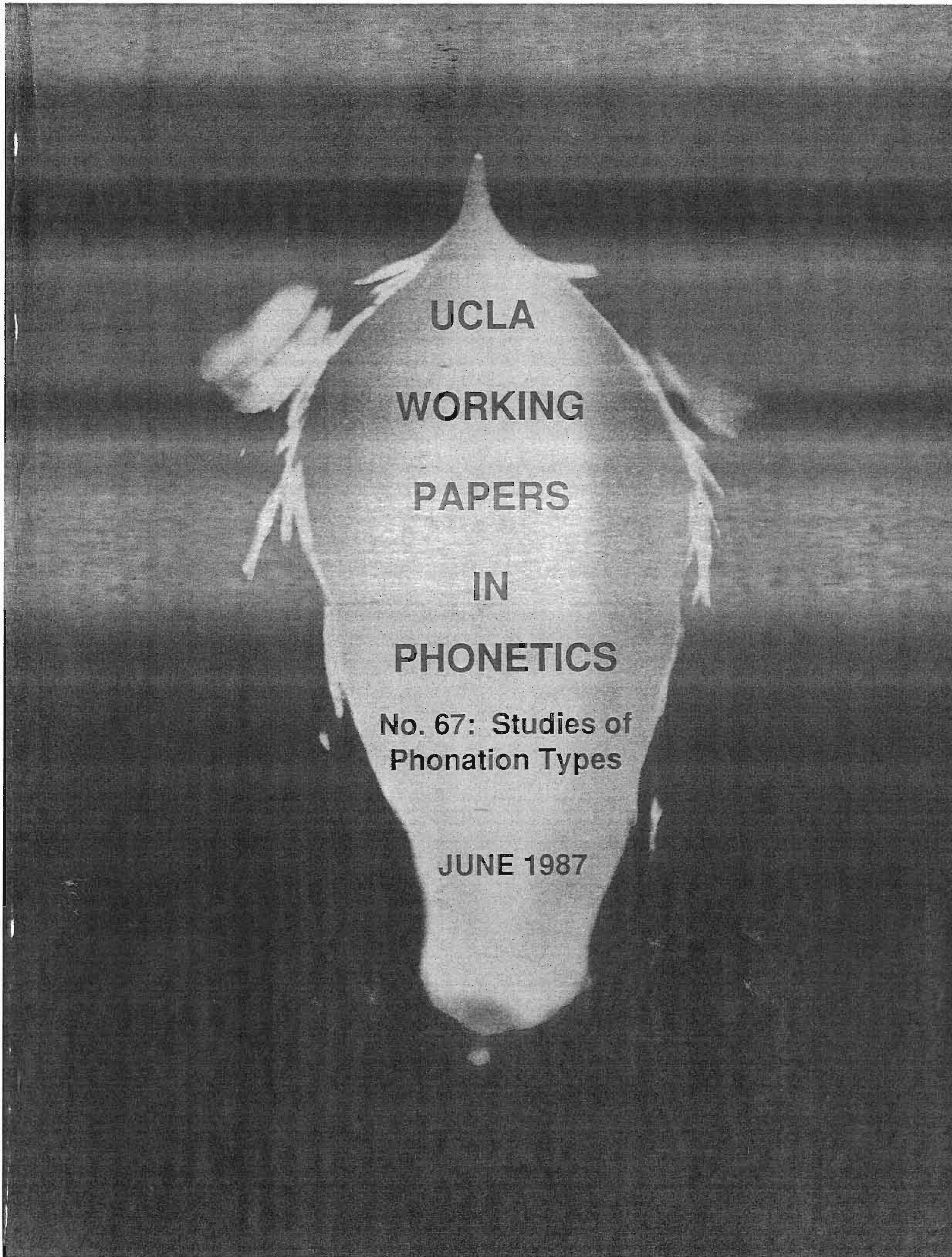


UCLA WORKING PAPERS IN PHONETICS

No. 67: Studies of Phonation Types

JUNE 1987

The UCLA Phonetics Laboratory Group

Stephen R. Anderson
Alice Anderton
Norma Antoñanzas-Barroso
Vernon Austel
Sue Banner
Marjorie Chan
Abby Cohn
Derek Collins
Sarah Dart
Sandy Disner
Prakash Dixit
Bill Dolan
Florence Dolphyne
Karen Emmorey
Vicki Fromkin
Bruce Hayes
Susan Hess
Marie Huffman
Michel Jackson
Victoria Jangozian
Pat Keating
Paul Kirk
Jody Kreiman
Jenny Ladefoged
Peter Ladefoged
Liao Rongrong
Mona Lindau
Priscilla McCoy
Ian Maddieson
Keiko Nagano
Ren Hong-Mo
Lloyd Rice
Mary Lee Scott
Henry Teheranizadeh
Therapan L. Thongkum
Neil Ticktin
Diana Van Lancker
Wanda White

As on previous occasions, the material which is presented in this volume is simply a record for our own use, a report as required by the funding agencies which support the Phonetics Laboratory, and a preliminary account of research in progress for our colleagues in the field.

Funds for the UCLA Phonetics Laboratory are provided through:

NSF grant BNS-8406131
USPHS grant 1 R01 NS18163
USPHS grant 1 RO1 NS22726
and the UCLA Department of Linguistics.

Correspondence concerning UCLA *Working Papers in Phonetics* should be addressed to:

Phonetics Laboratory
Department of Linguistics
UCLA
Los Angeles CA 90024-1543
(U.S.A.)

UCLA *Working Papers in Phonetics* is edited by Ian Maddieson.

** Cover design based on a photograph by John Ohala

UCLA Working Papers in Phonetics 67

June 1987

Anthony Traill and Michel Jackson	Speaker variation and phonation types in Tsonga nasals	1
Therapan L. Thongkum	Phonation types in Mon-Khmer languages	29
R.Prakash Dixit	Mechanisms for voicing and aspiration: Hindi and other languages compared	49
R. Prakash Dixit	In defense of the phonetic adequacy of the traditional term 'voiced aspirated'	103
Ian Maddieson and Susan Hess	The effect on F_0 of the linguistic use of phonation type	112
Peter Ladefoged, Ian Maddieson, Michel Jackson and Marie Huffman	Characteristics of the voice source	119
Peter Ladefoged, Ian Maddieson and Michel Jackson	Investigating phonation types in different languages	126
Therapan L. Thongkum	Another look at the register distinction in Mon	132
	Index of *UCLA Working Papers in Phonetics* #58-67	166

Speaker variation and phonation type in Tsonga nasals

Anthony Traill* and Michel Jackson

*(Department of Linguistics, University of California, Los Angeles and *University of the Witwatersrand, Johannesburg.)*

0. Abstract

Breathy voiced nasals in Tsonga differ in a number of respects from those found in languages such as Newari and Marathi. They sound less breathy, they induce very large perturbations in F_0, and auditorily they appear to be subject to systematic speaker variation affecting phonation type. The two questions of interest are firstly, whether there are measurable properties of these nasals that will allow one to discriminate reliably between the auditorily based groups and secondly, whether the speaker variation is constrained by a principle of relative invariance or is random. The study suggests that there are reliable speaker differences based inter alia on spectral tilt and that most, but not all, subjects produce the modal/ non-modal contrast in a relatively invariant way.

1. Introduction

Tsonga, a Bantu language of Mocambique and South Africa makes the familiar linguistic contrast between nasals produced with normal voicing (modal voice) and nasals produced with breathy voicing (also referred to as murmur). In this respect it is like Newari and Marathi, for example: all three languages can be said to make the same phonation contrast among nasal consonants. Despite the appropriateness of this cross-linguistic identification of sound types, there are a number of significant respects in which the phonetic details of Tsonga breathy voiced nasals differ from those of Newari and Marathi. Firstly, as casual listening will confirm, Tsonga breathy voiced nasals sound less breathy and secondly they are accompanied by a marked lowering of the fundamental frequency (Ladefoged l983). A third detail, hitherto unreported, involves systematic speaker variation in the pronunciation of the contrast: some individuals regularly pronounce the phonologically non-modal nasals with a phonation type that is auditorily not breathy but more akin to what might, just for convenience, be called "stiff voice". These language specific details are of more general interest because of their relevance to hypotheses about constraints on phonetic variation in phonation type both across and within languages.

The development of various objective measures for characterising linguistically significant differences in phonation type has been the subject of a number of recent papers from the UCLA Phonetics Laboratory (Ladefoged 1983; Maddieson and Ladefoged 1985; Kirk, Ladefoged and Ladefoged 1984; Ladefoged and Antoñanzas-Barroso 1984; Huffman 1987; Jackson *et al.* 1985a,b). An important goal of this program of research has been to find measures that will yield results that are comparable across both speakers and languages so that generalizations can be made about the phonetic categories available for phonation type contrasts. The well-known problem in this connection is that the "same" phonation type contrast in two languages, or in the speech of different speakers of the same language, may be produced in ways that are measurably distinct (Ladefoged 1985; Jackson *et al.* 1985b). In spite of this variability, phonologists and phoneticians do identify phonation type contrasts across languages and across speakers of the same language and this raises the question whether there is a principled basis to this such identifications.

One hypothesis to emerge from the above research proposes that individual variations in the pronunciation of a binary contrast, for example, will involve discrete pairs of points ordered along one articulatory or acoustic continuum. This hypothesis leads to the expectation that all the speakers of a language will produce a particular phonation contrast in the same way, but only relatively speaking (Kirk *et al.* 1984, p.103). There will thus be a relative invariance across speakers in the phonetic realization of a linguistic contrast and this will explain how an individual with, for example, a normally breathy voice can be successfully identified by other speakers or phoneticians as producing the same linguistic contrast between modal and breathy voice as an individual with a non-breathy modal voice.

Against this background the present study explores two problems. The first is whether there is an objective basis to the auditorily based division of Tsonga speakers into two non-modal phonation type groups. Presuming that these speaker differences are indeed measurable, the second problem is whether, as hypothesized, they qualify as more or less the same pronunciation of the linguistic contrast.

2. <u>Language Data</u>

The consonants of Tsonga may be divided into two classes according to their effect on a following high tone. Those that cause a high tone to become a rising tone are termed depressors (Louw 1968). The distinction between depressor and non-depressor nasals also involves a clear phonation type contrast. The latter are breathy (the literature on Tsonga uses the term "murmured" (Baumbach 1974; Louw 1968) and the former have modal voice. According to Louw, vowels following breathy voiced nasals are also breathy voiced: "it is important to note that murmur is carried over onto the following vowel. A correct phonetic transcription of a word such as <u>mhala</u> would thus be [mɦɑ̤lɑ]." (p.95). However, such a phonetic transcription is misleading because it fails to reflect the distinctive phonation type of the nasal itself. Moreover, it suggests that the main phonetic effect lies outside the nasal as it does in, for example, Hindi (in fact, to a speaker of Hindi, Tsonga breathy voiced nasals do not sound like those of Hindi (P. Dixit, personal communication)).

The following words illustrate this modal/breathy contrast. This set or subsets of it provided the data for the investigation. Because of the relatively small number of words in Tsonga that have breathy nasals it was not possible to select only minimal or near minimal pairs. However, in all these examples the vowel following the nasal was held constant, and care was taken to follow the nasals with both high and low toned vowels. Each word was checked for acceptability with all subjects before inclusion in the final list and it was established that each subject indeed made a phonation type contrast between the modal and breathy nasals.

High quality tape recordings were made of each word produced twice in isolation and once in the frame 'ndzirì____', 'I am saying____'.

mákálà	'slap'	m̤áwúrí	'august'
		m̤ándzé	'post'
màsásánì	'righteous person'	m̤àká	'matter'
màsàná	'sun's rays'	m̤àlà	'sp. of antelope'
náthà	'sieve finely'	n̤árù	'three'
		n̤ábyà	'jump up in surprise'
nàlá	'enemy'	n̤àngá	'hut for young girls'
ŋáwú	'miaow of cat'	ŋ̤wàtí	'sp. bird'
ŋáŋàní	'sp. bird'		

3. Subjects and procedures.

Fifteen subjects, five female and ten male ranging in age between 18 and 60 years participated in the investigation. Six of them, three female and three male were also the subjects of an aerodynamic investigation. All subjects were resident in the Giyani district of Gazankulu, South Africa, but they were not dialectally homogeneous. However, dialectal variation in Tsonga does not affect the breathy/modal distinction amongst nasals (Baumbach 1974). The original intention was that the aerodynamic investigation would supplement the perspective provided by acoustically based measures of the audio recordings. When all these data were being collected, the facts about individual differences in phonation type were not known. These differences became apparent only after the data had been recorded and so the subjects represent an entirely random and, unfortunately, unbalanced sample of the individual variations of interest. Moreover, the aerodynamic results, though interesting in their own right, turn out to bear only indirectly on the results obtained from the acoustic measures of phonation type.

The aerodynamic investigation involved simultaneously recording nasal and oral flow, laryngeal activity via a Fourcin laryngograph and the audio signal. All these data were recorded on a Mingograf 34 inkwriter. Nasal and oral flows were calibrated with a Gilmont flowmeter. The experimental set up is schematically given in Figure 1

Figure 2 illustrates various measurements made on or derived from these tracings:-
- the peak nasal flow;
- the nasal flow rise time, i.e. the time from the onset of nasal flow to peak nasal flow;
- the duration of nasal flow;
- the normalized nasal flow rise time, i.e. the rise time divided by the nasal flow duration;
- the nasal flow at four equidistant points through the nasal;
- the mean of the four nasal flow measurements;
- the oral flow at 25 ms after vowel onset and at 75 ms after vowel onset;

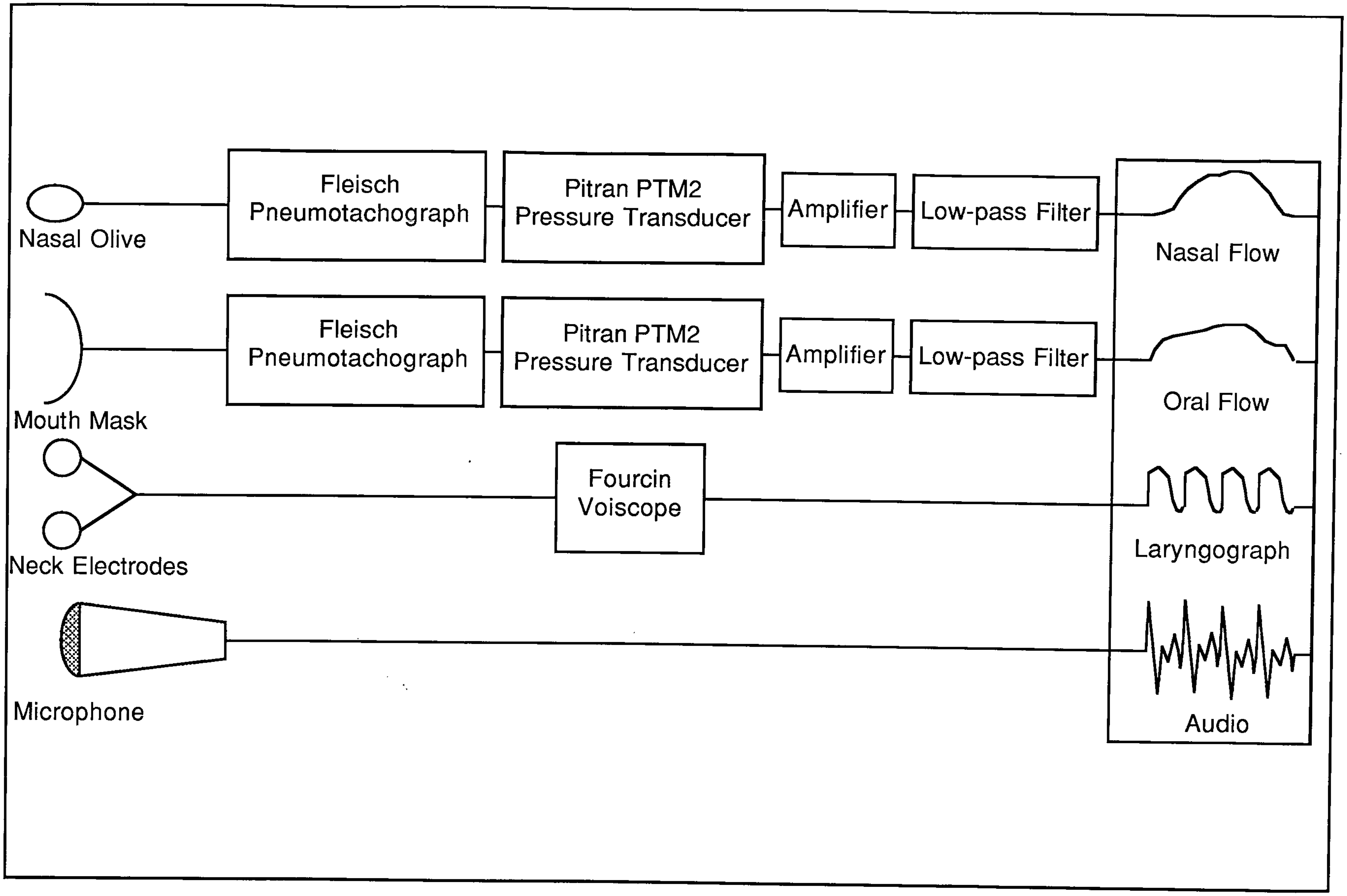
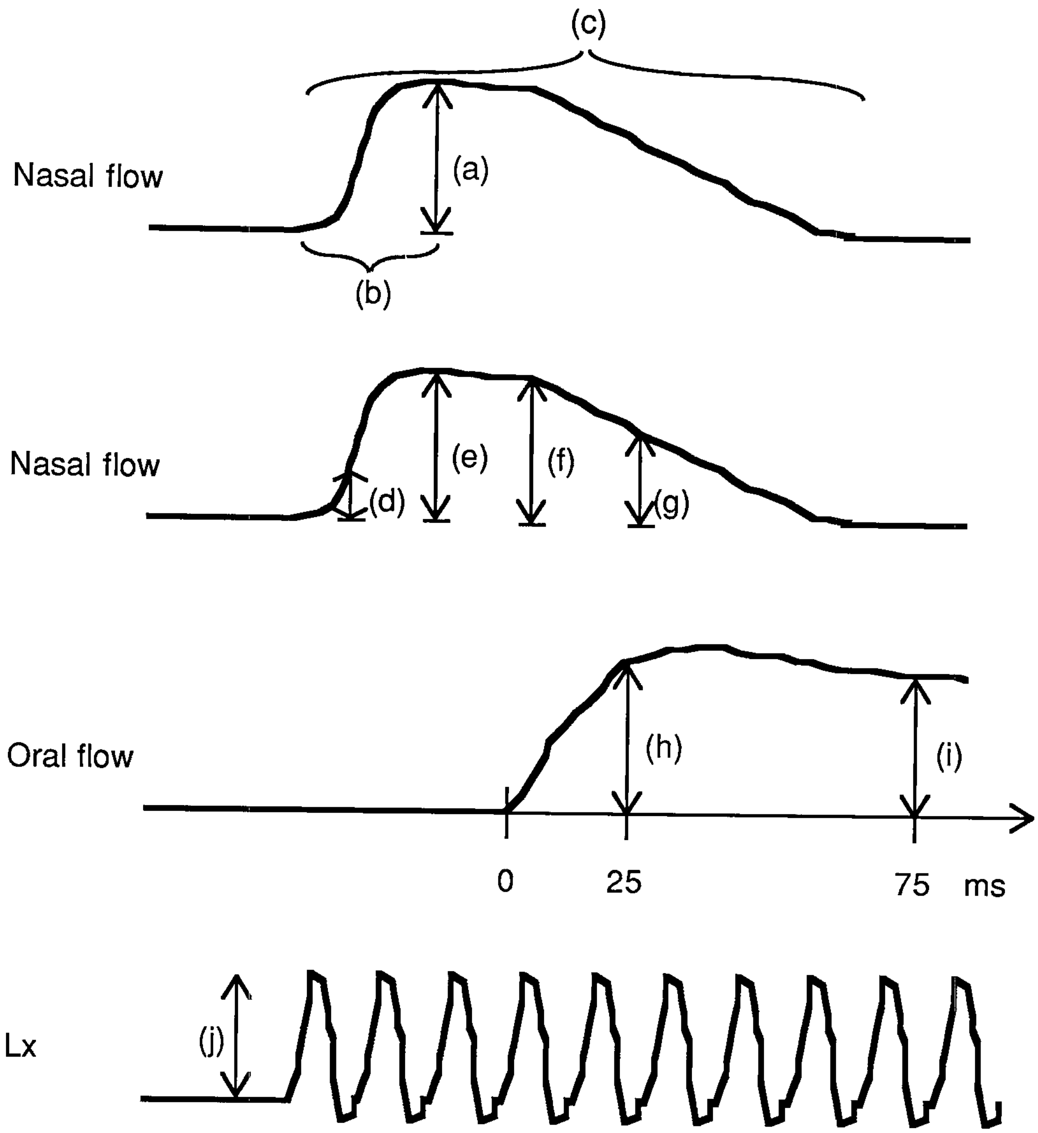


Figure 1. The experimental setup used for the aerodynamic investigation of modal and breathy voiced nasals



Key to measurements:

(a) Peak nasal flow
(b) Rise time
(c) Duration of nasal
(d)-(g) Nasal flows
(h)-(i) Oral flows
(j) Maximum laryngograph amplitude

Derived measurements:

(b), (c) → Rise time / Duration = Normalized flow rise time
(d)-(g) → Mean nasal flow
(h)-(i) → Mean oral flow

Figure 2. Measurements and derived measurements from the aerodynamic investigation.

- the mean oral of the oral flow measurements;
- laryngographic amplitude.

The appropriateness of most of these parameters for quantifying the distinction between breathy and modal nasals will be evaluated. However, a brief explanation of some of them is necessary. Firstly, measuring oral flow during vowels following nasals was included to test the claim mentioned earlier that vowels following breathy voiced nasals are also breathy voiced and indeed that breathy voiced nasals may most appropriatedly be transcribed with a period of breathy voiced aspiration before the onset of the (breathy voiced) vowel proper (recall Louw's transcription [mɦɑ̤lɑ] above). Secondly, laryngographic tracings provide a graphic display of impedance changes caused by different glottal settings, and it is to be expected that the phonation distinction involving nasals should be sytematically reflected in changes to the laryngographic waveform.

The audio recording of the fifteen subjects was used for a number of acoustic analyses. Narrow band spectrograms were analyzed for an accurate representation of F_0 variation through the nasal consonant and the following vowel, and wide band spectrograms provided formant frequencies and information about the general distribution and nature of spectral components. Measurements of jitter, shimmer, and signal-to-noise ratio were made using the CSpeech computer program (Milenkovic 1986). Jitter, shimmer, and signal-to-noise ration were calculated for 70 ms of the nasal and 50 ms of the contiguous portion of the following vowel. The obtained jitter measures were normalized across subjects to control for differing F_0 ranges, by dividing them by the mean pitch period for each subject. For these measures, the recordings were sampled at 8.333 kHz to 12 bits.

Other acoustic measures were determined using programs implemented in the WAVES system at the UCLA Phonetics Lab. Several spectral tilts and values of the breathiness index developed by Ladefoged and Antoñanzas-Barroso (1984) were calculated from each token. The recordings were sampled at 10 kHz to 12 bits.

Spectral tilt was measured by hand from averaged spectra derived by chirp z-transform (Jackson *et al.* 1985 b) over the frequency range 0-2500Hz. The measures of spectral tilt in this study are based on measurements of relative harmonic intensity in the audio signal. Although this technique has been employed succesfully for the analysis of phonation contrasts involving vowels (e.g. Maddieson and Ladefoged 1985; Kirk *et al.* 1984) it has apparently not been used for such contrasts in nasals and it was not *a priori* obvious how it should be applied or modified. Informal examination of spectrograms and spectra led to the conclusion that the relative intensity of the first harmonic could be reliably compared with the intensities of two higher points in the nasal spectrum in addition to comparing it with the intensity of the second harmonic.

