		RT		GU		PBP		PvB		tot.	
Number presented	12	2	6	2	11	5	4	4	8	1	41	14
Number correctly identified	10	0	2	0	9	0	4	1	6	0	31	1
% correctly identified	83	0	33	0	82	0	100	25	75	0	76	7

Table 16. IDENTIFICATION OF FILTERED MURMURED VOWELS GROUPED ACCORDING TO VOWELS (LISTENER RD).

	1. BS (160(200)-500)								2. HP 230							
Vowels	i		e		a		o		i		e		a		o	
	unfilt.		unfilt.		unfilt.		unfilt.		unfilt.		unfilt.		unfilt.		unfilt.	
	+	-	+	-	+	÷	+	÷	+	-	+	÷	+	-	+	-
Number presented	12	2	9	1	12	10	8	9	12	0	8	0	12	6	9	8
Number correctly identified	11	0	8	0	11	4	2	3	11	0	3	0	12	1	5	0
% correctly identified	92	0	89	0	92	40	25	33	92	0	38	0	100	17	56	0

The high percentage of correctly identified vowels for the band-stop filtered recording is not astonishing. There was no reason to expect that the murmured character of the vowels should be damaged as long as the fundamental was preserved. Of the 41 vowels heard correctly, when unfiltered, 32 (or 78%) were heard correctly when band-stop filtered. Moreover of the 22 vowels heard incorrectly when unfiltered, 7 (or 32%) were heard correctly when filtered.

Most of the mistakes (six out of nine) concern the vowel [o̲]. This can be explained by the fast that in [o̲] the whole spectral difference lies in the amplitude relation between the fundamental and the second harmonic and (particularly) the first formant, and these are cut out (or weakened) by the filtering.

It is more astonishing that 76% of the highpass filtered murmured vowels have been heard correctly (the only difference from the band-stop filtered vowels is that one vowel has been improved). As mentioned in section 4, the fundamental has only been attenuated about 25 dB, but even this attenuation must reverse the relations between the level of the fundamental and the formants completely, and in the test with unchanged vowels this relation had been found to be of importance for perception. It should, however, be kept in mind in the first place that distribution of spectral energy is not the only cue, and in the second place that the difference in distribution of spectral energy is not limited to the relation between fundamental and higher formants. Almost all mistakes concern [e̲] and [o̲] , whereas [a̲] and [i̲] are heard correctly, and in the latter vowels there is also a clear difference in the relation between the level of F_3 and F_4 (and partly F_2) and a difference in the level of H_3, i.e. the subformant of [a̲] and the first formant of [i̲] are somewhat raised and broadened. Murmured [i̲] sounds more open or "lax" than clear [i], also when highpass filtered. Moreover six of the 12 words with [a̲] were spoken by PvB and had strong noise.

8.3.4. Perception of words with murmured vowels after tape-cutting.

As some of the acoustic cues found were concentrated to the beginning of the vowel, it might be of interest to examine the effect of removing the beginning or the end of the vowel. The investigation undertaken with this purpose is very restricted and purely preliminary, and it was only possible to find one listener at the time (Dave).

Two isolated examples of [a̲] spoken by RD, and 12 words spoken by RD, GU, and PvB were used. Cuts were made at steps of approximately three cs both from the beginning and from the end of the vowel until about lo-12 cs were left. The initial and final consonants were cut off at the first step from the beginning and end respectively. The cuts were made by hand and controlled by means of spectrograms. After each cut the words were played over to a second tape recorder. Dave listened a couple of times to the word series in the order recorded, i.e. with diminishing length, and tried to decide when the murmured character of the word became dubious, and when the vowel became clear. It was not so that he found it increasingly difficult to decide whether the vowel was clear or murmured with diminishing length; at a certain point the vowel was heard as clear. This, again, shows that murmur is perceived as a marked feature.

The results of the listening showed, in the first place, that a vowel must have a certain duration in order to be perceived as murmured. Below about 12 cs almost all vowels are heard as clear. This may partly be due to the accompanying reduction of loudness.