The first higher point was located at the harmonic closest to 1.4 kHz where there was a regular reduction of intensity in the spectrum as determined from spectra and spectrograms. For convenience, these three measures of spectral tilt will be referred to as Nasal tilt 1, Nasal tilt 2, and Nasal tilt 3. If this point fell halfway between two harmonics, their amplitudes were averaged.

The second higher point was the most intense harmonic in the nasal formant above 2 kHz, again determined from spectra and spectrograms. In practice there was no difficulty in locating these points for each subject, though care had to be exercised to avoid identifying spurious peaks as harmonics.

Averaged spectra were also derived for the first 40 ms of the vowel immediately following the nasal as well as of the remainder of the vowel, and the H1-H2 measure of spectral tilt was calculated from these averaged spectra. These spectral tilts were calculated in order to test the claim that breathiness spreads into the vowel following breathy voiced nasals.

The breathiness index was computed for entire nasals, the first 40 ms of the following vowel, and the remainder of the vowel. Again, the breathiness index of the following vowel was calculated in order to test the claim that breathiness spreads into the following vowel.

The 14 acoustic measures made on each token are summarized below.

Table 1. Acoustic measures made on each token.

<u>NASAL</u>	<u>VOWEL ONSET</u>	<u>REST OF VOWEL</u>
Jitter	Jitter	
Shimmer	Shimmer	
S/N ratio	S/N ratio	
Nasal tilt 1 (H1-H2)	Tilt (H1-H2)	Tilt (H1-H2)
Nasal tilt 2 (H1-H at 1.4kHz)		
Nasal tilt 3 (H1-strongest H above 2.0 kHz)		
Breathiness index	Breathiness index	Breathiness index

4. <u>Results</u>

It was noted above that dialectal differences in Tsonga do not affect the phonation contrast in nasals. However, there is clear speaker variation in the production of the contrast between a modal and breathy voiced nasal. It seems that one group of subjects' breathy voiced nasals are consistently produced with an auditorily clear, breathy voiced phonation while those of the other group seemed to be consistently produced with a more constricted glottis yielding an auditory impression of a stiffer, more tense phonation. These distinctions cut across both sex and age differences. Despite these phonation type differences the non-modal nasals of both groups nevertheless produce the same depressor effect on following tones, that is they caused following high tones to become rising tones, and following low tones to become lower.

The 6 subjects who participated in the aerodynamic study do not represent adequate samples from these two groups and so the results cannot be used systematically to verify the difference. The results of this part of the study are therefore presented below mainly without regard to this issue, as interesting in their own right.

(a) Aerodynamic measures.

A higher air flow for breathy nasals than for modal nasals was expected, but was found only in the second half of the nasals, at and after the point of peak nasal flow. Nasal flow at points 1 and 2 show no significant differences between the two types of nasal. The normalized nasal flow rise time is not significantly different between breathy and modal nasals, though the peak flow is significantly higher in breathy nasals. As a consequence of the higher flow at the peak and in the second half of the nasal, the mean nasal flow is also higher in breathy nasals, though the significance level associated with the difference between the breathy and modal nasals is lower.

Similarly, the oral flow measured at 25 ms after vowel onset shows significant differences, while the flow measures at 75 ms after onset does not. The mean of these measures also shows significantly more flow for the breathy nasals, though again the significance level is lower. The significant differences found for other aerodynamic measures are presented in Table 2, which shows the grand means for all modal and breathy tokens for all subjects.

Table 2. Six measures of airflow (in ml./s.), showing significant differences between modal and breathy voiced nasals. (Mean of six subjects.)

	Modal (n=103)		Breathy (n=112)
nasal flow 1		n.s.	
nasal flow 2		n.s.	
nasal flow 3	70	$p < .01$	78
nasal flow 4	78	$p < .01$	88
mean nasal flow	63	$p < .05$	68
peak nasal flow	82	$p < .001$	93
oral flow 1	60	$p < .0001$	80
oral flow 2		n.s.	
mean oral flow	74	$p < .01$	90

These results show that airflow gradually increases during both modal and breathy voiced nasals, reaching greater flows in the latter part of the nasal, with the greatest flows being in the breathy voiced nasals. This greater flow persists into the early portion of the vowel, but not into the later portion. The laryngographic amplitudes for breathy and modal nasals differ significantly ($p < .0001$) with larger amplitudes for the latter. One might say that the joint effect of all these results is to show there is less glottal resistance to the airflow during breathy voiced nasals because of a less constricted glottis, a conclusion not at variance with conventional definitions of the phenomenon.

The most notable individual variation in these results involves a subject whose nasal flow patterns were different. The subject's phonologically breathy nasals had been auditorily classified as phonetically "non-breathy". Neither the aerodynamic nor the laryngographic measures distinguish the elements of this subject's linguistic contrast. It seems from this evidence that the basis of this subject's phonation contrast involves a laryngeal adjustment that is very different from that of the other subjects. This question will be addressed in greater detail below when acoustic measures for larger numbers of subjects are discussed.

A different case of individual variation involved another subject for whom there was a significant difference only in the first oral flow measure; flows for modal and non-modal nasals did not differ significantly during the nasal and there was no laryngographic difference. This subject's non-modal nasals had, however, been assessed as being breathy voiced.

(b)Acoustic measures

The so-called "depressor" effect of breathy voiced nasals on the tone on a following vowel is illustrated in Figure 3 where the mean tonal contours of all fifteen subjects for modal and breathy voiced nasals followed by high and low-toned vowels are displayed.

The dramatic effect on both high and low tones of a "depressor" nasal is clear from Figure 3. It is perhaps useful to express it as a percentage of the mean pitch of the corresponding non-depressed tone: depression of a high tone is about 20% during a breathy voiced nasal and 17% during the following vowel; depression of a low tone is about 16% during the nasal and 4% during the following vowel. These effects appear to be part of the phonology of the language in that they do not seem to be derivable as obviously natural consequencies of the glottal adjustment necessary for breathy voice per se. Two other features of these contours require comment. Firstly, the maximum degree of depression occurs at about the middle of the nasal. Its effect persists throughout a depressed high tone of the vowel and through about half a depressed low tone of the vowel. It is interesting to note that the time course of these F_0 contours closely matches the timecourse of the airflow changes reported above during breathy voiced nasals and following vowels: increased flows commence well into the nasal and persist into the following vowel. However it is sufficient to recall the behaviour of the exceptional subject referred to above, to realize that the tonal contour is not necessarily due to the glottal adjustments that lead to increases in airflow since this subject had the same tonal contours as the other subjects but without an increase in flow. Secondly, F_0 is relatively steady only during low toned modal nasals. In all the remaining cases F_0 is constantly changing during both nasal and vowel and through extensive ranges across male and female subjects. This will be an important factor to be remembered below when making inferences about spectral tilt from spectra which have not been inverse-filtered and which reflect averages over the time stretches.

It will be clear from Figure 3 that it is not possible to control completely for pitch changes in a study of this kind. However, an attempt has been made to minimize this

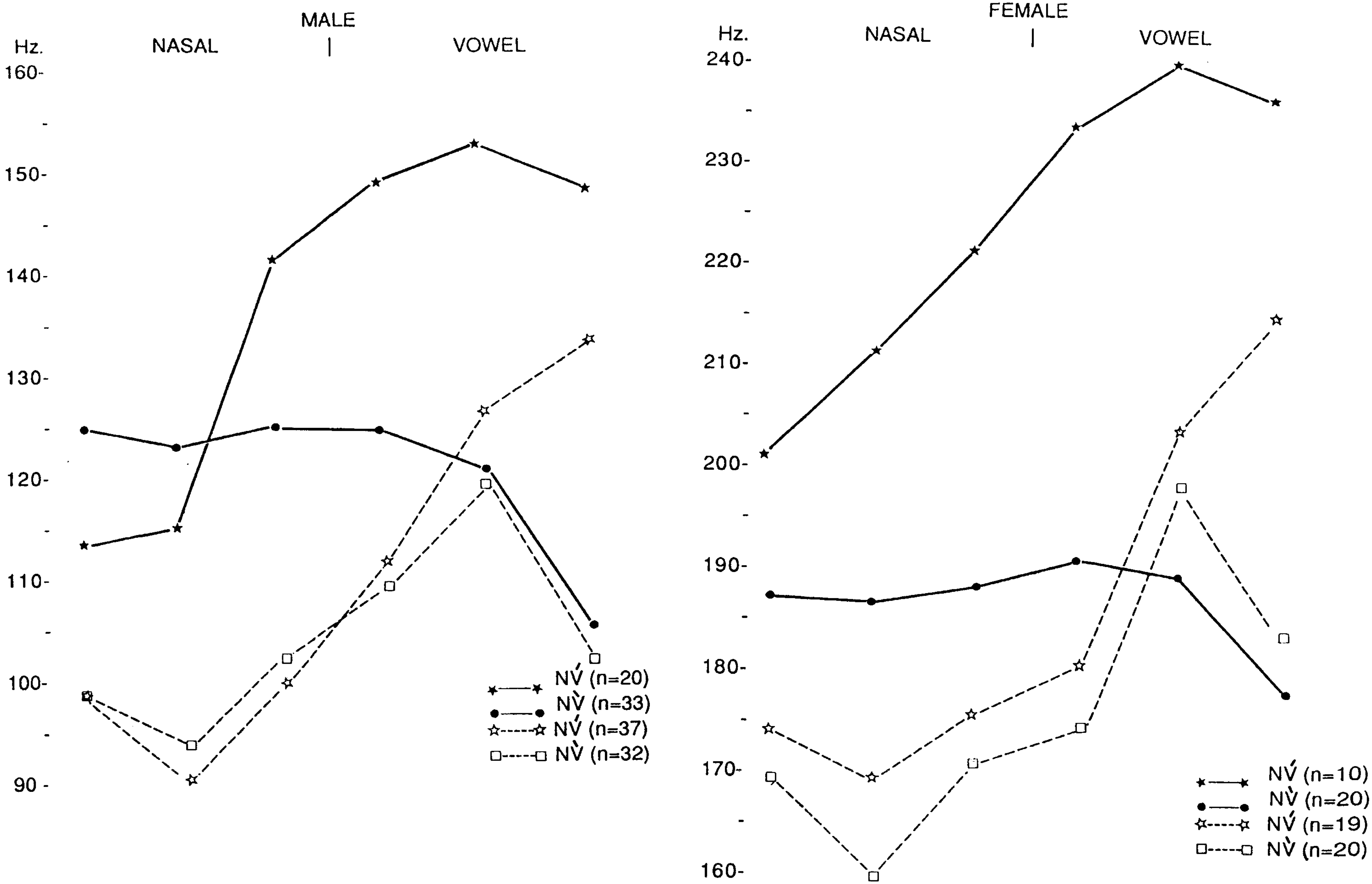


Figure 3. Mean tonal contours for modal and non-modal nasals followed by high and low vowels.

effect firstly by separating male and female subjects for purposes of statistical analysis and secondly by ignoring high toned tokens in the analysis of spectral tilt using chirp z-transforms. Unfortunately this reduces the number of tokens for each subject for the measures of spectral tilt to two each.

In order to control the effect of individual variation on the statistical analyses, modal values were subtracted from non-modal values to yield a "difference value" for each of these parameters.

Subjects were divided by sex and assigned to one of two "phonation type" groups: Group A contained four male and three female subjects whose non-modal nasals had been judged auditorily to involve breathy voicing and Group B contained six male and two female subjects whose non-modal nasals had been judged to involve a different (i.e. non-breathy voiced) phonation type. As noted above, the groups are of unequal size because subjects had not originally been selected with a view to testing a hypothesis concering individual differences in the production of non-modal nasals.

An analysis of variance was performed to see if any of the fourteen variables were significantly affected by the group classifications. For male subjects this revealed a significant group effect on Nasal tilt 2 ($p < .005$) and a probably significant effect on Nasal tilt 3 ($p < .05$). Weaker effects were found for jitter at vowel onset ($p < .10$) and spectral tilt at vowel onset ($p < .10$). For female subjects, probably significant effects were found for Nasal tilt 2 ($p < .05$) and the signal to noise ratio at vowel onset ($p < .05$). Discriminant analyses were then performed on the two variables Nasal tilt 2 and Nasal tilt 3 for the male subjects and on the two variables Nasal tilt 2 and signal to noise ratio at vowel onset for the females. The results for male subjects are presented in Table 3 in terms of the percentage of pairs of tokens correctly or incorrectly classified.

Table 3. Results for male subjects of a discriminant analysis using the two variables Nasal tilt 2 and Nasal tilt 3.

Pairs of tokens from group	Percentage classified as Group A	Percentage classified as Group B
A (4 subjects)	75 (12 tokens)	25 (4 tokens)
B (6 subjects)	17 (4 tokens)	83 (20 tokens)

Each individual is represented by the difference scores between two pairs of tokens. In no case are both pairs from a given speaker misclassified, hence no individual is completely misclassified although, as we shall see below, the confidence associated with particular cases varies. Thus the division of male subjects into two groups based on the auditory effect of their non-modal nasals appears to receive relatively good support from the results of a discriminant analysis using the two nasal tilt measures. Table 4 displays the misclassified cases.

Table 4. Four pairs of tokens misclassified by the discriminant analysis, male subjects. (Parenthesized probabilities refer to the other pair of tokens)

Subject from group	Individual	Probability of membership in Group A	Probability of membership Group B
A	NM	.22 (.98)	.78 (.02)
	PM	.36 (.93)	.64 (.07)
B	SD	.65 (.27)	.35 (.73)
	GSH	.58 (.21)	.42 (.79)

Table 4 gives the degree of confidence associated with each classification; the parenthesized probabilities refer to the other pair of tokens, correctly classified by the discriminant analysis. We shall delay discussion of these cases until a later stage when an attempt has been made to establish a "criterion probability" for successful classification.

The results of the discriminant analysis for female subjects appear in Table 5.

Table 5. Results for female subjects of a discriminant analysis using the two variables Nasal tilt 2 and signal to noise ratio at vowel onset.

Pairs of tokens from Group	Percentage classified as Group A	Percentage as Group B
A (3 subjects)	83	17
B (2 subjects)	25	75

The discrimination between Group A and B subjects is the same as that found for male subjects and once again, no subject is misclassified on both pairs of tokens. The misclassifications are shown in Table 6.

Table 6. Two pairs of tokens misclassified by the discriminant analysis, female subjects. (Parenthesized probabilities refer to the other pair of tokens).

Subject from group	Individual	Probability of membership in Group A	Probability of membership in Group B
A	SM	.15 (.79)	.85 (.21)
B	MBA	.53 (.01)	.47 (.99)

Discussion of these classification probabilities will be taken up below.

In order to refine the interpretation of these statistical tendencies a correlation analysis was performed on all variables in order to reveal their interactions. The two spectral tilt measures used in the discriminant analysis for males are positively correlated ($r \approx .5$, $p < .05$) and the two measures used for females are possibly weakly correlated ($r \approx .6$, $p < .10$). This may be interpreted to mean that both pairs of variables are directly contributing to the group classification rather than introducing obfuscating individual effects.

Since the subjects had been assigned to two groups according to the subjective criterion of their auditory impression, it could be argued that the division lacks independent validity and therefore that the success of the discriminant analyses may not reflect a real distinction between the subjects. An attempt was made to deal with this problem for male subjects only by conducting discriminant analyses on groups into which a few deliberately misclassified individuals had been placed (this could not be done in a meaningful way for female subjects because of the small number of subjects). In the first analysis two A and three B Group members were assigned to the wrong groups. The result of this for two individuals was that one A and one B Group member was misclassified on both pairs of tokens by the discriminant analysis, i.e. was reassigned to his correct group. Interestingly, none of the other subjects was misclassified but their probability of membership in their assigned group was far lower than the probability associated with their correct grouping in the original analysis. Moreover the overall confidence with which any pairs of tokens was correctly assigned to their groups was scarcely better than chance ($p \approx .58$ for both groups). This means that the deliberate misclassifications were lowering the degree of confidence of every other classification. In the second analysis a single misassigned Group B individual was replaced with one who represented an "extreme" case of the Group B pronunciation. A similar result was obtained: overall probability of correct classification was low for all cases ($p \approx .59$ for Group A and $p \approx .61$ for Group B) and the same subjects were misclassified. The "canonical" Group B individual was (incorrectly) classified as Group A but with a probability no greater than chance ($p \approx .51$ and .52 for the two pairs of tokens). It is reasonable to conclude therefore that the original groupings reflect something of an optimum with respect to the quantitative variables .

These results also provide a useful guide to setting a criterion probability for group membership. Based on the worst case of undetected misclassification with a probability of .73, it would be necessary to proceed with a criterion probability of at least .75 for a reasonably stringent evaluation of any successful classifications. This is extended arbitrarily to the female subjects. If a classification probability does not exceed the criterion value, it is not considered a reliable classification. Applying this criterion to Tables 4 and 6, one difference score from each of two Group A subjects, the male subject NM, and the female subject SM, is completely misclassified by exceeding the criterion. The Group B subject SD, on the other hand is not reliably classified into either group, since none of classification probabilities attain the criterion value.

Since the ANOVA results show only weakly significant effects, other combinations of variables chosen for their phonetic coherence were used to model the two groups and these combinations were subjected to discriminant analyses. The analysis for male subjects based on all seven nasal variables (jitter, shimmer, signal to noise ratio, three measures of tilt and breathiness index) fails to improve on the classification based only on two variables (Tilt 2 and Tilt 3), and provides precisely the same percentage success for Groups A and B. Again, in no case were both the pairs of tokens for any subject completely misclassified. The addition of four variables only marginally affects the degree of group membership of individuals without significantly improving on the overall classification. This suggests that some of these variables are merely acting to suppress variances (Cohen and Cohen 1975, pp 84 ff.) through redundancies. This supposition is supported by the negative correlation between nasal jitter and nasal signal to noise ratio ($r \approx -.7$, $p < .001$) and the positive correlation between the first nasal tilt and the breathiness index ($r \approx .7$, $p < .0005$).

For female subjects, the full set of nasal variables assigns subjects correctly in 100% of the cases. But there are more variables than subjects and once again strong positive and negative correlations exist between variables: nasal jitter and nasal shimmer ($r \approx .8$, $p < .005$), nasal jitter and nasal signal to noise ratio ($r \approx -.7$, $p < .02$), Nasal tilt 2 and breathiness index in the nasal ($r \approx -.7$, $p < .05$).

A further discriminant analysis based on only the three measures of spectral tilt during the nasal, provided an identical overall classification for male subjects as that using the above groups of variables, but with the same misclassifications as the analysis based on the two spectral tilt measures. For female subjects the discrimination between the groups improved with these three variables against the two variables selected from the results of the ANOVA: subjects in Group A were classified 100% correctly and one subject from Group B (MBA) was misclassified for one pair of tokens with a probability of membership in Group A of .51 (cf. .99 probability of membership in Group B for the other pair of tokens).

Male subjects were further modelled with the set of variables Nasal tilt 2,Nasal tilt 3 , tilt at the vowel onset, and jitter at the vowel onset. As noted above, the latter pair were found to be possibly weakly correlated with group membership in the ANOVA. This improved the classification of Group B subjects to 100% but classified Group A subjects only 63% successfully; again no subject was completely misclassified. Between these variables there is a positive correlation between the jitter measure and the third nasal tilt measure ($r \approx .5$, $p < .02$).

The breathiness index of the nasal discriminates between the two male groups very weakly (Group A 63%, Group B 67%) misclassifying 7 of the 10 subjects, but each on only one pair of tokens. With female subjects this variable classifies Group A subjects 83% successfully but fails to classify Group B in 50% of cases. No Group A or B subject is misclassified for both pairs of tokens.

In order to apply the criterion of successful classification adopted earlier, the results of these discriminant analyses should be expressed in terms of the overall probability associated with successful discrimination into the two groups. This probability is based on the statistic labelled "posterior probability of membership in group" provided for each individual by the discriminant analysis but in Table 7 below the effect of unsuccessful cases has not been computed.

Table 7. Probability of successful classifications into groups. The partly misclassified subjects are recorded by initials.

	Male subjects		Female subjects	
Variables	Group A	Group B	Group A	Group B
1. All nasal	.85 NM, PM	.80 SD,EB	.99	1.00
2. Selected by ANOVA	.84 NM,PM	.78 SD,GSH	.91 SM	.88 MBA
3. Three nasal tilts	.83 NM,PM	.78 SD,GSH	.88	.95 MBA
4. Breathiness index in nasal	.51 NM, PM, SN	.52 JM,MM,EB,GSH	.69 SM	.89 MBA,GS
5. ANOVA plus tilt and jitter at vowel onset	.98 MM,SN,RNG	1.00	–	–
6.Breathiness index at vowel onset	.58 NM,SN,RNG	.60 SD,JM,EB,SK	.58 SM	.63 MBA,GS
7.Tilt at vowel onset	.71 NM,SN,RNG	.63 JM,EB	.78 AM,CM	.73 GS

The sets of measures numbered 4,6,7 do not meet the criterion for one or both groups. Of those that do, the first was rejected above leaving 2, 3, and 5 for males and 2 and 3 for females as potential models of the phonetic behaviour of the groups. The third model for males does not improve on the second despite the additional variable and is therefore discounted. The fifth model has by far the highest posterior probability associated with it for male subjects, and contains properties of the nasal and vowel onset both of which have been traditionally regarded as features of the non-modal nasals. However, as has been noted, the difference measures of tilt and jitter at vowel onset are not statistically significant which means that the fifth model is less adequate than the second. For female subjects, the third model only marginally improves the discrimination at the cost of two additional statistically insignificant variables and it is discounted. Therefore, Tilts 1 and 2 best model the male group differences and Tilt 2 and signal to noise ratio the female group differences.

The results reported so far are all based on difference scores for each variable. In order to explore the phonetic details of each group's modal and non-modal nasals, mean scores for the two classes are recorded in Table 8 below.

Table 8. Mean scores for breathy and modal tokens for measures used in the discriminant analyses.

Speakers:		Male subjects		Female subjects	
Variable(s):		Group A (n=4)	Group B (n=6)	Group A (n=3)	Group B (n=2)
Tilt 1	M	-9.062	-20.363	23.833	23.438
	B	-3	-11.279	7.75	12.188
Tilt 2	M	18	-8.458	8.125	21.375
	B	19	2.188	31.125	18
Tilt 3	M	1.656	-17.188	20	30.375
	B	-2.625	-7.062	24	36.375
Tilt in vowel	M	-12.094	-16.396	4.875	0.375
	B	4.827	-11.104	-3.342	0.875
Breathiness index in nasal	M	.372	.398	.38	.68
	B	.387	.459	.193	.315
SN Ratio	M	--	--	15.389	14.978
	B	--	--	14.449	11.641
Jitter in vowel	M	.187	.103		
	B	.127	.146		

Because of the small numbers of tokens involved in measures of spectral tilt, tests of significance for the differences in Table 8 cannot be seriously considered.

The precise relevance of these numbers for the concept of spectral tilt in the nasal are best visualized in the form of schematic line spectra for each group reflecting the relative amplitudes of the first and second harmonics and the harmonic at about 1.4 kHz and the most intense harmonic in the formant above 2 kHz. Figure 4 presents this for male subjects and Figure 5 for female subjects.

The amplitude of the first harmonic, the reference harmonic in these spectra, has been set at 0 dB across all subjects so as to facilitate comparison. The amplitude of the higher harmonics therefore reflects the difference measures.

The nasal spectra for Group A males differ from those of the Group B males in terms of a general feature affecting both modal and non-modal nasals. Relative to the first harmonic, Group B subjects have more intensity at the higher points in the spectrum than Group A subjects. This may be interpreted to mean that on the whole, all Group A nasals are more breathy than all Group B nasals. The modal/non-modal contrast is made within these two phonatory settings. The relationship between modal and non-modal nasals is remarkably constant within each group for each measure of nasal tilt as can be seen in Table 9.

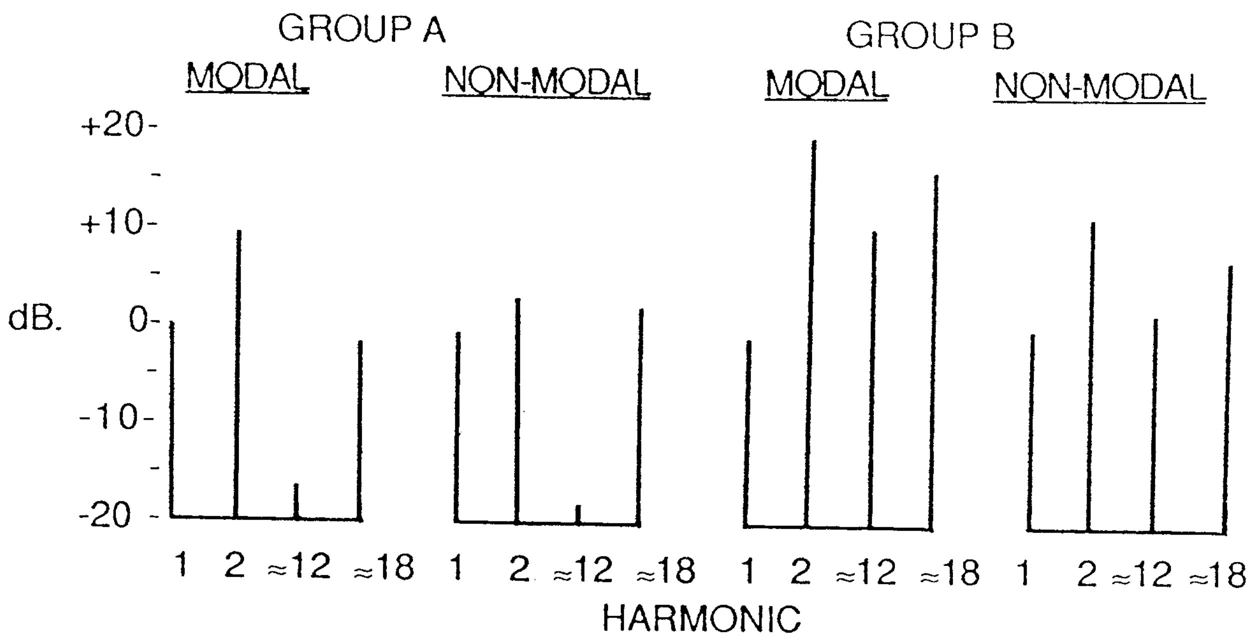


Figure 4. Schematic line spectra for modal and non-modal nasals for Groups A and B, male subjects

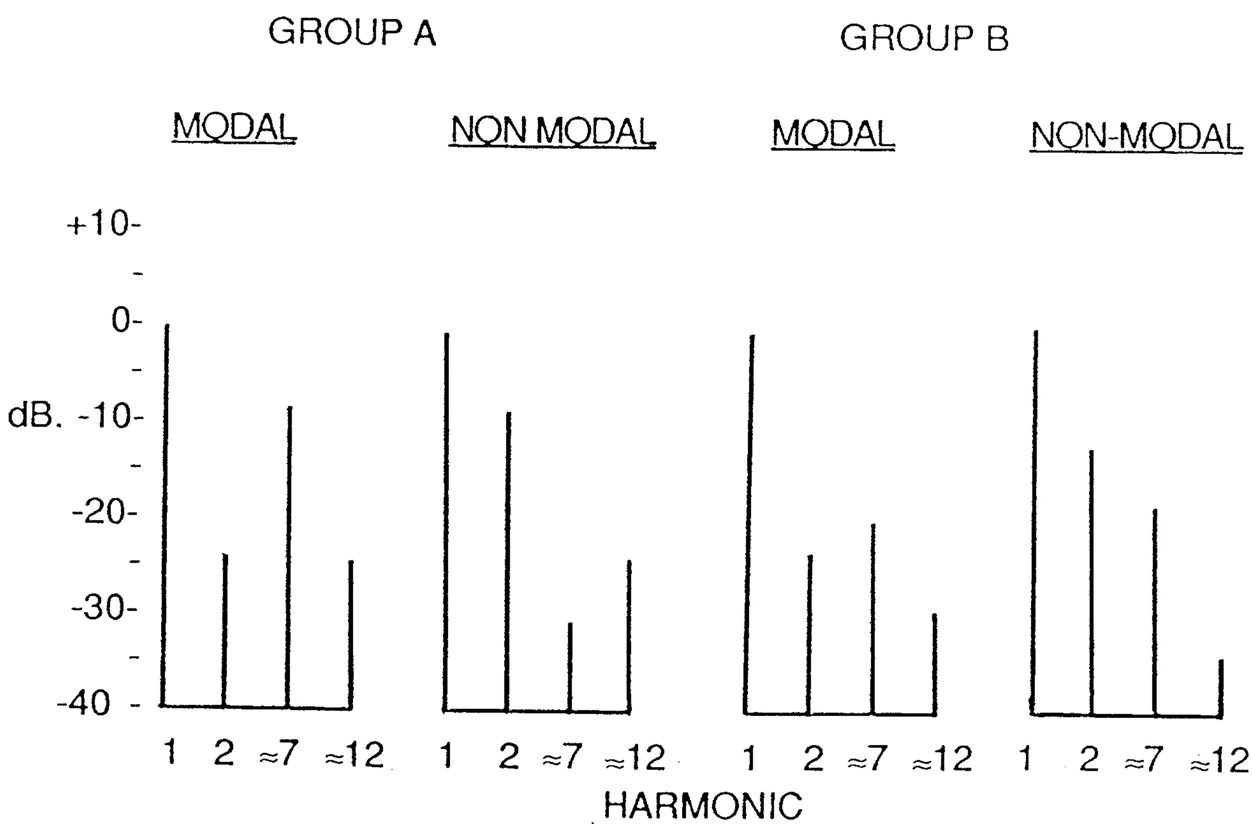


Figure 5. Schematic line spectra for modal and non-modal nasals for Groups A and B, female subjects.

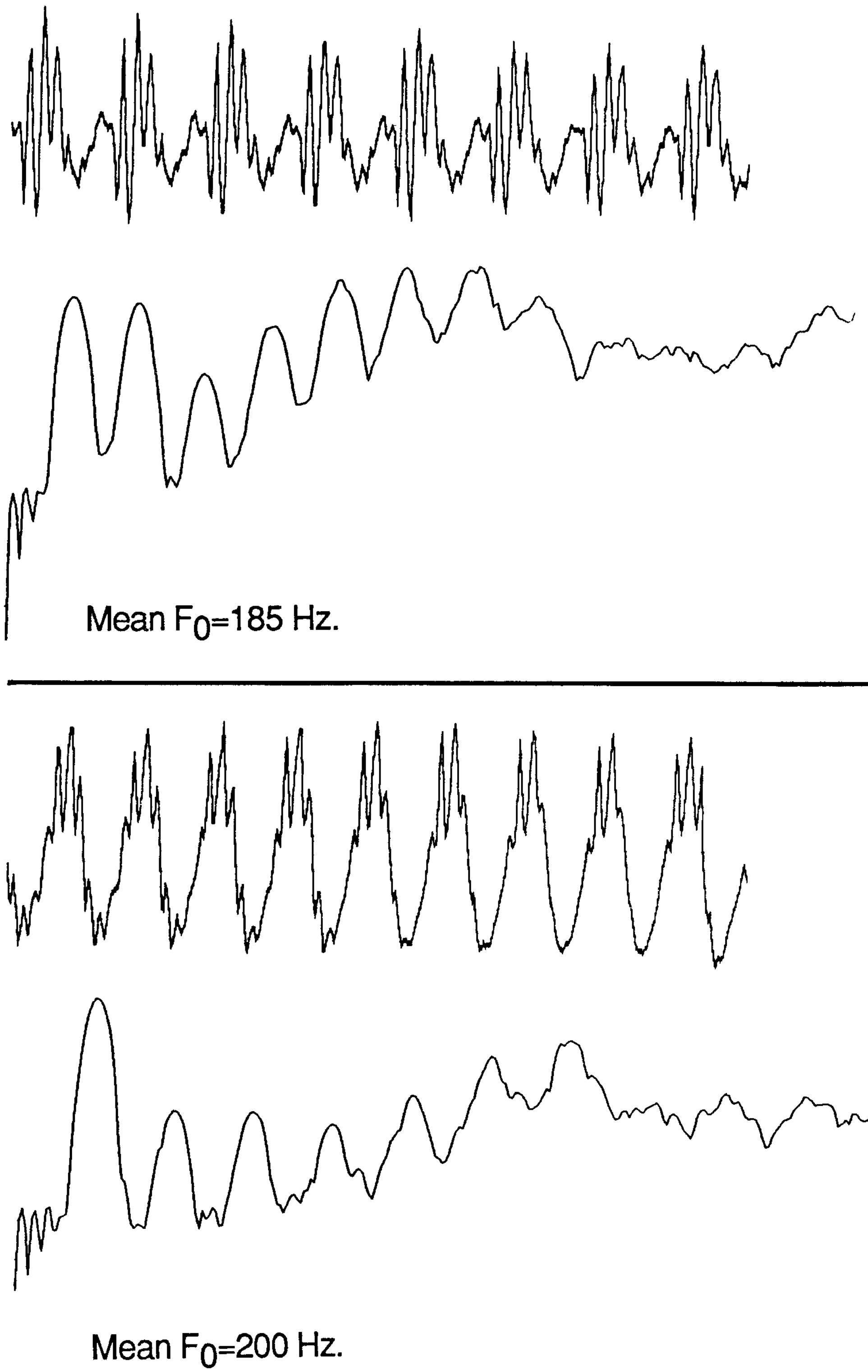


Figure 6 Waveforms and spectra illustrating that the subject below has a breathier phonation during modal nasals than the subject above.

Table 9. Amplitude differences between modal and non-modal nasals at three points in their spectra , for Groups A and B, male subjects.

	Tilt 1		Tilt 2		Tilt3	
Group	A	B	A	B	A	B
$H1\text{-}H_n$ difference (dB)	6	10	1	10	5	10

Group B subjects make a larger difference (by a factor of about 2) between their modal and non-modal nasals than do those in Group A. But in both Group A and Group B the difference between the first harmonic and higher harmonics is reduced throughout the spectrum during non-modal nasals. This means that for both groups there is relatively more energy in the fundamental during the non-modal nasals although in absolute terms this is manifested in prominently different ways acoustically and this presumably underlies the auditory distinction between the groups.

The tilt measured at vowel onset reverses the group trends in Table 9: Group A subjects use a much greater (by a factor of 3) spectral tilt adjustment than Group B subjects, with relatively more intensity in H1 following a non-modal nasal. However, in the rest of the vowel this difference is lost, suggesting that neither group has any notable phonation type distinction persisting in the later part of the syllable (although, as noted, certain tonal effects do persist).

The spectra for female subjects do not closely resemble those of their male counterparts in these respective groups. An examination of Figure 8 suggests that one reason for this is the strong first harmonic in the spectra for modal nasals and since this is the reference in these diagrams its effect will influence the the appearance of the rest of the spectrum. Two individuals in Group A and one in Group B are responsible for this effect; the remaining subjects do not have it. A plausible explanation for this exaggerated intensity of the first harmonic would be that it is reflecting a relatively high F_0 in the vicinity of the first formant of the modal nasal and the fact that the effect is attenuated when F_0 is lower during non-modal nasals would tend to support this. But this leaves the remaining cases unexplained. Their F0's are comparably high (they are only 11-15 Hz lower) yet they do not show a similar effect of the nasal formant. It seems therefore that an additional factor must be involved and it would be reasonable to suggest that some subjects have a relatively strong first harmonic anyway during modal phonation. This means that some female subjects have breathy modal voices. The waveforms and spectra in Figure 6 illustrate these differences for two cases.

The relative amplitudes of the higher harmonics in Figure 5 suggest that Group B non-modal nasal spectra roll off more evenly then the modal nasals and that Group A subjects have slightly more intensity higher in the spectrum for modal nasals compared to non-modal nasals. The difference in relative amplitude of the harmonics at about 1.4 kHz - approximately the seventh harmonic - was identified by ANOVA as significant. At this point in the spectrum the Group B female subjects resemble their male counterparts in having greater intensity during non-modal nasals, than the Group A subjects. However this parallel is not found with the harmonic above 2 kHz. The small number of tokens precludes any convincing interpretation of these differences in spectral tilt. At best one can say that there is an acoustic basis to the auditory difference between the groups but there is only a tenuous similarity between corresponding male and female groups. Finally, the female groups differ in the tilt at vowel onset. Whereas Group B subjects have virtually no noticeable distinction between vowel onsets preceded by modal and non-modal nasals, Group A subjects do, and it is in the direction of a reduction in the difference between H1 and H2 when a non-modal nasal precedes, as was the case with Group A male subjects.

5. Discussion

The focus of the preceding section has been to provide an adequate phonetic description of the distinction between modal and breathy voiced nasals in Tsonga and to identify which measures can be used to characterize the speaker variation in the production of this distinction.

The correct phonetic description of the distinction has been addressed in the literature on Tsonga, but answers have not been based on any objective investigation. The second issue has not been raised at all. The literature on Tsonga cited earlier claims that the non-modal nasals require a wider glottal opening than their modal counterparts and that this phonatory setting persists into the onset of the following vowel and even for the duration of this vowel, giving rise to breathy voice or murmur. Transcriptions such as [mɦa̤la] (Louw 1983) are explicit about the sequence of events that are supposed to be associated with a breathy voiced nasal and following vowel (the omission of the diaresis below [m] is surely unintentional or reflects a phonological bias) and an objective investigation of these segments should attempt to quantify them in terms of both nasal and vocalic variables. The aerodynamic measures that have been used on six subjects show that for some subjects modal and breathy voiced nasals indeed differ in terms of airflow during the nasal and in the onset of the adjacent vowel. Moreover the highly significant laryngographic amplitude reduction for the breathy voiced nasals is consistent with the interpretation that a glottal adjustment involving some degree vocal fold abduction is involved (Esling 1978), a situation that would be consistent with the observed increased airflow through the nasal cavities. These findings receive support from the measures of spectral tilt during nasals and vowel onsets and they confirm the correctness for some speakers of Louw's transcription above. The aerodynamic measures also reveal that nasal flow increases with time and that breathy voiced and modal nasals differ significantly only after their mid points. It seems clear , therefore, that for certain speakers the glottal gesture associated with breathy voicing has the latter part of the nasal and the onset of the vowel as its domain.

If one examines the comparative data in Table 10 below it seems likely that breathy voiced nasals arose through partial progressive assimiliation to the aspiration of a following voiceless stop:

Table 10. Comparative evidence

Tsonga	Zulu	Sotho	
m̤ala	impala	phala	sp. antelope
m̤isi	impisi	phiri	hyena

It is tempting to suggest that the phonetic details of breathy voicing in non-modal nasals in Tsonga are in fact a phonetically transparent reflex of earlier phonological aspiration. Certainly, the aerodynamic results presented so far would support such a conclusion. However the situation is not straight forward. It is complicated by two details, the first of which involves the innovative depressor effects of breathy voiced nasals and the second which arises from the consideration of individual differences in the pronunciation of these segments.

Non-modal nasals, like other depressor consonants in Tsonga, have local phonetic effects on the following tone and in addition they block the tone from spreading (incidentally, along with a phonetically heterogenous class of other segments (Louw 1968, p. 95)). While the origins of the tonal depression could be seen as a natural development from the phonation type innovation, the extent of the effects shown in Figure 3 are so striking as to represent an exaggeration or phonologization of the tendency for breathy voicing to co-exist with a decrease in F_0 (Laver 1981). The extent of this exaggeration can be appreciated if one compares the 35 Hz lowering of a high tone and the 22 Hz lowering of a low tone induced by breathy voiced nasals in Tsonga with, for example, the 12 Hz pitch lowering associated with lax voicing in Jingpho reported by Maddieson and Hess (1986). Incidentally, the duration of the effect (as can be seen in Figure 3) means that Tsonga is an exception to the claim of Stevens *et al.* (1986) that tone and pitch accent languages limit the duration of redundant consonantally induced pitch peturbation to very short intervals so as not to "...disturb the properties of other phonemic distinctions" (p.458). If the existence of such a prominent and reliable phonetic feature is linked to the observation that the non-modal nasals of some individuals do not have the same phonation characteristics as those of others, it might be suggested that in Tsonga the only reliable phonetic distinction among nasals is between those that lower tones and those that do not, rather than primarily involving a phonation type contrast with accompanying secondary tonal effects. The trends reflected in the spectra of Figures 4 and 5 suggest that this is not a complete picture of the case. Group B subjects are like Group A subjects in having particular spectral differences between modal and non-modal nasals that are not reducible to an effect of F_0 but which are consistent with a phonation contrast. It is true that F_0 changes in some female subjects have most likely affected the amplitude of the first harmonic through the influence of the nasal formant. But the differences observable between modal and non-modal nasals at higher frequencies are immune to this effect and are therefore probably reflecting a phonation contrast.

We may now address the question concerning the different phonetic bases of the phonation contrasts that distinguish the groups. Recall that in the Introduction a hypothesis was offered that predicted that the differences should lie along a single phonetic dimension, in this case one that can be labelled "degree of glottal constriction". The expectation would be that one of the groups might utilize one pair of values on this dimension for the modal/non-modal contrast whereas the other group would utilize a different pair of values. Are the spectra in Figures 4 and 5 compatible with this? Consider the male subjects first. The modal spectrum of Group A differs from that of Group B. Group B has a clear tilt upward in the higher frequencies with the first harmonic remaining weaker than all the others. Group A has less intensity throughout the spectrum and less of an intensity difference between the first and second harmonics. The conventional interpretation of these facts is that Group A males have breathier modal voices than Group B males. The non-modal spectra differ from the modal ones firstly by showing a reduction in the relative amplitude difference between first and second harmonic. This has been proposed as a reliable mark of breathy voicing (e.g. Ladefoged 1983) and since the groups do not differ significantly on this measure it would be reasonable to conclude that both sets of non-modal nasals are breathy voiced. At high frequencies in the spectra where the groups do differ significantly, relative similarities between the groups can be observed. The Group B non-modal spectrum continues to show a reduction in the difference in amplitude between the first and higher harmonics relative to modal voice, and there is generally far more energy higher in the spectrum than found for Group A. Group B shows a significant drop in energy at around 1.4 kHz but the harmonic above 2 kHz has the same intensity in both modal and breathy voice. However the overall envelopes for A and B are the similar. Figure 7 displays these differences and similarities clearly.

Figure 7 reveals very effectively why Group A and B non-modal nasals should sound so different: the overall spectral energy for Group B non-modal nasals exceeds that of Group A modal nasals and the spectram in the high frequencies for Group B non-modal nasals is faltter than that of Group A's modal nasals. Although these are large acoustic differences which give rise to prominent auditory effects it is nevertheless possible to interpret Figure 7 as representing two groups of speakers producing the same linguistic contrast in nasals in relatively the same way (but not to the same degree) on the phonetic dimension of glottal constriction. These facts about Tsonga nasals would thus support the Kirk *et al.* (1984) hypothesis about the organization of individual phonetic differences involving phonation contrasts.

Some problems nevertheless remain. Firstly it is not clear what Group B subjects are doing to achieve their relatively larger spectral adjustments between modal and non-modal nasals (see Table 9) without sounding breathy voiced. Secondly, the limited laryngographic and flow data that bear on this show that for one Group B subject, modal and non-modal nasals do not differ in terms of vocal fold abduction and flow, whereas for two Group A subjects the folds do abduct and flow increases significantly to produce breathy voicing during non-modal nasals. It remains possible therefore that different articulatory strategies underly the relative acoustic similarities between the groups.

Another difference between the groups, althought not a statistically significant one involves the spectral tilt at vowel onset. Group A subjects have a prominent modal/breathy difference of 10 dB, but Group B subjects have a milder one of 5 dB. It has been noted that these spectral differences do not persist during the remainder of the vowel. One may conclude that there is indeed a period of breathy voicing following breathy voiced nasals, but that this is prominent in some speakers and not in others.

The schematic spectral envelopes for the female subjects appear in Figure 8 below.

As has already been noted, these differences at best only partially resemble those of the male subjects, and on the basis of this evidence it seems unlikely that the female subjects are making the phonation contrast in the same way as the male subjects. Difficulties arise when one attempts to infer phonation differences from Figure

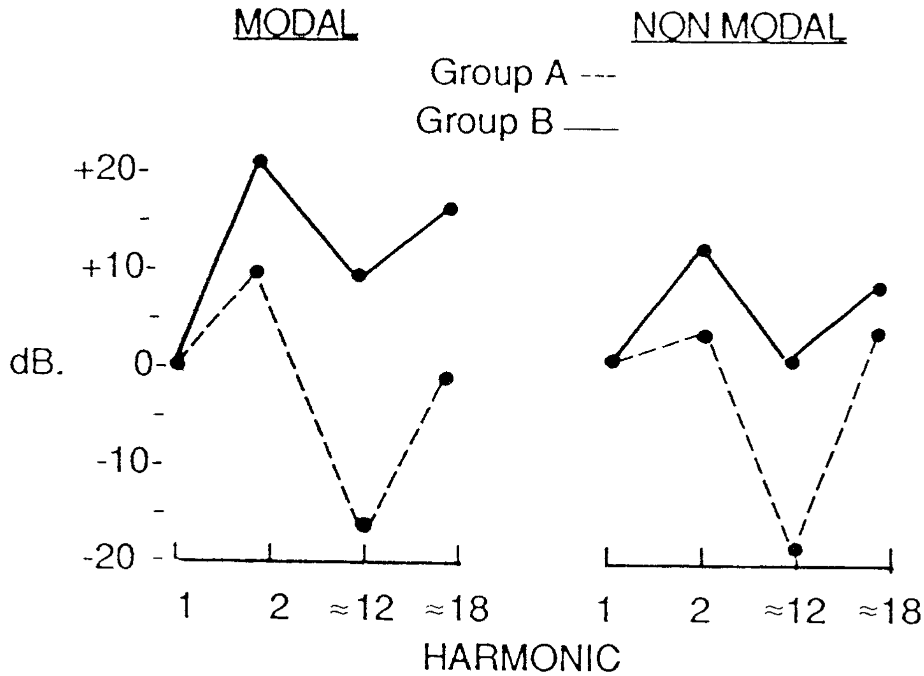


Figure 7. Envelopes for modal and non-modal nasals for Groups A and B male subjects .