In the second place, it was obvious that cutting from the beginning is more detrimental than cutting from the end. When about 11 cs are cut off (from 5 to 16 cs for the different vowels) the vowel is heard as clear (with one exception), irrespectively of the length of the remaining part (which was from 11 to 45 cs). Eleven of the twelve words had also been represented in the test with unchanged vowels, and all identified as murmured, and seven of the words had been presented in the test with highpass filtering and all, except one, had been identified correctly in this test, too. This means that shortening of the vowel, particularly from the beginning, is more detrimental than removing the fundamental. And as an inspection of the spectrograms has shown that the specific spectral structure is kept throughout the vowel, it is obvious that the reasons for the importance of the beginning of the vowel must be sought in other types of cues. The cues concentrated in the beginning are low start of the tone movement, tone dip, and, sometimes, weak intensity. Almost all the vowels in this test had the small tone dip. Dave has this tone dip very often in his speech, and it seems to be of importance for his perception also. It was unfortunate that RT, who does not have this tone dip, and for whom it did not seem to play any role in the test with unchanged vowels, could not take part in this test.

One of the words was heard as clear as soon as the consonant was cut off. This word did not have the characteristic spectral distribution. -

One word was not influenced at all by the tape cutting, namely PvB's /pa̲ɽ/. This word had strong noise.

8.3.5. Moreover, a detailed examination of the answers to the test with unchanged vowels has shown that none of the acoustic cues are necessary, and none is alone sufficient. For all the cues cases can be found where a correct identification has taken place, although the cue was lacking, and for all the cues cases can be found where the cue is present in a strong degree, and nevertheless the vowel has not been identified as murmured. Generally a certain number of cues are combined, and the lack of one may be compensated for by a stronger degree of one of the others.

9. GENERAL CONCLUSIONS.

The phonetic investigation of murmured vowels in Gujarati presented on the preceding pages has brought the following results:

1. On the physiological level murmured vowels are characterized by a strong airflow. This is a very stable feature. It seems to be due to the presence of a small opening in the rear part of the glottis. Since murmured vowels have, in spite of this opening, the same physical intensity as clear vowels, a stronger activity of the expiratory muscles may be assumed.

2. Acoustically murmured vowels are characterized by one or more of the following features: longer duration, lower start of the fundamental frequency, sometimes a small tone dip in the beginning, higher rise, longer distance to the tonal peak, sometimes lower intensity at the **start**, longer distance to the intensity peak, relatively strong level of the fundamental compared to the second harmonics, and to formant 1, 2, and 4, but not to the third harmonic and to formant 3, sometimes a momentary weakening of formants 2-4 in the beginning of the vowel accompanied by slight noise,

a slight rise in the frequency of formant 1, more damping, asymmetry of the oscillogram, noise in the higher formants.

All these features may be explained by the opening of the glottis with the exception of the longer duration, which is probably due to the historical origin of the murmured vowels as a fusion of vowel with [h] (compare also that isolated murmured vowels seem to be shorter than clear vowels).

The most stable of these cues are duration, distribution of spectral energy, and distance to the tonal peak (the stability of the latter may, however, partly be a consequence of the differences of duration). The least stable features are lower intensity at the start, momentary weakening of higher formants, formant frequency, damping, and asymmetry.

3. The cues which are most important for perception seem to be distribution of spectral energy, low frequency start, and a dip of frequency at the beginning.

None of these cues are necessary, and none is alone sufficient. We are thus faced with a situation where a large number of instable acoustic cues correspond to a simple physiological difference. And murmured vowels may be quoted as an example in favour of the motor theory of speech perception (9), according to which the motor center is involved in the identification of incoming signals. It may well be that a murmured vowel is identified by the speaker as a vowel which he would produce with strong airflow.

The general interest of the investigation lies in this point.

Postscriptum:

For an additional investigation made after the completion of the paper, see p.84.

Inverse filtering analysis.

By inverse filtering the ripples of the formants can be removed from the oscillogram, so that only the glottal curve is left. Such a filtering of one example of the word pair /bar - ba̲r/, spoken by RD, has been undertaken in the Speech Transmission Laboratory of the Royal Technical High School in Stockholm by J. Lindquist. The result is seen below (Fig. 27). It is obvious that the shape of the curve is more sinusoidal and the closure phase is shorter in the murmured vowel than in the clear vowel. This points to a relatively stronger fundamental (as it was also found in the spectral analysis). The murmured vowel has also a somewhat slower fall of the curve corresponding to a slower closing movement of the vocal cords (cp. the suggestion made by Gunnar Fant concerning the asymmetry of the oscillogram 6.5.4.c).

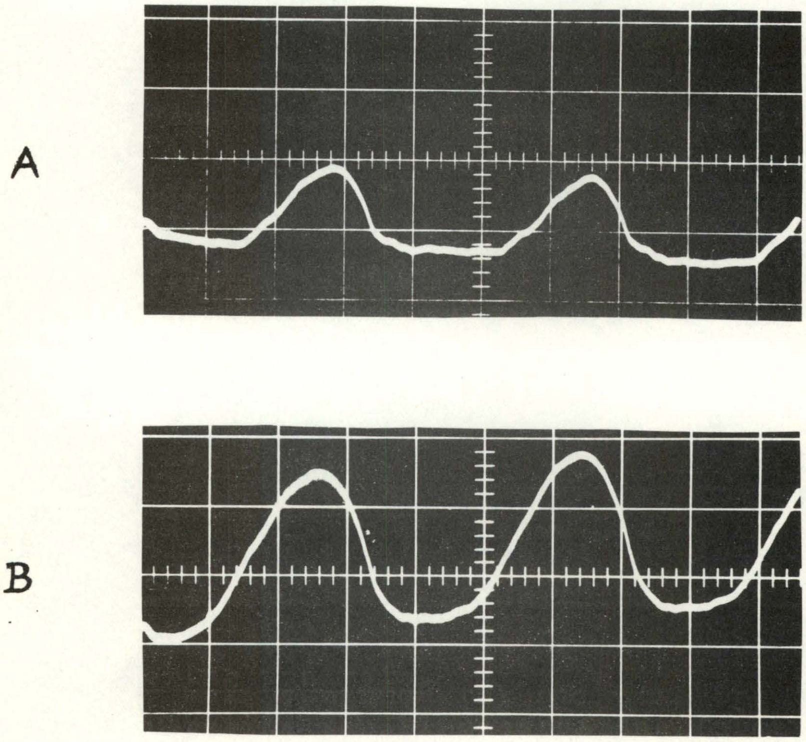


Fig. 27. Inverse filtering curve of the beginning of the vowel in A. /bar/, B. /ba̲r/, spoken by RD.

REFERENCES

1. P.B. Pandit, "Nasalisation, Aspiration and Murmur in Gujarati", Indian Linguistics 17 (1955-56) 1957, pp. 165-172.
2. R. Dave, "A Formant Analysis of the Clear, Nasalized and Murmured Vowels in Gujarati", this report (ARIPUC 2/1967), pp. 119-132, to be published in full in Indian Linguistics.
3. G. Fant, K. Fintoft, J. Liljencrantz, B. Lindblom and J. Mártony, "Formant Amplitude Measurements", JASA 35 (1963), pp. 1753-62.
4. Jørgen Rischel, "Instrumentation for Vowel Synthesis", ARIPUC I/1966, pp. 15-21.
5. Jørgen Rischel, "Heterodyne Filter", ARIPUC 1 (1966), pp. 13-14, and "The Heterodyne Bandpass Filter of the Institute of Phonetics", this report, (ARIPUC 2/1967), pp. 20-34.
6. K. Ch. Bahl, "Tones in Punjabi", Indian Linguistics 17 (1955-56) 1957, pp. 139-147, and Ved Kumari Ghai, "Word Tones in Dogri", this report (ARIPUC 2/1967), pp.133-157, to be published in full in Indian Linguistics.
7. K. Fintoft, B. Lindblom, and J. Mártony, "Measurements of Formant Level in Human Speech", QPSR 2/1962, Speech Transmission Laboratory, Stockholm, pp. 9-17 (p. 16).
8. P.B. Pandit, "E and O in Gujarati", Indian Linguistics XV, 1955-56 reprint edition, pp. 617-55 (p. 628).
9. Ilse Lehiste, "Acoustical Characteristics of Selected English Consonants", International Journal of American Linguistics 3o, Part IV, 1964, 197 pp., see particularly Chapter 3, pp. 141-18o: "Acoustic Characteristics of /h/ and Whispered Speech".
1o. Alvin Liberman, "Some Results of Research on Speech Perception", JASA 29 (1957), pp. 117-22, and A.M. Liberman, F.S. Cooper, Katherine S. Harris, and P.G. MacNeilage, "A Motor Theory of Speech Perception", Proceedings of the Speech Communication Seminar, Stockholm 1962, Volume 2, D3, and the Proceedings of the Seminar on Speech Production and Perception Leningrad 1966, to be published in Zeitschrift für Phonetik und allgemeine Sprachwissenschaft. -