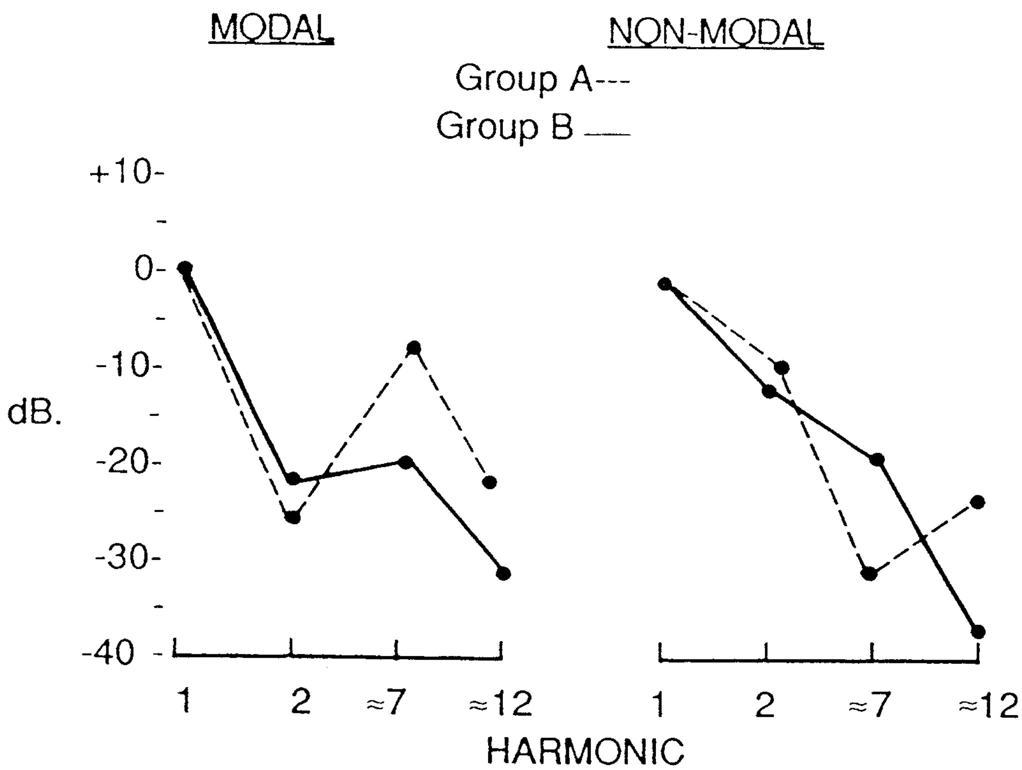


Figure 8. Envelopes for modal and non-modal spectra for Groups A female subjects.

8. The reduction in the difference between H1 and H2 for the non-modal nasals of both groups would normally be taken as consistent with a more breathy phonation. The confounding factor is the extremely strong first harmonic found in the modal nasals which equally suggests breathy (indeed even breathier) voice. As we have seen, there does not seem to be a single explanation for this strong H1 found for some but not all female subjects. Since these subjects do not sound as if their modal nasals are breathier than their non-modal nasals we must conclude that the interpretation of the H1-H2 measure for these modal nasals remains obscure. If the groups are compared across phonation types it is clear that there is an acoustic basis for their auditory difference. At 1.4 kHz the envelopes for non-modal nasals for the Group A female subjects show a relative drop in intensity similar to that seen with Group A male subjects. This may account for the perceptual resemblance between them as against Group B subjects. But parallells between the sexes are not preserved in the Group B. For convenience, the sex and group differences at higher points in the spectrum are summarized in Figure 9.

It seems that the Group B females are an exception to the Kirk *et al.* hypothesis because it shows that not all Tsonga speakers make the modal/non-modal phonation contrast in relatively the same way. However one categorized the difference in terms of phonological features it would be the case that Group B subjects preserved the contrast and even the same physical scale for its implementation as Group A subjects, but mapped the contrast in the reverse way so to speak. Of course, this interesting possibility should be treated with the caution appropriate for the small corpus of data on which it rests.

The female groups also differ at vowel onset on the measure signal to noise ratio (SNR). Group B has a lower SNR following non-modal nasals than Group A. A sensible interpretation of this, namely, that Group B's vowel onsets are noisier because of breathy voicing following a non-modal nasal, is complicated by other conflicting evidence. Firstly, the tilt measure at vowel onset suggests that Group A maintains the adjustment for breathy voicing into the vowel but Group B does not. Secondly, three of the four Group A subjects participated in the aerodynamic investigation and they produced significantly high flows at vowel onset following non-modal nasals. Given these facts it is not clear how the SNR measure should be interpreted. This is unfortunate because this measure and Nasal tilt 2 model the difference between the female groups most successfully and this failure to interpret one of the components means that a coherent picture of the phonetic basis of the difference cannot be offered.

In summary, distinctive speaker differences observed in the pronunciation of the modal/non-modal nasal contrast are relatively invariant for male subjects. Certain measures of spectral tilt during nasals reliably discriminate between the two groups of subjects, showing that they differ in terms of their relative degree of breathiness. Group A subjects locate their contrast on the scale of glottal constriction at the less constricted end while Group B subjects use values corresponding to greater constriction. There is no evidence that the salient pitch perturbation associated with non-modal nasals is a factor in the group differences since all subjects (male and female) make the same pitch adjustments. The phonetic behaviour of the female subjects has been less easy to interpret. Speaker differences resembling those of the male counterparts can be observed but the parallels are incomplete. Althought the data are limited, a significant difference in spectral tilt during modal and non-modal nasals exists. The non-modal

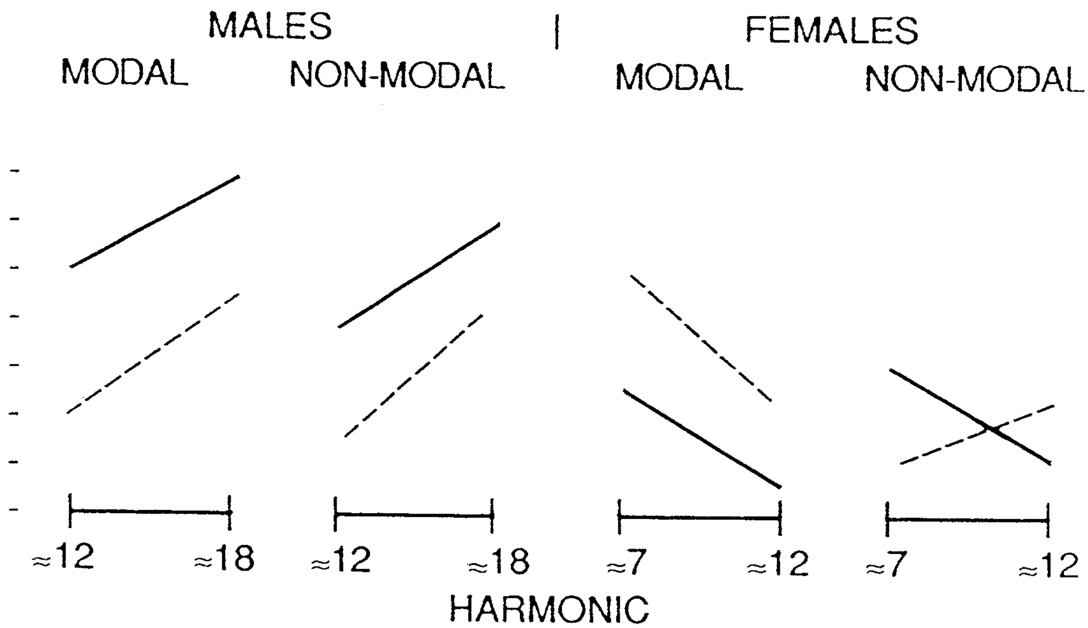


Figure 9. Schematic spectra of Tilt 2 and 3 showing Group A (dashed line and Group B (solid lines) differences for modal and non-modal nasals for all subjects.

nasals of group B subjects evidently sound different from those of Group A subjects because there is more intensity in the spectra of the former in the vicinity of 1.4 kHz. In this respect male and female subjects resemble one another. Group A female subjects preserve the spectral similarity with Group A males more generally by having less intensity in the higher frequencies compared to modal nasals. Group B female subjects show an unexpected reversal of this feature because their modal nasals have less intensity in the higher frequencies. Thus speaker variation in these female subjects does not appear to obey the constraint that that all speakers of a language will make the same phonation contrast in relatively the same way. Loosely speaking, what these subjects appear to be doing is just making some phonation contrast between the types of nasals.

An alternative view would be that these are cases where a primary phonation contrast has been restructured as a pitch distinction and therefore that these data are not relevant to the present hypothesis about constraints on speaker variation. Indeed, one should consider whether this suggestion could justifiably be extended to all subjects. Under such a proposal Group B females would not be exceptional since all speakers would be described as producing a primary and consistent pitch contrast between nasal consonants, with individuals varying in the phonation type accompanying the pitch adjustment. Pitch lowering can be achieved with different glottal settings so it would not be surprising to find modal voice or varying (subsidiary) phonation types or degrees of one phonation type associated with pitch lowering. This view obviously places a quite different perspective on the speaker variation in phonation type described above since there would no longer be an expectation that they should be structured according to a principle of relative invariance. On the contrary, any or no phonation type changes would be expected. There are two reasons for rejecting this demotion of phonation type to the level of an arbitrary and even gratuitous side effect of pitch lowering in Tsonga. Firstly, if it were correct we should expect to find cases where pitch lowering took place without any phonation change, i.e. with modal voice in both sets of nasals. This is never the case. Evidence from spectral measurements show that the non-modal nasals are always associated with spectral changes that are evidence for a phonation difference. Secondly, Tsonga speakers contrast higher and lower pitches independently of depressor consonants as part of the tonal system of their language, and when they do this, lower pitches are not associated with the phonation differences found with nasals. These observations show that both pitch and phonation changes are necessary in the pronunciation of nasal contrasts but the question of whether either is primary is left open.

Acknowledgements

This work was supported by a grant to Peter Ladefoged and Ian Maddieson from the Department of Health and Human Services for the Linguistic Use of Laryngeal and Aerodynamic Processes. Many members of the Phonetics Laboratory contributed in various ways to this study and their assistance is gratefully acknowledged. Particular thanks are due to Norma Antonanzas-Barroso, Peter Ladefoged, Jenny Ladefoged, Ian Maddieson and Rong Rong Liao.

References

Baumbach, E. J. M. (1974) *Introduction to the Speech Sounds and Speech Sound Changes of Tsonga.* Pretoria: van Schaik.

Cohen, J. & Cohen, P. (1975) *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences.* Hillsdale NJ: Lawrence Erlbaum Associates.

Esling, J.K. (1978) *Voice Quality in Edinburgh. A Sociolinguistic and Phonetic Study.* Unpublished PhD. thesis, University of Edinburgh, Scotland.

Huffman, M.K. (1987) Measures of phonation type in Hmong. *J. Acoust. Soc. Amer.* 81 (2): 495-504.

Jackson, M., Ladefoged, P., Huffman, M. K. & Antoñanzas-Barroso, N. (1985a) Measures of spectral tilt. *UCLA Working Papers in Phonetics* 61: 72-78. also abstracted in *Acoust. Soc. Amer.* 77 S1.

____. (1985b) Automated measures of spectral tilt. *UCLA Working Papers in Phonetics* 62: 77-88.

Kirk, P. L., Ladefoged, P. & Ladefoged, J. (1984) Using a spectrograph for measures of phonation type in a natural language. *UCLA Working Papers in Phonetics* 59: 102-113.

Laver, J. (1981) *The Phonetic Description of Voice Quality.* London: Cambridge University Press.

Ladefoged, P. (1983) The linguistic use of different phonation types. In *Vocal Fold Physiology: Contemporary Research and Clinical Issues* (D. Bless & J. Abbs, eds.), pp 351-360. San Diego: College Hill Press.

Ladefoged, P. & Antoñanzas-Barroso, N. (1984) Computer measurements of breathy voice quality. *UCLA Working Papers in Phonetics* 61: 79-86. also abstracted in *J. Acoust. Soc. Amer.* 75 S1.

Louw, J.A. (1968) *The Intonation of the Sentence and its Constituent Parts in Xhosa and Tsonga.* Pretoria: Human Sciences Research Council.

Louw, J.A. (1983) Some tone rules of Tsonga. Afrika und Ubersee, Band LXV1, 13-24.

Maddieson, I. & Ladefoged, P. (1985) 'Tense' and 'lax' in four minority languages of China, UCLA Working Papers in Phonetics, 60, 59-83. also abstracted in *J. Acoust. Soc. Amer.* 75 S1.

Maddieson, I. & Hess, S. (1986) "Tense" and "lax" revisited: more on phonation types in minority languages of China. *UCLA Working Papers in Phonetics* 63: 103-109. also presented at the 19th Conference on Sino-Tibetan Languages and Linguistics.

Milenkovic, P. (1986) Least mean square measures of voice perturbation. Department of Electrical and Computer Engineering, University of Wisconsin-Madison.

SAS Institute Inc. (1982) *SAS User's Guide: Statistics.* Cary NC: SAS Institute Inc.

Stevens, K.N., Keyser, S.J. & Kawasaki, H. (1986) Toward a phonetic and phonological theory of redundant features. In *Invariance and Variability in Speech Processes* (J.S. Perkell & D.H. Klatt, eds.) pp. 426-463, London: Lawrence Erlbaum Associates.

Phonation types in Mon-Khmer languages

Theraphan L. Thongkum
Chulalongkorn University

1.Introduction

Southeast Asian (SEA) languages can be classified typologically into three types: (a) tone languages, (b) register languages, and (c) neither tone nor register languages. In tone languages, the pitch by itself may distinguish the lexical meanings of words. A register language may be defined as a language that has a lexically contrastive register complex (a combination of phonation type, pitch, vowel length, vowel quality, etc.). The term register is used by Henderson (1952) as a phonological concept. It is a cover term not only for laryngeal activity but also for a cluster of activities in the vocal tract. Phonation-type language can be an alternative term.

In this paper, only the second type will be considered, since phonation type plays an important role. A large number of Mon-Khmer languages, such as Phalok, Wa, Chong, Mon, Bru, Kui, So, Nyah Kur, Western Khmer and so forth, have two lexically contrastive phonation types: clear (normal, modal) voice vs. breathy voice. The following examples are drawn from a Western Khmer dialect, that has been investigated by Gérard Diffloth and myself. The data are from my fieldnotes collected in September, 1985. Register distinction has been lost in most Khmer dialects. This dialect is very conservative, and therefore of special interest to Mon-Khmer comparativists.

	Clear voice		Breathy voice
pəmat	'gall bladder'	ma̤t	'mouth'
phlow	'pathway'	phlo̤w	'thigh'

Even though quite rare, some languages can have a three-way or four-way contrast of phonation type, for instance, clear voice vs. creaky voice vs. breathy voice in Didrá (Gregerson and Smith, 1973: 145), clear voice vs. clear followed by creaky voice vs. breathy voice vs. breathy voice followed by creaky voice in Chong (Thongkum, in press).

Some languages that make phonation-type distinctions lexically also do so morphologically. A change of phonation type can cause a shift of grammatical function; for example, in some cases, the clear phonation in Kui may relate nouns and verbs which originally have breathy phonation:

phɛʔ	'mush-like food'	phɛ̤ʔ	'to be mushy'
luum	'a bite, a mouthful'	lṳum	'to eat by stuffing whole chunks of food into one's mouth'

In some tone languages of the area certain tones are associated with a given phonation type: for instance, the mid-falling tone in white Hmong (e.g. tɔ̤^{31} 'stool') and the low-rising tone in Tai Wang (e.g. ta̰a^{313} 'eye'), which are always accompanied by breathy phonation and creaky phonation, respectively.

Many scholars of SEA linguistics have written about register and its phonetic exponents, e.g. Henderson (1952, 1965, 1977), Shorto (1967), Smith (1968), Glover (1971), Egerod (1971), Pittman (1978), Matisoff (1973), Gregerson and Smith (1973), Jenner (1974), Gregerson (1976), Huffman (1976), Weidert (1978), Ferlus (1979) and Thurgood (1980). Gregerson (1976) points out the need for experimental research on the subject of register in SEA languages. Putting emphasis on what is going on in the vocal tract, i.e. the tongue-root position, Henderson (1977), unlike Gregerson, suggests that register phenomena are more associated with laryngeal activities. She also draws attention to different kinds of phonation type that play a linguistic role.

In this paper I will present a summary of my instrumental studies of the phonetic manifestation of phonation-type distinction in three Mon-Khmer languages: Nyah Kur (Chao Bon), Kui (Suai) and Chong. (See Thongkum 1982, 1985 and in press for details.)

The phonetic instruments used in the studies are as follows:

- Kay Sona-Graph 6061-B;
- Fundamental Frequency Meter, type FFM 650 (F-J);
- Intensity Meter, type IM 360 (F-J);

- Electro-glottograph, type EG 830 (F-J);
- Electro-aerometer, type EA 510/4 (F-J);
- Mingograf 34 T (Siemens AB).

2. Definition of "phonation type"

Different "modes of vocal fold vibration" cause different phonation types. Catford (1964: 30-35, and 1977: 95-101) classifies phonation types in terms of the type of stricture and the location of stricture. Phonation stricture type consists of breath (voiceless), whisper, voice, creak and stop. These five major phonatory stricture types can occur in many combinations: breathy voice, whispery voice, whispery creak, voiced creak (creaky voice), whispery voiced creak. There are four major locational types: glottal (full glottal), ligamental (anterior), arytenoidal (posterior) and ventricular. The laryngeal modifications of phonation comprise vocal fold modification, upper larynx constriction, and vertical displacement of the larynx.

Abercrombie (1967: 26) distinguishes four states of the glottis; open glottis (breath state), glottis in vibration (voice state), narrowed glottis (whisper state), and closed glottis.

Ladefoged (1971: 8) describes seven states of the glottis, namely: voice, voiceless, aspiration, murmur (breath voice),laryngealization (creaky voice), glottal stop, and whisper.

Sprigg (1978: 3-17) defines phonation types in terms of the activity of the vocal cords and the posture of the glottis. The term activity consists of two systems: voicing and trillization, which enter into various combinations with each other and with the terms of posture. The voicing system comprises voice and voicelessness, whereas the trillization system is classified into glottal trill and clear. The various shapes or postures of the glottis are breath, closed arytenoidal, approximated and ligamental.

Laver (1980: 109-140) classifies modes of laryngeal vibration in phonatory settings into two major types: simple and compound phonation types. The former consists of modal voice, falsetto, whisper, creak, harshness, and breathiness. The latter consists of whispery creak, whispery voice, whispery falsetto, creaky voice, creaky falsetto, whispery creaky voice, whispery creaky falsetto, breathy voice, harsh voice, harsh falsetto, harsh creak, harsh whispery voice, harsh whispery falsetto, harsh creaky voice, harsh creaky falsetto, harsh whispery creaky voice, and harsh whispery creaky falsetto.

Abercrombie's classification of the states of the glottis is too broad. This is because he aims only at describing segmental features, i.e. consonants and vowels. Breathy phonation and creaky phonation are discussed separately under the features of voice quality and of voice dynamics. As a result, his presentation of phonation types lacks unity. The terms and definitions proposed by Catford, Sprigg and Laver are very complicated and too difficult to cope with. They may be useful to phoneticians who specialize in artificial speech and speech pathologists who work on abnormal speech, but for the linguistic investigation of natural speech, I feel that they are rather unnecessary.

As for register distinction in SEA languages, only three kinds of phonation type need be distinguished: clear (or normal or modal) voice, breathy voice and creaky voice. Thus, the seven states of the glottis proposed by Ladefoged are adequate for linguistic purposes, at least for describing and explaining phonation-type phenomena in SEA languages.

Catford (1977: 100) gives a definition of VOICE as follows: "glottis closed, and vocal folds subjected to varying degrees of tension, such that they vibrate (at volume-velocities from about 50 cm^3/s upwards and from about 2 to 3 cm H_2O up to about 30 cm H_2O subgoottal pressure), emitting periodic high-velocity puffs of air, generating the periodic sound known as 'voice'." This refers only to 'chest voice', which is the basic form of voice.

In producing BREATHY VOICE, the arytenoids are apart while the ligamental vocal cords are vibrating (Ladefoged, 1971: 8). Laver (1980: 132-5) points out that in comparison with modal voice, breathy voice is always accompanied by slight audible friction, high rate of air-flow and low degree of laryngeal effort, and that some of the acoustic energy is also lost by the damping effect of the general relaxation of muscles of the whole vocal system of lax voice, of which breathy voice is almost always a component.

The terms glottal fry, vocal fry, CREAKY VOICE and laryngealization have been used more or less synonymously. When laryngealized sounds are produced, "the arytenoid cartilages are pressed inwards so that the posterior portions of the vocal cords are held together and only the anterior (ligamental) portions are able to vibrate. The result is a harsh sound with a comparatively low pitch" (Ladefoged 1971: 14-15).

3. Phonation type and its phonetic correlates in Nyah Kur (Chao Bon), Kui (Suai) and Chong

This section presents a summary of my instrumental analysis of Nyah Kur, Kui and Chong vowels in respect to fundamental frequency, formant frequency, duration, intensity, power spectra, laryngeal waveforms and airflow. These represent the phonetic realization of phonation type distinctions in the three languages investigated.

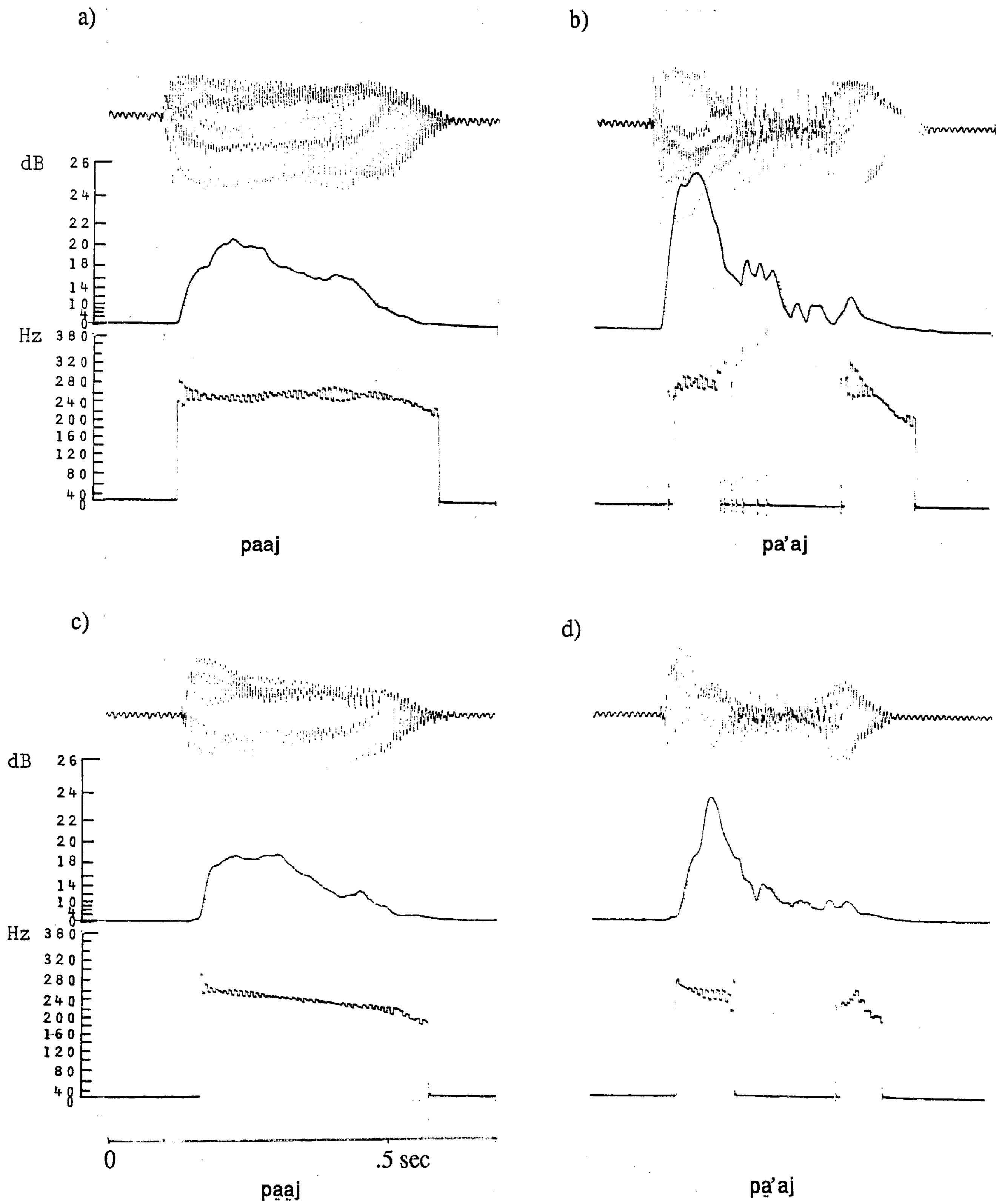
3.1 Fundamental frequency

In Nyah Kur and Kui breathy voice vowels have lower F_0 than clear voice vowels in all types of syllable structure. F_0 and F_0 contour in Chong (see Figure 1) are somewhat different, since Chong has four distinctive phonation types: clear voice vs. clear-creaky voice vs. breathy voice vs. breathy-creaky voice; for example, **paaj** (clear voice) 'to throw', **pa'aj** (clear-creaky voice) 'palm of the hand and foot', **pa̤a̤j** (breathy voice) 'to wear across', and pa̤'aj (breathy-creaky voice) 'two'.

Although Chong breathy vowels can be perceived auditorily as having the lowest pitch, their F_0 does not start low at all. From 0% up to 30% of the vowel duration, breathy vowels can have even higher F_0 than clear vowels. The point at which the higher pitch of clear vowels and the lower pitch of breathy vowels can be differentiated is from 30% up to the end of the vowel duration.

In Chong, only clear and breathy vowels can occur in syllables ending in /h/ or /ʔ/ (glottal stop). In CVh syllables, both types of vowel have rising F_0 contour; however, breathy vowels do not always have lower F_0 that clear vowels. In CVʔ syllables, clear vowels have rising F_0 contour with an abrupt fall at the end, whereas breathy vowels have falling F_0 contour.

As is shown in Figure 2, in syllables in which all four phonation types occur, clear-creaky vowels have the highest fundamental frequency. Only the clear part of these vowels could be measured due to the irregulatities of the glottal period in the creaky part. However, it is possible to conclude that clear-creaky and breathy-creaky



ure 1: Tracings of a) paaj, b) pa'aj, c) pa̱a̱j, and d) pa̱'aj in Chong.

Duplex oscillogram (top line)

Intensity curve (middle line)

Fundamental frequency curve (bottom line)

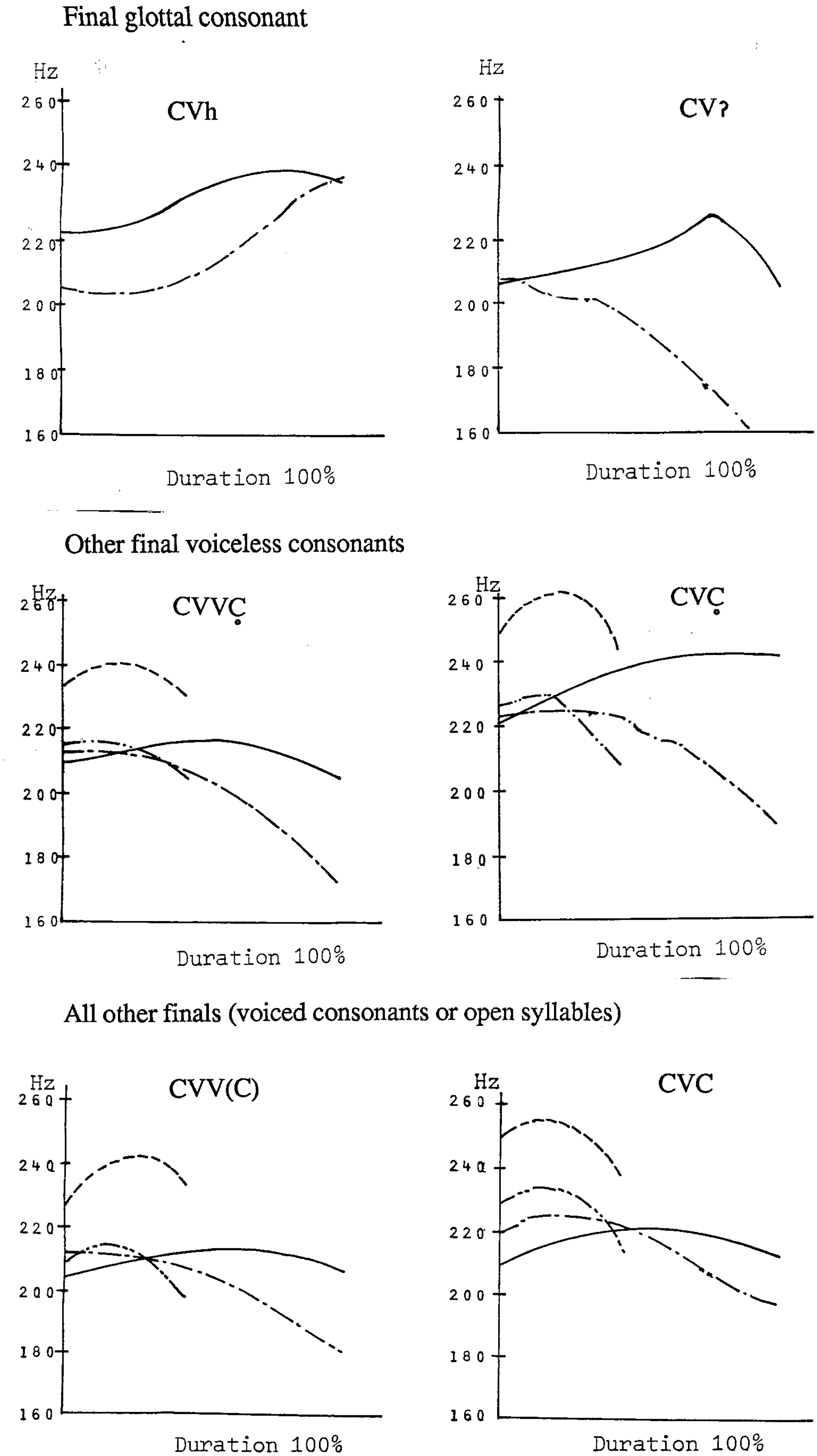


Figure 2: Mean F0 values (in Hz) of clear, clear-creaky, breathy, and breathy-creaky vowels in 6 syllable types in Chong.

——— clear ---- clear-creaky

—·— breathy —··— breathy-creaky

vowels have high rise-fall and high fall F_0 contours, respectively. The occurrence of laryngealization seems to be the cause of the falling F_0 contour. This may offer an explanation of how falling tones arise in some tone languages. The creaky phonation may be absent in some instances; as is shown in Figure 3; thus clear-creaky vowels become clear vowels with higher fall F_0 contour, and breathy-creaky vowels become clear vowels with lower fall F_0 contour.

3.2 Formant frequency

The acoustical measurements of short and long vowels in Nyah Kur and Kui confirm my auditory judgment that each pair of vowels, clear voice vs. breathy voice, have similar quality. The vowel formant charts do not exhibit any obvious patterns that breathy voice vowels are always more close or more open than their clear voice counterparts. The minor differences, although they do exist, are not systematic.

In Chong, however, breathy and breathy-creaky vowels have a lower F_1, and thus sound higher than clear and clear-creaky vowels. This finding supports Gregerson's hypothesis that in Mon-Khmer languages clear vowels which are produced with retracted tongue-root are always more open than breathy vowels which are produced with advanced tongue-root (Gregerson, 1976). The Chong vowel formant charts exhibit obvious patterns: clear voice and clear-creaky voice vowels are more open than the breathy and breathy-creaky voice vowels. However, there are alternative possible explanations. If the tongue-root positions were the same, but the larynx was always lower, F_1 would be lower, making the vowel appear more close. Breathy vowels in many Mon-Khmer languages are produced with lowering of the larynx (Henderson 1952, Gregerson 1976, etc.).

3.3 Duration

Fischer-Jørgensen (1977) and Kirk et al (1984), in their studies of Gujarati and Jalapa Mazatec, respectively, say that breathy vowels in those two languages have longer duration than clear vowels. Perhaps this tendency could be attested only in languages that have no distinctive vowel length. In my Kui data, although breathy short vowels have longer duration than clear short vowels, it works in the opposite way for long vowels. I suspect that differences in duration caused by differences in phonation types might not be important when the language investigated possesses phonological length. The results of the measurements of Chong vowels makes me even more confident of this fact.

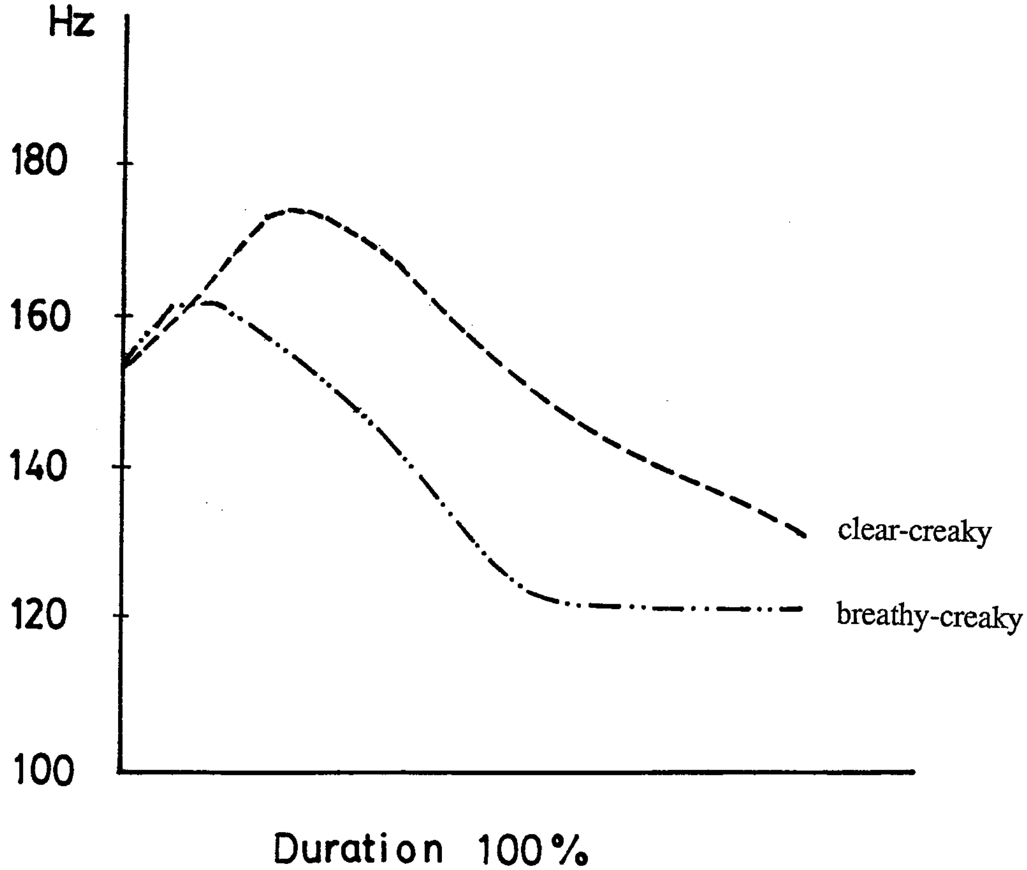


Figure 3: Mean F0 values (in Hz) of clear-creaky and breathy-creaky vowels in CVVC̥ syllable type when creaky phonation disappears and is replaced by falling F0 contour in Chong.

Regarding clear-creaky and breathy-creaky vowels, in Chong, the creaky part of breathy-creaky vowels is longer than that of clear-creaky vowels, although the overall duration of breathy-creaky vowels is shorter than that of clear-creaky vowels. It is also noticeable that breathy-creaky vowels are shorter than the other kinds of vowels no matter in what type of syllable.

3.4 Intensity

Clear vowels in the Kui data have higher amplitude than breathy vowels. The measurements of the overall intensity of Chong vowels agree very well with what has been found in Kui, i.e. clear vowels and clear-creaky vowels have higher amplitude than breathy and breathy-creaky vowels. The loss of intensity is due to leaking glottis during breathy phonation (Fischer-Jørgensen, 1977: 119), which is an inefficient method of transforming physiological into acoustic energy. Furthermore, the acoustic energy is lost by the damping effect of the general relaxation of the muscles of the whole vocal system in lax voice (Laver, 1980: 135), and by the loss of energy through the largely open glottis into the subglottal cavities.

Figure 4 shows that the intensity curves of clear and breathy vowels (on the left of the figure) look similar, i.e. more bell-shaped, in that the highest peak in amplitude is not much above those on either side; but those of clear-creaky and breathy-creaky vowels (on the right of the figure) look more cone-shaped, with a high peak amplitude and a sudden drop of intensity when clear and breathy vowels become creaky, caused by the abrupt closing of the vocal folds.

3.5 Power spectra

Kirk et al (1984) point out that the power spectra enable phoneticians to quantify the relative amount of energy in different harmonics. For their study of phonation types (clear vs. creaky vs. breathy) in Jalapa Mazatec, they measured the difference in dB between the intensity of the fundamental and the intensity of the first formant. They conclude, "There is considerable variation from speaker to speaker in the three phonation types; but for each speaker on this measure the value for creaky voice is less than that for modal voice, and the value for modal voice is less than that for breathy voice" (Kirk et al, 1984: 109).

Following their procedure, the same kind of measurements were made for Chong.

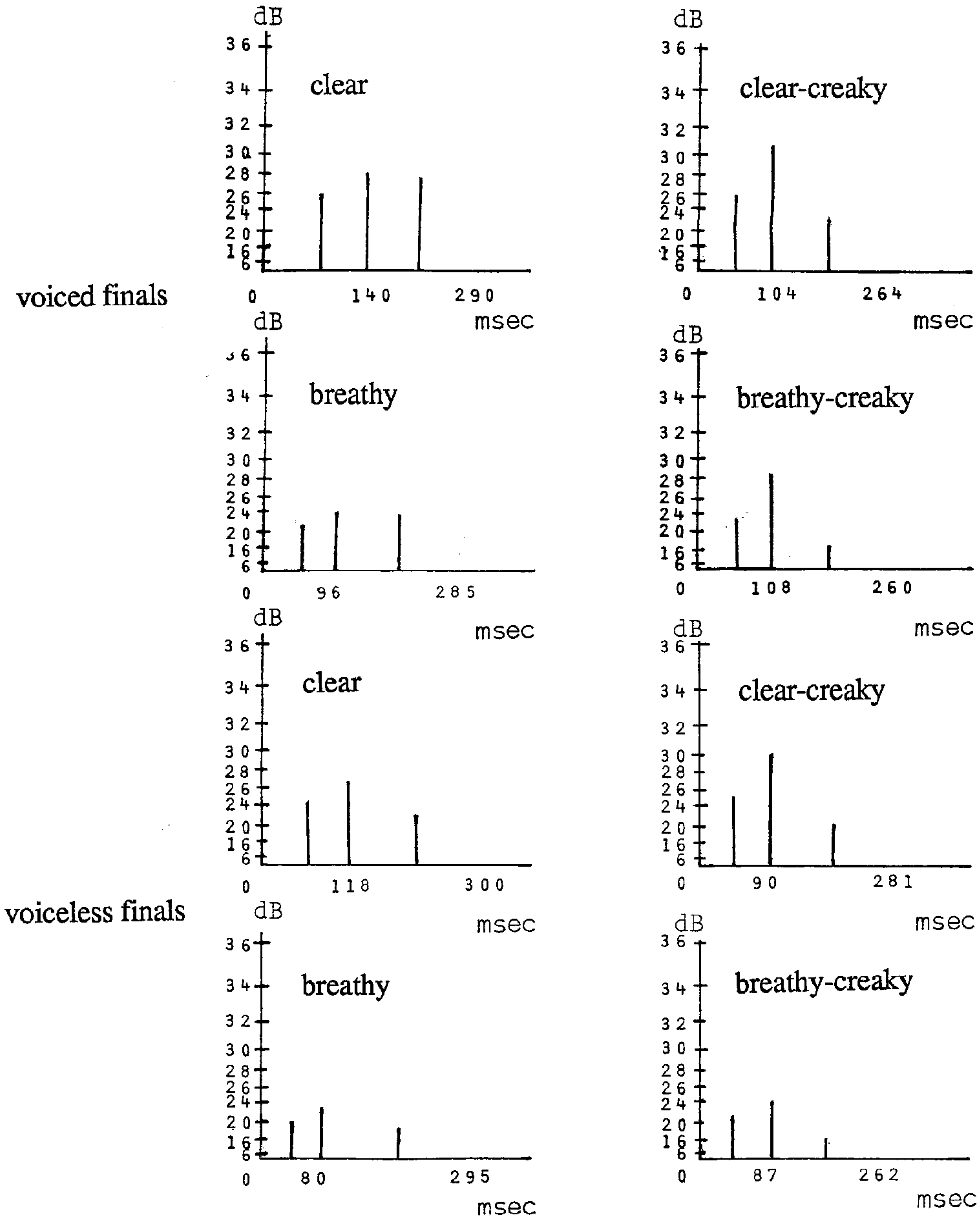


Figure 4: Intensity curves and mean amplitude (in dB) of clear, clear-creaky, breathy, and breathy-creaky long vowels in Chong.

Unfortunately, the results of the measurements did not meet my expectation; i.e. the measure did not separate out the phonation types successfully, in that the value for modal voice was higher than that for creaky voice which was not as expected, although it was less than that for breathy voice, as expected. The measure cannot be used for comparing phonation types in Chọng, because the vowels of each phonation category differ in quality. Moreover, Chong has two dynamic or combined phonation types: clear followed by creaky phonation and breathy followed by creaky phonation, which can cause problems for the measure.

3.6 Laryngeal vibration and airflow

No quantitative measurement was attempted, but as can be seen clearly in Figure 5, the laryngograms of the four sets of Chong vowels look different. At the onset of clear and clear-creaky vowels, there is a relative rise of the signal which may be produced by the raising of the larynx for normal voicing, and on the other hand, there is a relative fall of the signal at the onset of breathy and breathy-creaky vowels which may be caused by the lowering of the larynx during breathy phonation. There is also a relative sharp rise of the signal after the release of final consonant following clear-creaky and breathy-creaky vowels.

Regarding airflow, shown in Figure 6, the most prominent characteristic of Chong breathy vowels is strong airflow. This is due to the presence of an opening in the posterior part of the glottis and/or an increased activity of the expiratory muscles when breathy vowels are produced (Fischer-Jørgensen, 1967: 153). During the creaky part of the clear-creaky and breath-creaky vowels, there is a sudden drop of airflow caused by a rapid closing and tightening of the vocal folds.

4. Comments on Gregerson's hypothesis of tongue-root and register

Gregerson (1976) associates first register with open (clear or creaky) vowels and second register with close (breathy) vowels. He draws examples from several Mon-Khmer languages of Vietnam. He claims that Mon-Khmer register distinctions are based mainly on the advancement or retraction of the tongue-root, and that other phonetic parameters, such as larynx height, tongue height, phonation type, consonant voicing, pitch, and so forth, are somewhat secondary. All of these phonetic features are the effects of an underlying opposition between tongue-root advancement vs. tongue-root retraction (Gregerson, 1976: 338).

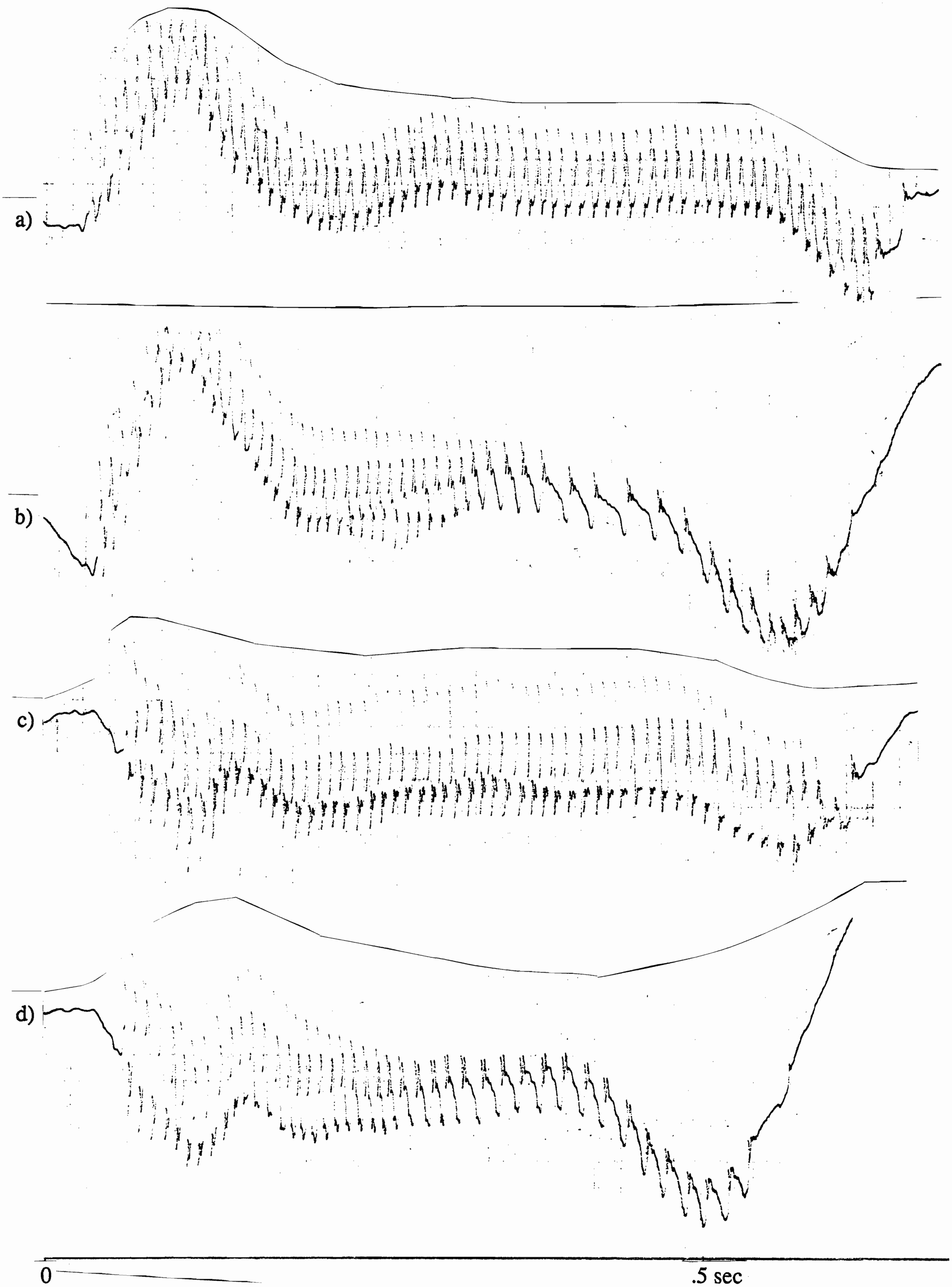


Figure 5: Laryngograms (laryngeal waveforms) of a) ceet, b) ce'et, c) pe̤e̤t, and d) pe̤'et in Chong.

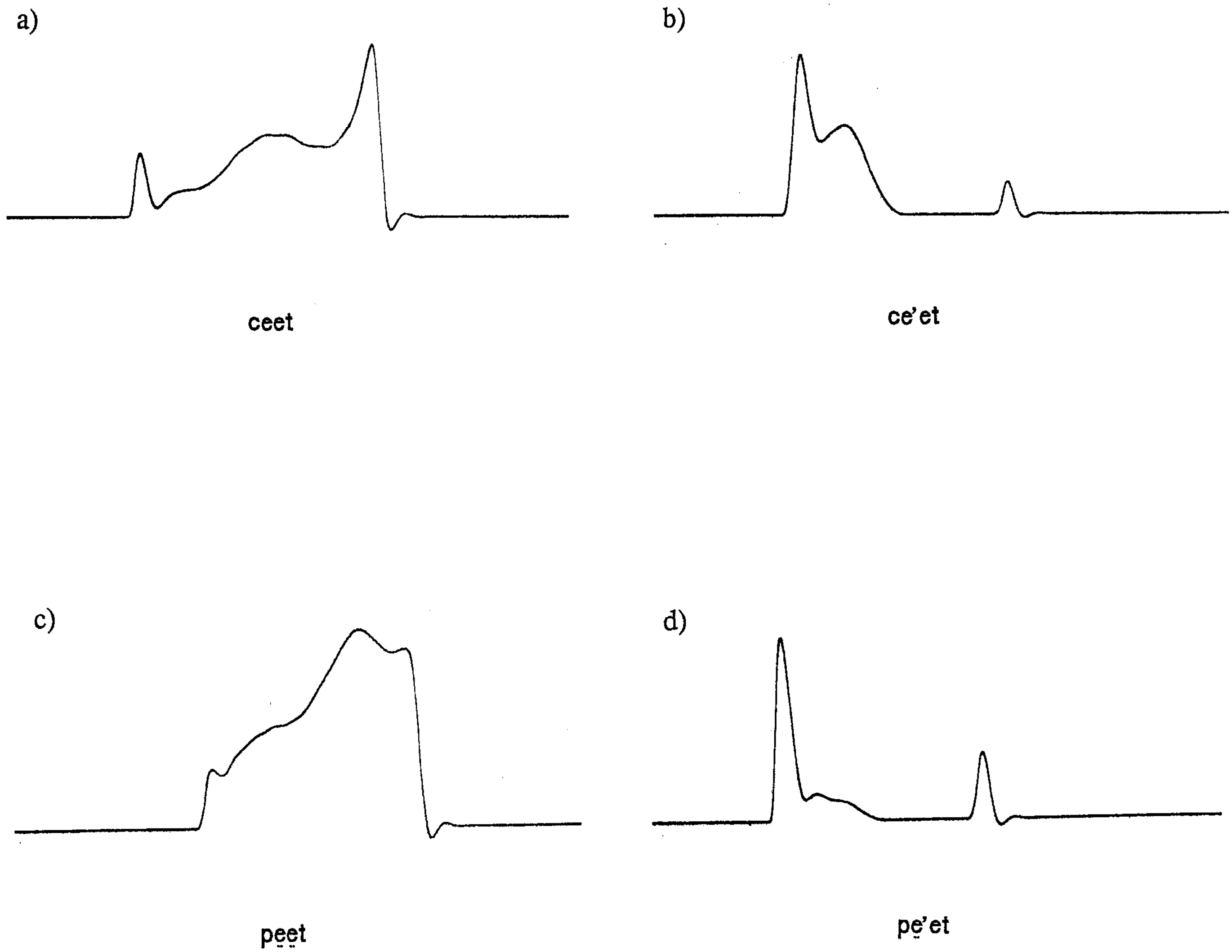


Figure 6: Airflow curves of a) ceet, b) ce'et, c) pe̤e̤t, and d) pe̤'et in Chong.

He also claims that the advancement or retraction of the tongue-root constitutes a major air-stream regulation. Although he recognized other phonetic features associated with register phenomena in Mon-Khmer languages, he seems to regard them as 'natural collaboration' notably tied to tongue-root articulation. As for pitch, he says:

> ... pitch is never a major feature. It is more often absent than present in register descriptions in various languages. This may be interpreted as some kind of support for the view that laryngeal activities are at least partially independent of tongue movement. (Gregerson, 1976: 354)

In my opinion pitch differences are always present in register languages. There is a closeknit connection between pitch and phonation type. Unfortunately, it has always been ignored or unheard by Mon-Khmer specialists.

Articulatorily, when a vowel sound is produced, the factors involved are:

- length of the vocal tract;
- vocal tract shape as affected by:
    - distance between the tongue body and the superior or posterior surface of the vocal tract;
    - distance between the tongue-root and pharyngeal wall;
    - posture of the lips; and
- laryngeal activities.

It seems to me that pairs of Kui and Nyah Kur vowels, clear voiced vs. breathy voiced, systematically differ from each other only in the laryngeal dimension; the rest of the differences are minor.

To test Gregerson's hypothesis, x-ray photographs of the Nyah Kur speaker were made in order to observe the movements and positions of his tongue when pronouncing seven pairs of Nyah Kur vowels. The speaker had a perfect profile for x-raying because he had no teeth; therefore, there were no molars to block the lateral view of the tongue. The radiological technique used in the experiment was fluoroscopy by image intensifier, together with a video recording system. The tracings of the x-ray photographs (Figure 7) show that for each pair of vowels, clear voice vs. breathy voice, there are no obvious differences in tongue-root positions.

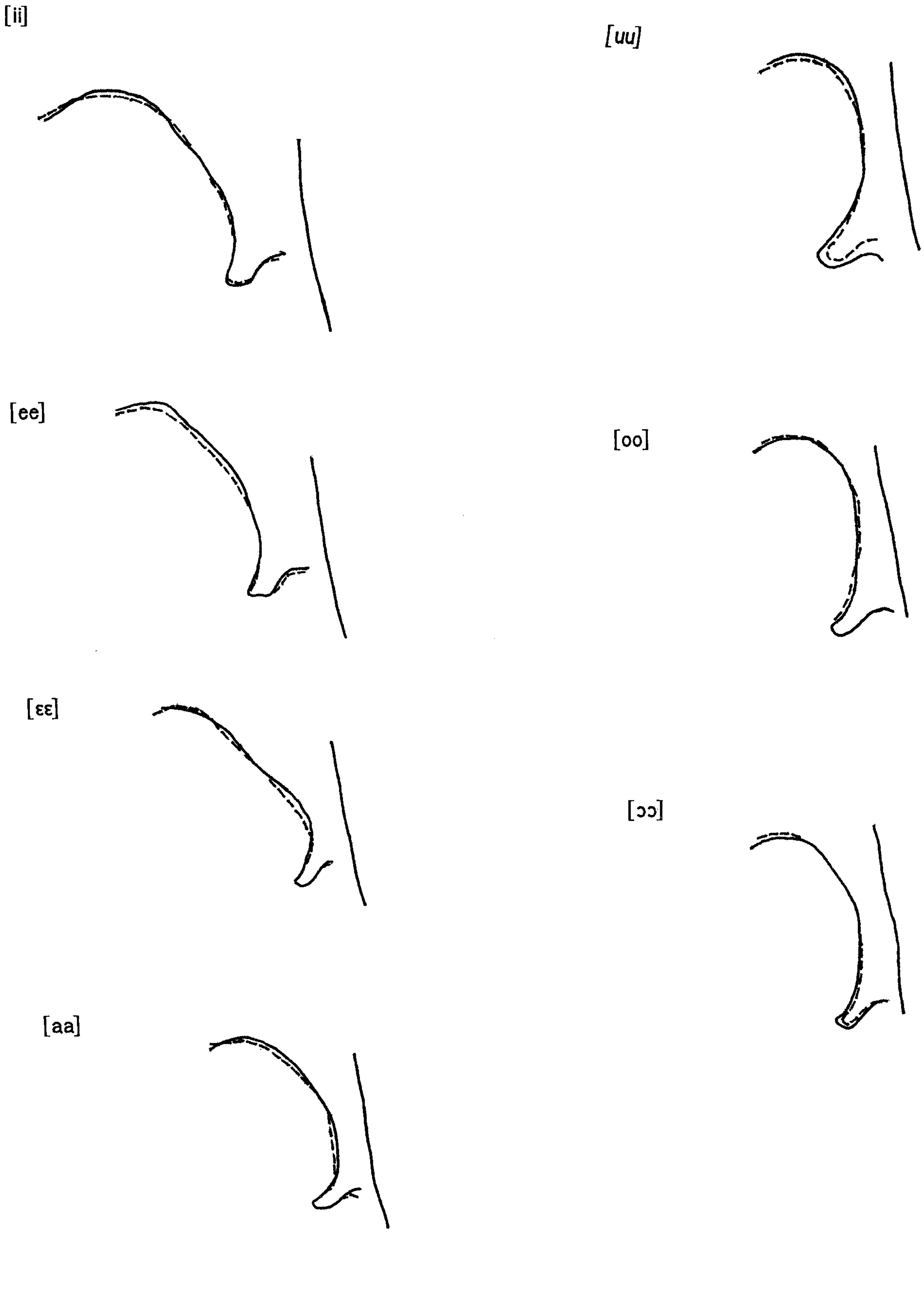


Figure 7: X-ray tracings of tongue positions of clear voiced (solid line) and breathy voiced (dotted line) vowels in Nyah Kur.

5. Conclusion

In conclusion, at the linguistic level, register phenomena should be described in terms of multi-dimensional features or a set of articulatory or acoustic parameters as proposed by Ladefoged (1980). At an abstract phonological level, less specific terms, such as TENSE vs. LAX seem to be more appropriate for an analysis of phonological patterns. For phonologists who care for exactness, a feature like GLOTTAL STRICTURE (Ladefoged 1971) or GLOTTAL APPERTURE (Ladefoged 1980) might be more attractive.

Acknowledgements

I would like to express my sincere thanks to Dr. Praphot Khlongsusuek (M.D.) and Dr. Sillawat Arthacinda (M.D.) for their collaboration. My appreciation goes to the Toyota Foundation and the Research Section, Chulalongkorn University for providing funds which made the research reported in this paper possible. I also would like to thank Peter Ladefoged for his comments. Without his help and encouragement, this paper could never have been completed.

References

Abercrombie, D. (1967). *Elements of General Phonetics.* Chicago: Aldine Publishing Company.

Catford, J.C. (1964). Phonation types: the classification of some laryngeal components of speech production. In *In Honour of Daniel Jones,* David Abercrombie et al, eds., pp. 26-37. London: Longmans.

________ (1977). *Fundamental problems in phonetics.* Edinburgh: Edinburgh University Press.

Egerod, S. (1971). Phonation types in Chinese and Southeast Asian languages. *Acta Linguistics Hafniensia* 8.2: 159-71.

Ferlus, M. (1979). Formation des registres et mutations consonantiques dans les langues Mon-Khmer. In *Mon-Khmer Studies* 8, Philip N. Jenner, ed., pp. 1-76. Honolulu: The University Press of Hawaii.

Fischer-Jørgensen, E. (1977). Phonetic analysis of breathy (murmured), vowels in Gujarati. *Annual Report of the Institute of Phonetics of the University of Copenhagen* 2: 105-56. (Reprinted from 1967.)

Glover, W. (1971). Register in Tibeto-Burman languages of Nepal: a comparison with Mon-Khmer. *Papers in South East Asian Linguistics, No. 2, Pacific Linguistics, Series A - No. 29*: 1-22.

Gregerson, K.L. (1976). Tongue-root and register in Mon-Khmer. In *Austroasiatic Studies 1*, Philip N. Jenner et al, eds., pp. 323-66. Honolulu: The University Press of Hawaii.

______ and Smith, Kenneth D. (1973). The development of Todrah (?) register: *Mon-Khmer Studies 4*, David D. Thomas and Nguyen Dinh Hoa, eds. Saigon. Pp. 143-84.

Henderson, E.J.A. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies* 14: 149-74.

______ (1965). The topograpy of certain phonetic and morphological characteristics of South East Asian languages. *Lingua* 15: 400-34.

______ (1977). The larynx and language: a missing dimension? *Phonetica* 34: 256-63.

Huffman, F.E. (1976). The register problem in fifteen Mon-Khmer languages. *Austroasiatic Studies 1*, Philip N. Jenner et al, eds., pp. 575-89. Honolulu: The University Press of Hawaii.

Jenner, P.N. (1974). The development of the registers in Standard Khmer. *South East Asian Linguistics Studies, Pacific Linguistics Series C-No. 31*, Nguyen Dang Liem, ed., pp. 47-60.

Kirk, P.L. et al. (1984). Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics* 59: 102-13.

Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics.* Chicago: The University of Chicago Press.

______ (1980). What are linguistic sounds made of? *Language* 56.3: 485-502.

Laver, J. (1980). *The Phonetic Description of Voice Quality.* Cambridge: Cambridge University Press.

Matisoff, J.A. (1973). Tonogenesis in Southeast Asia. In *Consonant Types and Tone, Southern California Occasional Papers in Linguistics No.1*, Larry H. Hyman, ed., pp. 71-95. Los Angeles: The Linguistics Program, University of Southern California.

Pittman, R.S. (1978). The explanatory potential of voice-register phonology. *Mon-Khmer Studies* 7, Philip N. Jenner, ed., pp. 201-226. Honolulu: The University of Hawaii Press.

Shorto, H.L. (1967). The register distinctions in Mon-Khmer languages. *Wissenschaftliche Zeitschrift der Karl-Marx-Universität, Leipzig* 16: 245-8.

Smith, K.D. (1968). Laryngealization and delaryngealization in Sedang phonemics. *Linguistics* 38: 52-69.

Sprigg, R.K. (1978). Phonation types: a re-appraisal. *Journal of the International Phonetic Association* 8.1-2: 2-17.

Thongkum, T.L. (1982). Register without tongue-root in Nyah Kur (Chao Bon). Paper presented at the 15th International Conference on Sino-Tibetan Languages and Linguistics, August 1982, Beijing.

______ (1985). An acoustic study of the register in Kui (Suai). Paper presented at the

18th International Conference on Sino-Tibetan Languages and Linguistics, August 1985, Bangkok. (To appear in the proceedings.)

______ (in press). An instrumental study of Chong registers. In *Essays on Mon-Khmer Linguistics in Honour of H.L. Shorto*, Jeremy Davidson, ed.

Thurgood, G. (1980). Consonants, phonation types and pitch height. *Computational Analyses of Asian and African Languages* 13: 207-19.

Weidert, A.K. (1978). The register system of northern Khasi. Paper presented at the 10th International Congress of Anthropological and Ethnological Science, Central Institute of Indian Languages.

Mechanisms for voicing and aspiration: Hindi and other languages compared

R. Prakash Dixit*
Phonetics Laboratory, Department of Linguistics
University of California, Los Angeles, CA 90024-1543

Abstract

Dynamic glottal adjustments during bilabial plosives of Hindi were investigated using photo-electric glottographic techniques. Actual and nonsense words of the CV, VCV and VC type were used, where C was one of the plosives /p ph b bh/ and V was the vowel /i/. The vowel following C in CV words and preceding C in VCV and VC words was stressed. The words were embedded in the carrier sentence /didi ___ boltje/ 'Elder sister ___ (please) say'. The results show that: the voiced unaspirated plosives are generally produced with an adducted glottis and vibration of the vocal folds; the unvoiced unaspirated plosives may be produced either with an adducted or an abducted glottis, but with non-vibrating vocal folds; the voiced aspirated plosives are produced with an adducted glottis during part of the closure interval and with an abducted glottis during part of the closure interval and all of the noise interval, but with vibrating vocal folds throughout. The unvoiced aspirated plosives are produced with an abducted and quiescent glottis. The glottis is narrowly open during the unvoiced unaspirated plosives, moderately open during the voiced aspirated plosives and widely open during the unvoiced aspirated plosives. Glottal abduction for the unvoiced unaspirated plosives begins somewhat before the articulatory closure and ends immediately after articulatory release, for the voiced aspirated plosives it begins appreciably before and ends much after the articulatory release, and for the unvoiced aspirated plosives it begins somewhat before the articulatory closure and ends much after the articulatory release. These results are compared with those reported from other languages. General mechanisms for the control of voicing, devoicing and aspiration associated with plosive production are discussed.

1. Introduction

In recent years, research efforts by a number of investigators have been directed toward the collection of experimental data on speech related dynamic adjustments of the glottis, particularly those that are used in the production of homorganic plosive stops. To contribute to these efforts Dixit and MacNeilage (1974) presented some photo-electric glottograms on glottal fricative and bilabial plosives of Hindi at the 87th Meeting of the Acoustical Society of America, New York. To the author's knowledge, no glottographic data were reported before that time on Hindi (or any other language) which uses mutually independendent but intersecting phonetic dimensions of voicing and aspiration contrastively, yielding four manner categories of homorganic plosives; namely, unvoiced unaspirated, unvoiced aspirated, voiced unaspirated and voiced aspirated. There is fiberoptic data reported by Hirose, Lisker and Abramson (1972) on five different categories of stops, including four plosives and one implosive. However, their subject was a trained phonetician and a native speaker of American English. Otherwise, nearly all of the previously reported data on glottal dynamics, pertaining to the production of plosives, came from languages that possess either a two-way manner contrast like Danish (Frøkjaer-Jensen, Ludvigsen and Rischel, 1971), Dutch (Slis and Damsté, 1967), English (Abramson, Lisker and Cooper, 1965; Lisker, Abramson, Cooper and Schvey, 1969; Lisker, Sawashima, Abramson and Cooper, 1970; Sawashima, 1970; Sawashima, Abramson, Cooper and Lisker, 1971; Fujimura and Sawashima, 1971), Japanese (Sawashima, 1968a, 1968b; Sawashima, Hirose, Kiritani and Fujimura, 1968; Sawashima and Miyazaki, 1973; Sawashima and Niimi, 1974) and Swedish (Lindqvist, 1972), or three-way manner contrast like Korean (Kim, 1967, 1970; Kagaya, 1971, 1974; Hirose, Lee and Ushijima, 1974).

Since Dixit and MacNeilage's (1974) report, Kagaya and Hirose (1975), and Benguerel and Bhatia (1980) have published glottographic data on Hindi. Similar data on Maithili have been published by Yadav (1984). Like Hindi, Maithili is also a four-category language belonging to the Indo-Aryan family of languages. Some further glottographic data have also appeared on two- and three-category languages, such as Cantonese (Iwata, Sawashima and Hirose, 1981), Danish (Fukui and Hirose, 1983; Hutters, 1985), Dutch (Yoshioka, Löfqvist and Collier, 1982), English (Yoshioka, Löfqvist and Hirose, 1981; Lisker and Baer, 1984), French (Benguerel, Hirose, Sawashima and Ushijima, 1978), Fukienese (Iwata, Sawashima, Hirose and Niimi, 1979), Icelandic (Pétursson, 1976; Löfqvist and Yoshioka, 1981), Japanese (Sawashima, Hirose, Ushijima and Niimi, 1975), Korean (Sawashima and Park, 1979; Sawashima, Park, Honda and Hirose, 1980), Mandarin (Iwata and Hirose, 1976), Swedish (Löfqvist and Yoshioka, 1980), and Tibetan (Kjellin, 1977).

The present study is based on a much larger and different set of photo-electric glottographic recordings than the earlier report. Its purpose is three-fold: (1) to provide further qualitative and quantitative data on temporal coordination of glottal and supra-glottal articulatory events in the production of Hindi plosives; (2) to compare these data with the similar data from the languages cited above; (3) and to discuss general mechanisms for the control of voicing, devoicing and aspiration associated with various homorganic plosives.

2. Method

The subject (RPD) for the present study was a male native speaker from western Uttar Pradesh (a Hindi-speaking state) in northern India. He had no history of voice or hearing disorders.

Minimal pairs of monosyllabic and bisyllabic actual and nonsense words constituted the speech sample. They were of the type: CV, VCV and VC, where C represented the bilabial plosives /p/ (unvoiced unaspirated [p]), /ph/ (unvoiced aspirated [pʰ]), /b/ (voiced unaspirated [b]), /bh/ [voiced aspirated [bʱ]) and V was the hight front unrounded vowel /i/. The vowels following C in the CV words and preceding C in the VCV and VC words were stressed. Thus, the initial plosives occurred in the prestressed prevocalic position, the medial plosives in the poststressed intervocalic position and the final plosives in the poststressed postvocalic position. All of the test words were embedded in a carrier sentence /didi___boltje/ 'Elder sister ___ (please) say'.

A microphone (Shure 565) and a photo-electric glottograph (F.J. Electronics, Denmark) were used to record the audio signal and to register light passing through the glottis, respectively. The photo-electric glottograph, described in detail elsewhere (Frøjkjaer-Jensen, 1967), yields quite reliable information about both the dynamic adjustments of the glottis and the presence or absence of glottal pulsing (vocal fold vibration). In brief, the photo-electric glottograph is composed of a 150 watt halogen lamp with blower, reflector, lens, and heat-absorbing filter; a light-directing 40 cm long (tapered) polished acrylic rod; a light-transducing silicon photo-transistor with a specially constructed wide-angle lens mounted in a 2 mm diameter polyethylene tube; and an amplifier based upon an integrated circuit followed by two emitter followers. The photo-electric glottograph has a frequency range of DC to 10kHz and a dynamic range of 50 dB with reference to total darkness.

Cold light directed through the rod penetrates the tissue of the neck just below the cricoid cartilage of the larynx. Variations in the amount of light passing through the glottis, which for the most part reflect the variations in the glottal area (Frøkjaer-Jensen, 1967), are picked up by the photo-transistor contained in the flexible and transparent polyethelene tube. The tube is inserted through the nose into the pharynx and partly swallowed into the esophagus to ensure the relative stability of the light-transducer against the movements of the tongue and epiglottis. For the present experiment, the photo-transistor was positioned appreciably above and behind the tip of the epiglottis.

The signals from the microphone and the photo-electric glottograph were simultaneously recorded on an FM data tape recorder (Honeywell 7600) as the sentences containing test words were spoken by the subject at his normal conversational level of pitch and loudness. They were later reproduced on paper by using an ink-jet oscillograph (Siemens Oscillomink, Model E) for visual inspection and selection of artifact-free tokens for measurements. Figures 1,2, and 3 show such oscillographic reproductions.

Each one of these figures consists of four graphs containing two oscillographic traces. The top trace represents the audio (AUD) signal from the microphone and the bottom trace the signal from the photo-electric glottograph. In the photo-electric glottographic (PEG) signal, the upward and downward deflections of the trace indicated opening and closing gestures of the glottis, respectively. The left half of these figures displays the unvoiced (UNVD) plosives and the right half their voiced (VD) cognates; the top half of these figures shows the unaspirated (UNASP) plosives and the bottom half their aspirated counterparts.

Hand-drawn vertical lines in each one of the graphs of these figures mark certain articulatory-acoustic events. Specifically, (from left to right) the first line marks the approximate point of articulatory contact (AC) or the beginning of closure interval; the second line marks the point of articulatory release (AR) or the end of closure interval; and the third line marks the end of noise interval or the onset of the vowel (OV) following initial and medial plosives or the onset of the voiced segment (a bilabial plosive) occurring initially in the next word belonging to the carrier sentence. It should, however, be mentioned that the second line in /ibi/ (Figure 2, Graph 2) marks not only the point of articulatory release but also the onset of the vowel following the plosive. Further, the second line in the utterance containing /ip/ (Figure 3, Graph 1) indicates the onset of vocal fold vibration for the initial /b/ of the next word, whereas in the utterance containing /ib/ (Figure3, Graph 2) the second line is drawn more or less arbitrarily to separate the final /b/ of the test word from the initial /b/ of the next word belonging to the carrier sentence. Both final /p/ and /b/ of the test words were unexploded; hence, the final /p/ and /b/ of the test words and the initial /b/ of the next word belonging to the carrier sentence had a common closure interval.

Numerals 1,2, and 3 on the PEG traces indicate the initiation of the glottal opening gesture (IGO), the peak of glottal opening (PGO) and the termination of the glottal

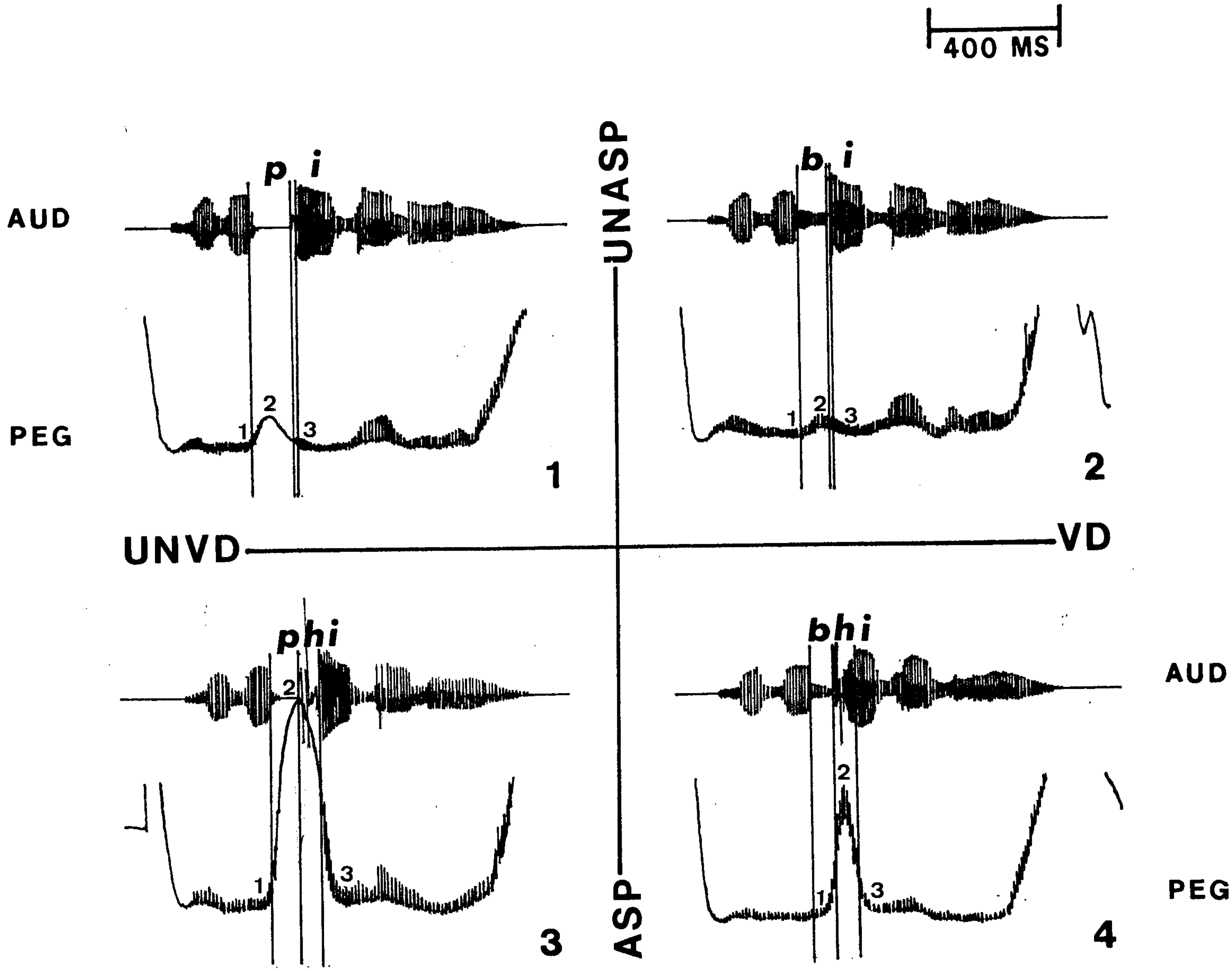


Figure 1. Oscillographic traces for the audio (AUD) signals and the photo-electric glottograms (PEG) for the test words /pi/, /bi/, /phi/, and /bhi/ spoken in the frame sentence /didi ___ bolɭje/.

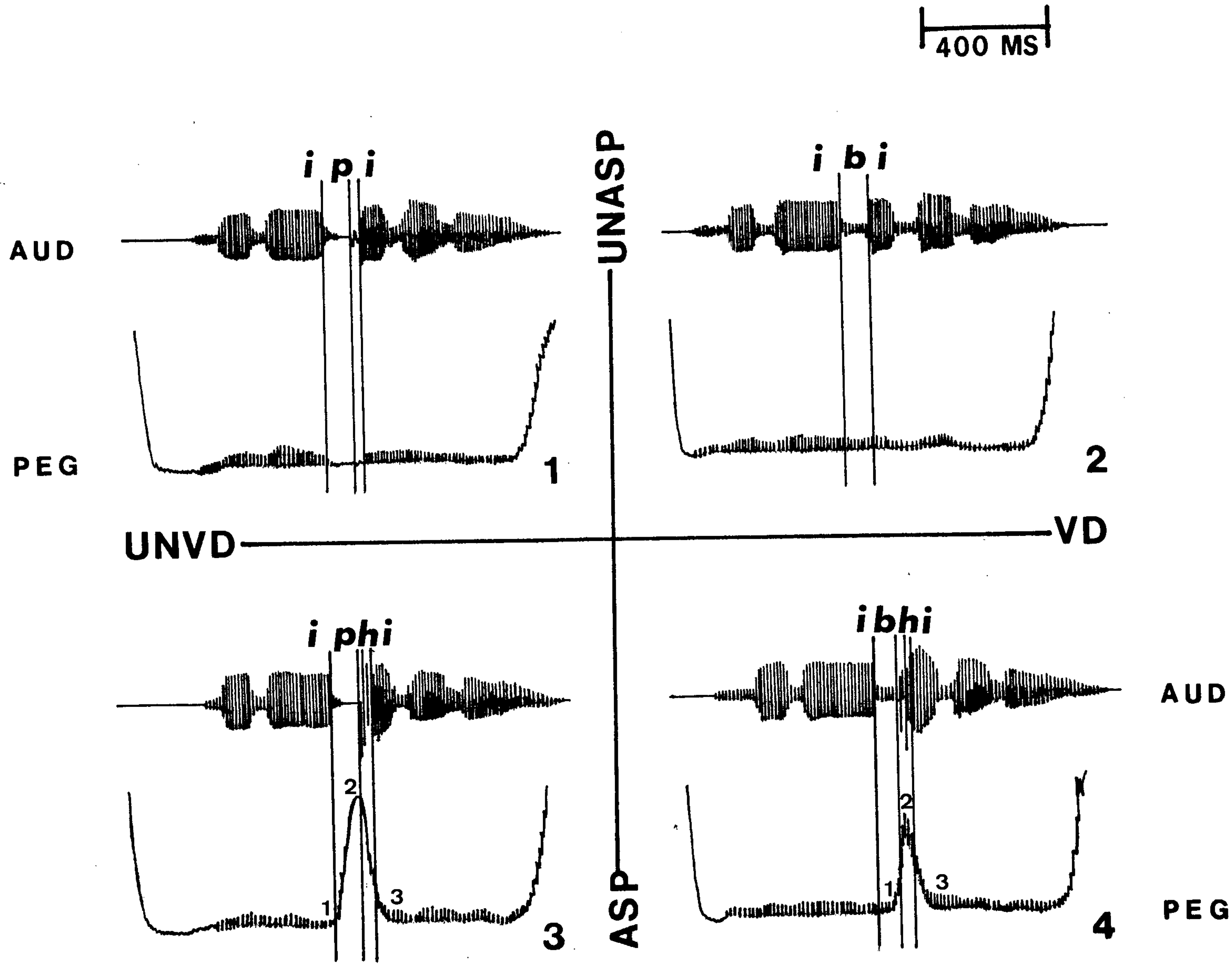


Figure 2. Oscillographic traces for the audio (AUD) signals and the photo-electric glottograms (PEG) for the test words /ipi/, /ibi/, /iphi/, and /ibhi/ spoken in the frame sentence /didi ___ bolʎje/

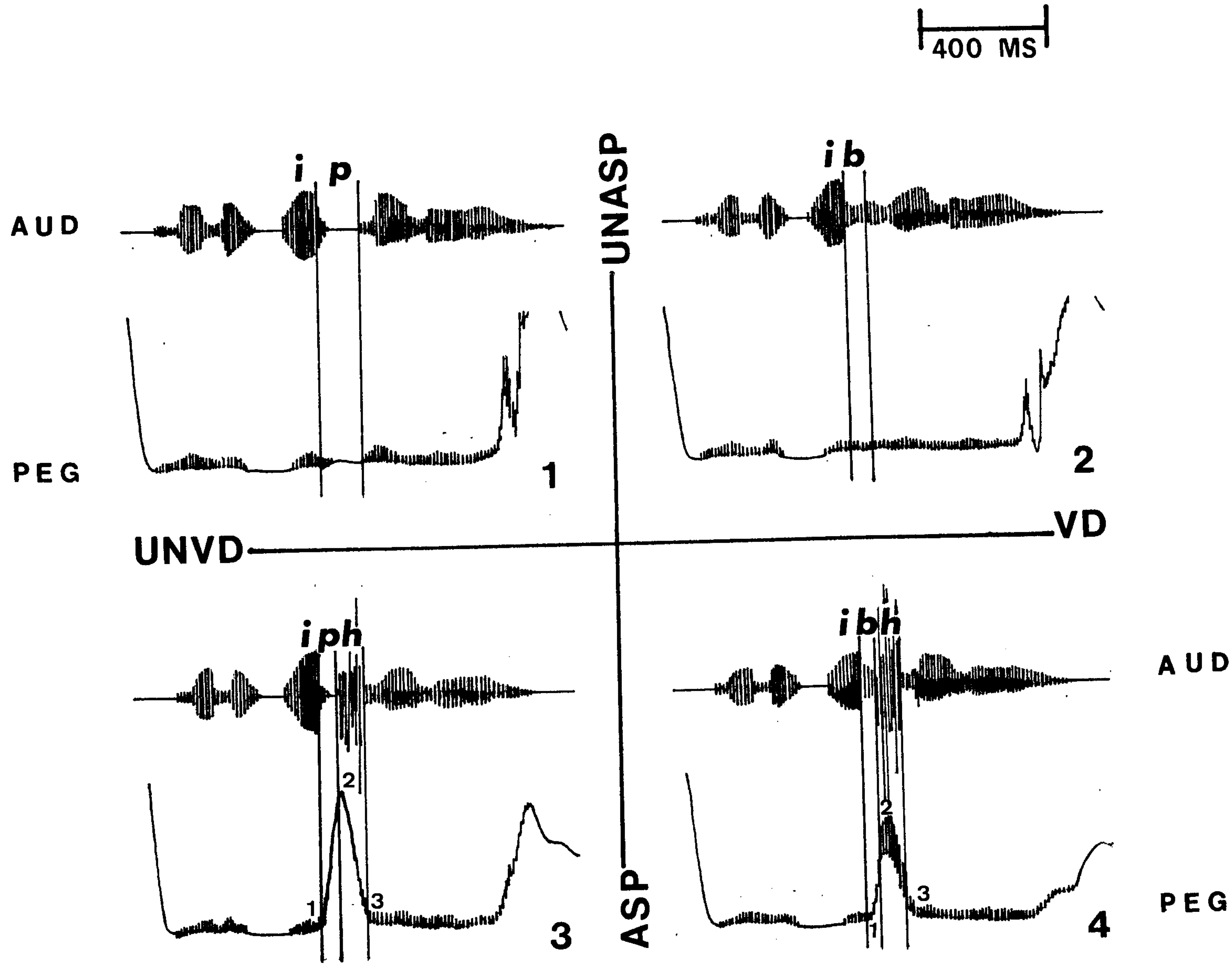


Figure 3. Oscillographic traces for the audio (AUD) signals and the photo-electric glottograms (PEG) for the test words /ip/, /ib/, /iph/, and /ibh/ spoken in the frame sentence /didi ___ boltje/.

closing gesture (TGC), respectively.

Selected tokens of each utterance were used to measure time intervals between certain glottal and articulatory events indicated by numerals on the PEG curves and vertical lines drawn through AUD signals and PEG curves in Figures 1,2, and 3. Thus the following measurements were made:

(1) time interval between initiation of glottal opening and articulatory contact,
(2) time interval between peak of glottal opening and articulatory contact,
(3) time interval between peak of glottal opening and articulatory release,
(4) time interval between termination of glottal closing and articulatory release,
(5) time interval between termination of glottal closing and vowel onset,
(6) duration of closure interval DCI,
(7) duration of noise interval DNI,
(8) height of the peak glottal opening PH.

A general purpose laboratory computer (PDP-12, Digital Equipment Corporation) was utilized for measuring and averaging.

We will refer back to the raw data, presented above, in the following sections.

3. Results

Table I shows quantitative data in the form of averages based on 15 to 20 measures of time intervals between initiation of glottal opening gesture and articulatory contact (column 2), peak of glottal opening and articulatory contact (column 3), peak of glottal opening and articulatory release (column 4), termination of glottal closing gesture and articulatory release (column 5) and termination of glottal closing gesture and vowel onset (column 6) for the different manner categories of initial, medial and final bilabial plosives. All values for the time intervals are in milliseconds.

Column 7 of Table I shows duration values in milliseconds for the closure interval of various bilabial plosives. The duration values for the closure interval of unvoiced plosives in the word-initial position and unexploded plosives in the word-final position may raise questions as to what these values really represent, since the point of articulatory contact does not show up in the acoustic signal for the former, and the point of articulatory release does not occur for the latter and both these points are important references for measuring the closure interval. In this table, values for the closure duration for the initial plosives represent the measures of duration from the offset of the final vowel of the word /didi/ in the carrier sentence to the point of articulatory release of these plosives. On the other hand, in the word-final position where unexploded plosives /p/ and /b/ shared closure duration with the initial plosive /b/ of the word /bolije/ in the carrier sentence, the determination of the closure duration for /b/ was in essence

==

Table I. Mean values (in ms) of time intervals between initiation of glottal opening gesture (IGO) and articulatory contact (AC), peak of glottal opening gesture (PGO) and AC, PGO and articulatory release (AR), termination of glottal closing gesture (TGC) and AR, TGC and onset of vowel or voiced segment following the plosive of interest--immediately or across word boundary (OV). A negative value indicates AC, AR, or OV preceding a related point, such as IGO, PGO, TGC, on the PEG curves. Columns 7 and 8 show the mean durations (in ms) of the closure interval (DCI) and the noise interval (DNI), respectively. Column 9 shows the peak heights (PH) of glottal opening (in mm).

	Intervals Between							
Plosives	IGO	PGO		TGC				
	AC	AC	AR	AR	OV	DCI*	DNI*	PH*
1	2	3	4	5	6	7	8	9
Initial								
p	16	-64	60	-40	-24	124	16	06
b	Generally, no light passing through the glottis registered					88	00	00
ph	18	-92	00	-120	-60	92	60	37
bh	-40	-112	-32	-118	-62	80	56	19
Medial								
p	Generally, no light passing through the glottis registered					120	10	00
b	No light passing through the glottis registered					92	00	00
ph	12	-92	08	-92	-44	100	48	25
bh	-44	-92	-24	-100	-52	68	48	13
Final								
p	Generally no light passing through the glottis registered					120	00	00
b	No light passing through the glottis registered					60	00	00
ph	20	-72	-08	-104	-32	64	72	22
bh	-16	-68	-20	-104	-40	48	64	11

* See text for explanation of what these measures represent.

arbitrary, whereas for the /p/ it was based on that portion of the common closure interval during which the glottal pulses were absent.

Column 8 of Table I shows the duration of the noise interval of different bilabial plosives. In aspirated plosives this is customarily called "the period of aspiration," though this is not very satisfactory, as the noise interval of an aspirated plosive consists not only of aspiration noise but also of transient and frication noises. The values shown represent measures of duration from the point of articulatory release to the onset of regular glottal vibration for the vowel following word-initial and medial plosives and for the voiced segment following word-final plosives and belonging to the next word in the carrier sentence. Judging from the oral air pressure and oral air flow curves for these plosives, presented in Dixit and Brown (1978), the closure durations and noise durations (Table I) seem quite reasonable.

Column (9) of Table I displays peak height (PH) values in millimeters for the degree of glottal opening during various bilabial plosives. Since photo-electric glottograph records cannot at present be calibrated, the PH measures of the degree of glottal opening have only comparative value.

Numerical data presented in Table I will be described in conjunction with Figure 4. Both Table I and Figure 4 summarize most of our glottographic results on bilabial plosives.

In Figure 4, variations in glottal area (GA) as a function of time and heights of peak glottal opening are shown along the ordinate in arbitrary scale, and time is shown along the abscissa in milliseconds (ms). Solid curves represent the unvoiced aspirated plosives; dashed curves represent the voiced aspirated plosives.; and dotted curves represent the unvoiced unaspirated plosives. No curves are shown for the voiced unaspirated plosives and the word medial and final unvoiced unaspirated plosives, since little or no light passing through the glottis was registered during their production. The curves in this figure are aligned in relation to the instant of articulatory release. The top panel of this figure displays the glottal area curves for the initial plosives, the middle panel for the medial plosives and bottom panel for the final plosives. Numerals 1,2, and 3 on the glottal area curves are the same as those on the PEG curves in Figures 1, 2 and 3; that is, they represent initiation of glottal opening, peak of glottal opening and termination of glottal closing gesture, respectively. Short vertical lines on the glottal area curves of the voiceless plosives mark the cessation of

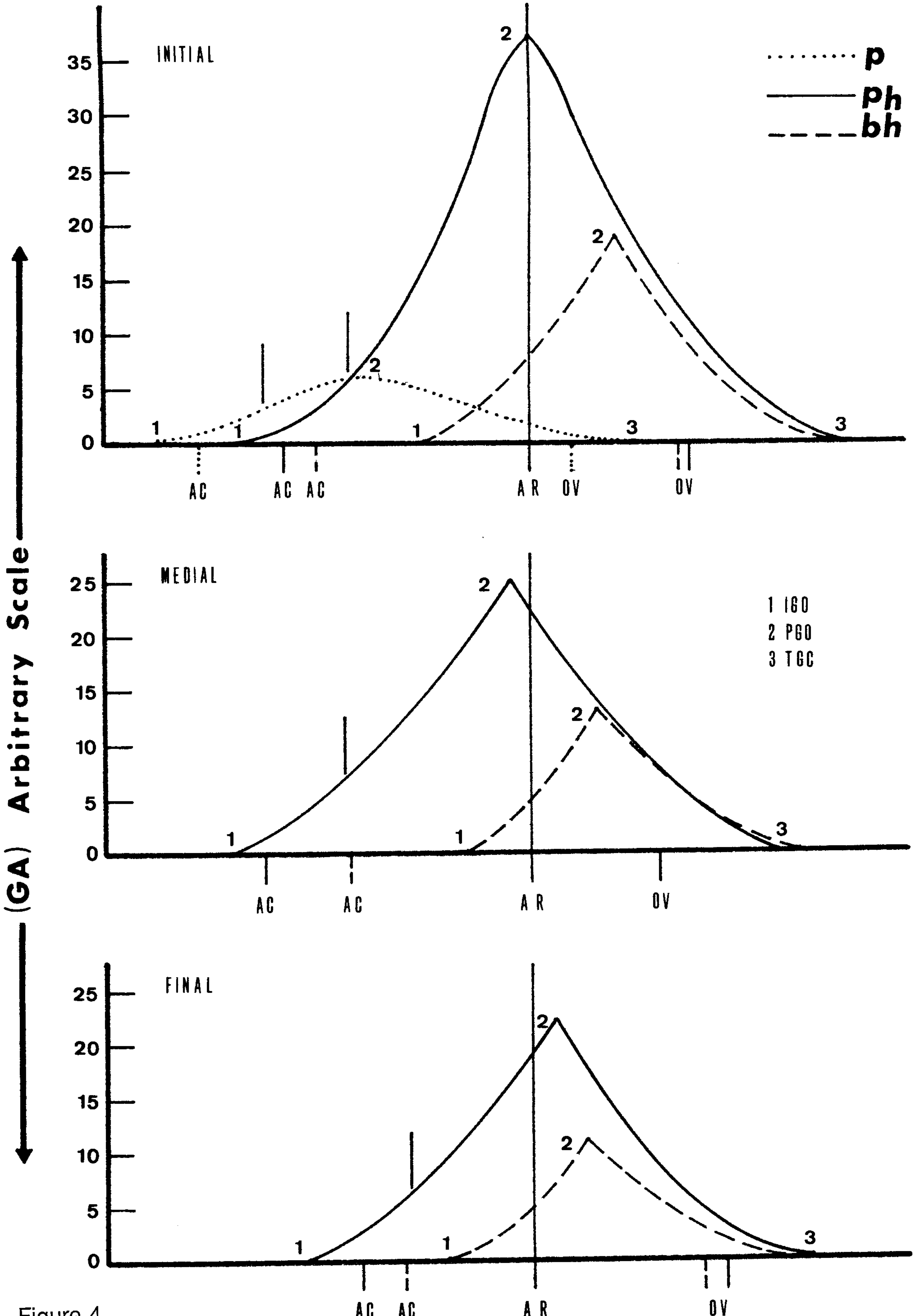


Figure 4.

Superimposed GA (glottal area) curves comparing temporal course of glottal dynamics (opening and closing gestures of the glottis) with respect to supraglottal articulatory adjustments (articulatory contact and release) for different manner categories of plosives. Line-up point: AR (articulatory release).

voicing continuing from the preceding voiced environment into the early part of closure interval. The rest of the closure interval of the unvoiced plosives was indeed voiceless (see Figures 1,2 and 3). The glottal area curves in this figure, like the PEG curves in Figures 1,2 and 3, not only represent the variations in the glottal area as a function of time, but are also the basis for the timing relationship in Table I.

On the basis of the data presented in Table I and Figure 4 (also Figures 1,2 and 3) the following observations can be made:

The voiced unaspirated plosive /b/ in all word positions and the unvoiced unaspirated plosive /p/ in word-medial and final positions were, generally, produced with an approximated glottis. It should, however, be mentioned that a few tokens of word-initial /b/ and word-medial and final /p/ were produced with a slight glottal opening. On the other hand, the word-initial /p/ was produced with an appreciable glottal opening. The opening gesture of the glottis generally began somewhat earlier than the instant of articulatory contact, the peak of glottal opening occurred approximately mid-way between the beginning of the closure and the release (that is, near the middle of the closure interval), and the termination of the glottal closing gesture occurred immediately after the instant of articulatory release.

Irrespective of voicing, all aspirated plosives in all three word positions were, invariably, produced with an open glottis. For the unvoiced aspirated plosive /ph/, the opening gesture of the glottis generally began somewhat before the instant of articulatory contact, the peak of glottal opening occurred at or closely around the instant of articulatory release, and the termination of the closing gesture of the glottis occurred much after the release in the early part of the following vowel or voiced segment. On the other hand, for the voiced aspirated plosive /bh/, the opening gesture of the glottis usually started near the middle of the closure interval, the peak of glottal opening occurred considerably after the instant of articulatory release, generally near the middle of the noise interval, and the termination of the glottal closing gesture occurred much after AR in the early part of the following vowel or voiced segment.

It should be further observed that the timing of the initiation of the glottal opening gesture in relation to articulatory contact for the unvoiced plosives /p/ and /ph/ (irrespective of aspiration) was quite similar, but the timing of the peak of glottal opening and the termination of the closing gesture of the glottis in relation to the points of articulatory contact and articulatory release was very different. The timing of the termination of glottal closing gesture in relation to the point of articulatory release for the aspirated plosives /ph/ and /bh/ (irrespective of voicing) was almost identical, but the timing of their initiation of glottal opening gesture and peak of glottal opening with respect to articulatory contact and articulatory release differed considerably. Between the voiced aspirated plosive /bh/ and the unvoiced unaspirated plosive /p/, on the other hand, there were absolutely no similarities with respect to the timing of glottal gestures vis-a-vis the timing of supraglottal articulatory events. The timing of the initiation of glottal opening gesture, peak of glottal opening and termination of glottal closing gesture in relation to the points of articulatory contact and articulatory release was always very different in these plosives. Thus, the temporal courses of the glottal dynamics for different manner categories of plosives were quite different.

Besides the timing difference between the opening-closing gestures of the glottis and the supraglottal articulatory adjustments, there were also remarkable differences in the degree (size or extent) of glottal opening (as reflected in the PGE and glottal area curves and the peak height measures) among different categories of plosives (Figures 1,2,3 and 4 and Table I). The degree of glottal opening systematically decreased in descending order from the unvoiced aspirated plosives to the voiced aspirated plosives to the unvoiced unaspirated plosives (ph > bh > p). The glottis was widely open during the unvoiced aspirated plosives, moderately open during the voiced aspirated plosives, narrowly open during the unvoiced unaspirated plosives, and approximated during the voiced unaspirated plosives. Glottal width during the voiced aspirated plosives was approximately half of that during the unvoiced aspirated plosives while during the unvoiced unaspirated plosives it was approximately a third of that during the voiced aspirated plosives. (Since the word-medial and final unvoiced unaspirated plosives were produced with the approximated glottis the above statement is applicable only to the word-initial unvoiced unaspirated plosives.) Both the timing of glottal opening-closing gestures and the degree of glottal opening appear to be conditioned by the acoustic characteristics of various plosives, the aerodynamic requirements for which differ from one manner category of plosives to another.

Further, the degree of glottal opening for various plosives in the word-initial (prestressed, prevocalic) position was found to be greater than for similar plosives in the word-medial (poststressed, intervocalic) or the word-final (poststressed, postvocalic) positions. In the latter two positions, the word-medial plosives showed somewhat larger glottal opening than the comparable word-final plosives (Figure 4 and Table I).

Also, the degree of glottal opening was found not to relate systematically to either the duration of closure or the duration of noise. For example, the closure for the unvoiced unaspirated plosives was invariably longer than for the voiced aspirated and the unvoiced aspirated plosives (Table I), but the degree of glottal opening for the former category of plosives was much smaller than for the latter two categories of plosives (Figure 4). Moreover, the duration of noise (aspiration) was about the same for the voiced aspirated and the unvoiced aspirated plosives (Table I), yet the extent of glottal opening was substantially greater for the latter than for the former.

Finally, no apparent direct correlation between the degree of glottal opening at the time of articulatory release and the degree (duration) of aspiration was observed. For instance, the unvoiced aspirated plosives in the word final position had a longer period of aspiration than the same plosives in the word-medial or word-initial positions (Table I); however, the degree of glottal opening at the time of oral release for these plosives in the word-final position was relatively smaller than in the other two word positions (Figure 4). Furthermore, the degree of glottal opening at the time of oral release for the unvoiced aspirated plosives was much greater than for the voiced aspirated plosives (Figure 4), but the degree or duration of aspiration for the two types was about the same (Table I).

Figure 5 shows ranges of variation for the averaged values of initiation of glottal opening gesture, peak of glottal opening and termination of glottal closing gesture in relation to the point of articulatory release. Filled circles stand for the averaged values (as shown in Table I) and solid horizontal lines represent the ranges of variation. In each row from left to right: the first line depicts the ranges of variation for initiation of glottal opening gesture, the second line for peak of glottal opening and the third line for termination of glottal closing gesture. The ranges are for word-initial /p/, /ph/ and /bh/, and word-medial and word-final /ph/ and /bh/. Short vertical lines to the left of the articulatory release mark the points of articulatory contact, while those to the right of articulatory release indicate the onset of the following vowel. It is of interest to note that the ranges of initiation of glottal opening gesture, peak of glottal opening and termination of glottal closing gesture for the unvoiced unaspirated plosives do not overlap with the similar ranges for the voiced aspirated or the unvoiced aspirated plosives. The ranges of initiation of glottal opening gesture and peak of glottal opening for the voiced aspirated and the unvoiced aspirated plosives are also mutually exclusive, except in the final position where

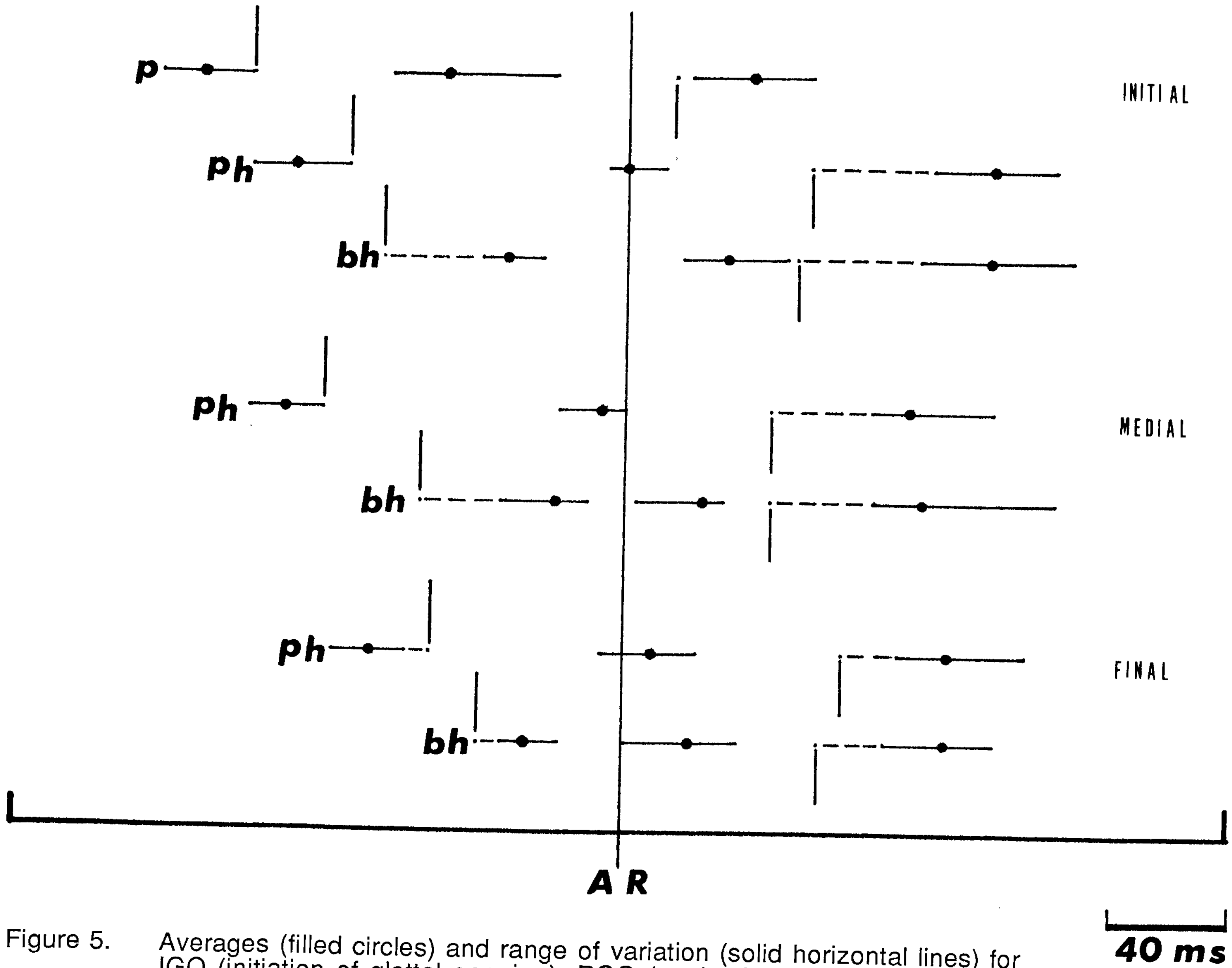


Figure 5. Averages (filled circles) and range of variation (solid horizontal lines) for IGO (initiation of glottal opening), PGO (peak of glottal opening), TGC (termination of glottal closing) gestures in relation to AR (articulatory release). Left to right: first line = IGO, second line = PGO, third line = TGC.

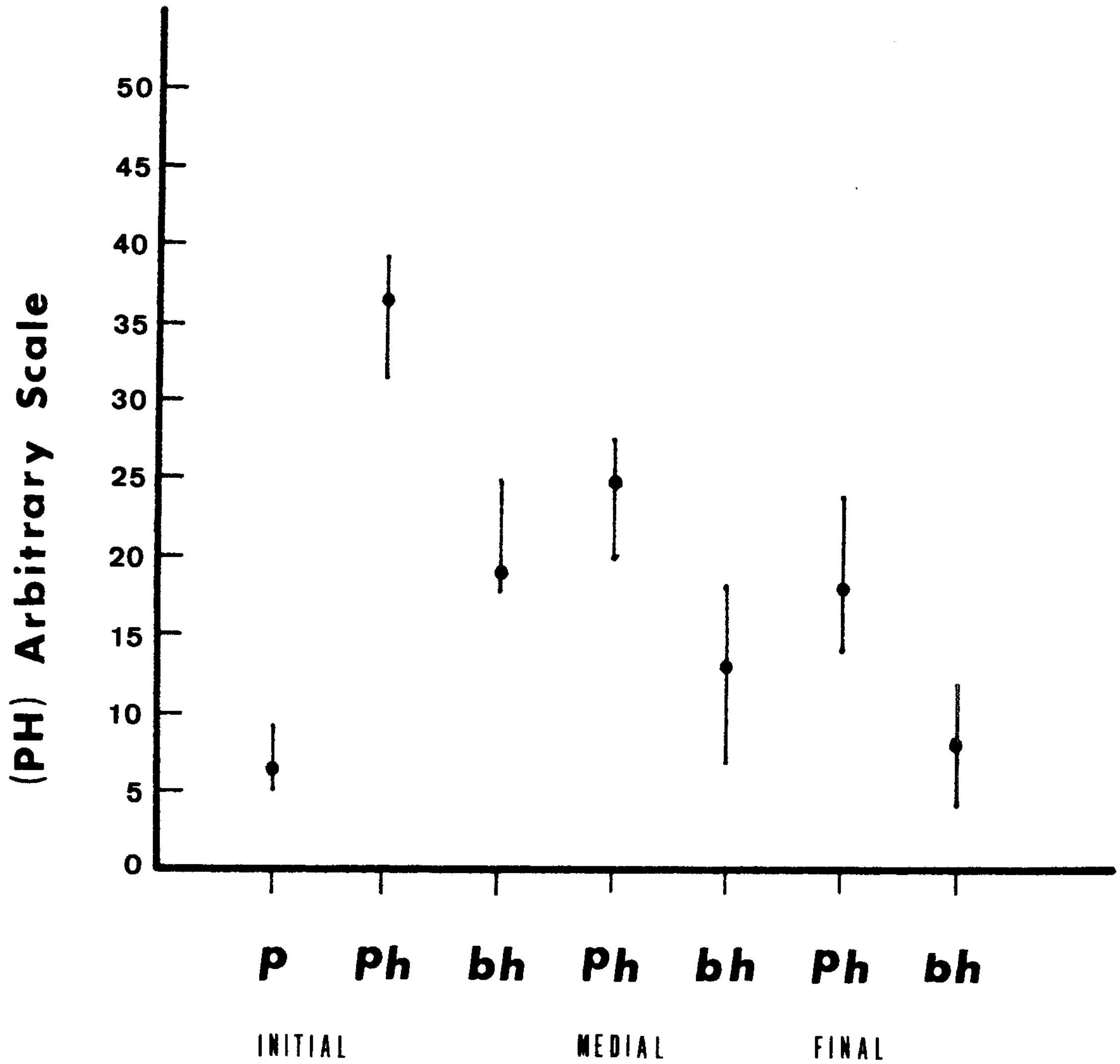


Figure 6. Averages (filled circles) and ranges of variation (solid vertical lines) for PH (peak heights) of glottal opening.

considerable overlap in the peak of glottal opening ranges of these plosives occurs. A large number of peaks of glottal opening for the word-final unvoiced aspirated plosives, like those for the voiced aspirated plosives, occurred appreciably after the release, hence the overlap. On the other hand, the termination of glottal closing gesture ranges of the voiced aspirated and the unvoiced aspirated plosives are more or less common since the closing gesture of the glottis for these plosives terminates at about the same time after the release. The overlap or lack of it in the linitiation of glottal opening gesture, peak of glottal opening and termination of glottal closing gesture ranges of the unvoiced unaspirated, the unvoiced aspirated and the voiced aspirated plosives on the whole conforms quite well with the temporal patterns of glottlal area curves in Figure 4 as expected.

Figure 6 illustrates the ranges of variation for the averaged values of peak heights (PH) of glottal opening measured from the base-line during word-initial /p/, /ph/ and /bh/, and word-medial and word-final /ph/ and /bh/. Filled circles represent the averaged values (as shown in Table I) and solid vertical lines represent the ranges of variation of the peak height. Although the measures of peak height given along the ordinate are in millimeters, the scale is in fact arbitrary. It simply compares the relative extent of glottal opening for different manner categories of plosives; it does not indicate actual size of glottal opening for these plosives. Clearly, the peak height ranges for different manner categories of plosives in a particular word position are non-overlapping.

4. Cross-language Comparison of Dynamic Glottal Adjustments

This section compares glottographic data from Hindi plosives to those from other languages. Since glottal behavior for various sounds is largely determined by their phonetic characteristics, it is imperative to have adequate phonetic information on the sounds involved in a cross-language comparison of glottal dynamics. Classificatory labeling (used for lexical or underlying representation) of various sounds may often be misleading with respect to their phonetic realization. Hence, the following digression.

In two category languages, Danish, Dutch, English, French, Icelandic, Japanese and Swedish are said to utilize voiced versus unvoiced contrast to keep their plosives apart. The voiced and unvoiced plosives in Dutch (Cohen, Ebeling, Fokemma and von Holk, 1961; Lisker and Abramson, 1964; Slis and Damsté, 1967), French (Delattre, 1966; Armstrong, 1959; Benguerel et al., 1978), and Japanese (Bloch, 1950; Han, 1961) are phonetically voiced unaspirated and unvoiced unaspirated, respectively.

However, in Danish (Fischer-Jørgensen, 1954, 1980; Frøkjaer-Jensen et al., 1971; Hutters, 1985) both voiced and voiceless plosives are phonetically voiceless. Word initially (in the stressed strong position) these plosives are distinguished by the presence of aspiration in the unvoiced set and absence of it in the so-called voiced set. Thus, word-initially Danish has two phonetic categories of plosives: the unvoiced unaspirated and the unvoiced aspirated. Word-finally there is free variation in unvoiced unaspirated and unvoiced aspirated plosives. In other positions only unvoiced unaspirated plosives occur; there the distinction is said to be of stop versus fricative. The voiced and unvoiced plosives of Icelandic are also essentially voiceless (Malone, 1923; Pétursson, 1976; Löfqvist and Yoshioka, 1981). In word-initial position the so-called voiced plosives of Icelandic are phonetically unvoiced unaspirated, whereas the voiceless plosives are unvoiced aspirated. In addition, Icelandic has unvoiced preaspirated plosives which occur in intervocalic position. In English (Trager and Smith, 1951; Bronstein, 1960; Lisker and Abramson, 1964; Ladefoged, 1975) and Swedish (Lindqvist, 1972; Fant, 1973) the allophonic distribution of voiced and unvoiced plosives appears quite similar. The so-called voiced plosives of English and Swedish are truely voiced in word-medial intervocalic position; in other positions they may or may not be voiced. In word-initial and final positions they are often produced without voicing and thus can be called phonetically unvoiced unaspirated. The unvoiced plosives of English and Swedish, on the other hand, are aspirated in word-initial and prestressed positions, whereas in unstressed position and after /s/ they are unaspirated. In word-final position they may be aspirated or unaspirated.

Cantonese (Chao, 1947; Lisker and Abramson, 1964; Iwata et al., 1981) and Mandarin (Chao, 1949; Iwata and Hirose, 1976) are said to distinguish their plosives by aspirated versus unaspirated contrast. Both aspirated and unaspirated plosives of Cantonese and Mandarin are voiceless. The aspirated versus unaspirated distinction, however, is found only in syllable initial position in these languages. In syllable final position, Cantonese has only unvoiced unaspirated plosives which are produced without plosion and are generally glottalized. On the other hand, Mandarin has no plosives in the word-final position (Kuraishi, 1963; Iwata et al., 1981).

In three-category languages, Fukienese (Chao, 1934; Iwata et al., 1979) and Tibetan (Goldstein and Nornang, 1970; Kjellin, 1976, 1977) are said to possess the voiced unaspirated, unvoiced unaspirated and unvoiced aspirated plosives. This three-way contrast, however, is lost word-medially and finally in Tibetan (Kjellin, 1976) and syllable-finally in Fukienese (Iwata et al., 1979). In Tibetan, word-medially and finally only voiced unaspirated "or at least unaspirated" plosives occur, whereas in

Fukienese the only plosives that occur syllable-finally are the unvoiced unaspirated plosives which are produced without plosion and may even be glottalized. The three manner categories of Korean, on the other hand, are all voiceless differing from one another primarily in the degree of aspiration and tensity (Lisker and Abramson, 1964; Kim, 1965, 1970). They have often been called tense unaspirated, lax slightly aspirated, and tense heavily aspirated. The lax slightly aspirated plosives of Korean have voiced unaspirated variants in intervocalic position. In final position only unvoiced unaspirated (unexploded) plosives occur.

In four-category languages, plosives are separated on the basis of intersecting contrasts of voicing and aspiration. Thus, Hindi (Dixit, 1963, 1975; Lisker and Abramson, 1964; Kelkar, 1968), and Maithili (Yadav, 1979, 1984) possess phonologically as well as phonetically voiced unaspirated, unvoiced unaspirated, voiced aspirated and unvoiced aspirated plosives. They occur in all word positions.

It should be clear from the above description that classificatory (lexical) labels, such as "voiced", "voiceless", etc., may not always signify the same phonetic category in all these languages. The qualitative and quantitative values of what these labels stand for may vary within the same language as well as across different languages. Variability of phonetic details appears to be a rule in experimental studies. Although, we will be using classificatory labels as headings of sub-sections in the ensuing discussion of glottal dynamics, we will refer back to the preceding phonetic information about different manner categories of plosives when their classificatory labels and phonetic realizations are at variance.

4.1 Voiced unaspirated plosives

Qualitative and quantitative data on dynamic glottal adjustments presented in this study showed that almost all of the tokens of voiced unaspirated plosives of Hindi, irrespective of their phonetic environment and position in a word, were produced with an approximated glottis. The vocal folds remained in voicing position and continued vibrating throughout the closure interval of these plosives. Similar results on the voiced unaspirated plosives of Hindi have been reported by Kagaya and Hirose (1975) and Benguerel and Bhatia (1980). These results are further confirmed by the glottographic data on the voiced unaspirated plosives of Dutch (Slis and Damsté, 1967), French (Benguerel et al., 1978), Fukienese (Iwata et al., 1979), Japanese (Sawashima, 1968a, 1968b; Sawashima et al., 1975), Maithili (Yadav, 1984), and Tibetan (Kjellin, 1977). Like Hindi, in these languages also the voiced unaspirated plosives were, in general, produced with an approximated and vibrating glottis. A similar statement as the above can be made about the voiced unaspirated variants of the lax slightly aspirated plosives of Korean, which occur in intervocalic position (Kagaya, 1971, 1974). Even in English (Abramson et al., 1965; Lisker et al., 1969, 1970; Sawashima, 1970; Sawashima et al., 1970; Cooper et al., 1971) and Swedish (Lindqvist, 1972) where the so-called voiced plosives may be realized phonetically as voiced unaspirated or unvoiced unaspirated, depending on their phonetic environment and position in a word (Section 4), most tokens of these plosives were produced with the vocal folds in voicing position. Only a small number of these plosives in English and Swedish showed some glottal opening during their production. A few tokens of the voiced unaspirated plosives in Dutch, French, Korean and Hindi were also produced with a slight opening of the glottis. Such occasional glottal openings during these plosives have been ascribed to active adjustment of the arytenoid cartilages in English (Lisker et al., 1970; Sawashima, 1970) and to a rise in supraglottal air pressure in Swedish (Lindqvist, 1972).

In Hindi, such occasional glottal openings during voiced unaspirated plosives, which were observed only in the word-initial prestressed position, were probably the consequence of rise in subglottal air pressure. Dixit and Shipp (1985) have reported an increase in subglottal air pressure for these and other plosives in the stressed position. Although, the level of EMG activity in the adductory muscles (the interarytenoid = IA, the lateral cricoarytenoid = LCA, and the thyroarytenoid = TA or vocalis = VOC) of the larynx was found to be somewhat lower for the voiced unaspirated plosives (irrespective of their position in a word) than for the surrounding vowels (Dixit, 1975), it cannot be related to the slight glottal opening which occurred during a few tokens of these plosives in word-initial position, since all other tokens of the voiced unaspirated plosives were produced with a closed glottis. EMG data from the abductor muscle (the posterior cricoarytenoid = PCA) of the larynx on these plosives are unavailable. (See also Hirose and Gay, 1972; Benguerel et al., 1978; and Collier et al., 1979.)

The foregoing comparison of the glottographic results from different languages suggests that irrespective of whether the voiced plosives were realized as voiced unaspirated or unvoiced unaspirated, the glottal gesture in these languages was in most cases for the voiced (phonetically voiced unaspirated) plosives, that is, the vocal folds were in voicing (approximated) position during their production.

Contrarily, in Danish (Frøkjaer-Jensen et al., 1971; Fukui and Hirose, 1983; Hutters, 1985) and Icelandic (Pétursson, 1976; Löfqvist and Yoshioka, 1981) the so-called voiced plosives were not only produced with the interruption of vocal fold vibration but also with an open glottis. Obviously, they were realized phonetically as unvoiced unaspirated plosives in the two languages (Section 4). In fiberoptic films on Icelandic obstruents Löfqvist and Yoshioka (1981) observed a small "spindle shaped opening in the membranous portion of the glottis" during the so-called voiced plosives. Similar observations were made by Hutters (1985) on the so-called voiced plosives of Danish. However, her fiberoptic still (page 5) of such a plosive reveals appreciable separation of the vocal processes. The EMG data presented by Hutters point in the same direction, since the PCA shows a peak of activation and the IA and VOC display a trough of suppression associated with the production of the so-called voiced plosives of Danish. Electromyograms revealing similar reciprocal patterns of activation and suppression between laryngeal abductor and adductors, respectively, were earlier reported by Fischer-Jørgensen and Hirose (1974) for the so-called voiced plosives of Danish. Clearly, the opening-closing gestures of the glottis for these plosives resulted from active adjustment of the arytenoid cartilages. Thus, it appears that the intended glottal gesture in Danish, unlike other languages discussed above, was for the unvoiced unaspirated plosives rather than for the so-called voiced plosives. However, the relative size of glottal opening for these plosives was considerably smaller as compared to that for the unvoiced aspirated plosives in Danish, than it was in Hindi and most other languages. In fact, Fukui and Hirose (1983) have stated that the "relative size of the glottal opening for aspirates and inaspirates in Danish resembles Hindi." Although the peaks of glottal opening in relation to the release occurred earlier for voiceless inaspirates in Danish than those in other languages, the timing of initiation of glottal opening and termination of glottal closure in relation to the points of articulatory contact and release was quite similar in Danish and the other languages discussed above. Thus, from the production standpoint the so-called voiced plosives of Danish are indeed unvoiced unaspirated.

Similar statements as were made above with respect to the so-called voiced plosives of Danish can perhaps be made about the so-called voiced plosives of Icelandic, since in both these languages they are realized phonetically as unvoiced unaspirated. Löfqvist and Yoshioka's (1981) glottal transillumination data shows very small glottal opening during voiceless inaspirates as compared to that during voiceless aspirates. On the other hand, in photo-glottographic data reported by Pétursson (1976) the glottal opening for the voiceless inaspirates is considerably more than half of that for the voiceless aspirates. The difference in the two sets of glottographic data on the so-called voiced (actually unvoiced unaspirated) plosives of Icelandic may simply indicate inter-subject variability. However, in the absence of EMG data from the abductor and adductor muscles of the larynx it would be inadvisable to say that the intended glottal gesture observed in the glottographic data for the so-called voiced plosives of Icelandic was in fact for the unvoiced unaspirated plosives and that it was actively maneuvered; although this may well be the case.

4.2 Unvoiced unaspirated plosives

In the present study all word-initial prestressed unvoiced unaspirated plosives of Hindi were produced with an open glottis and interruption of vocal fold vibration after a few initial pitch periods during their closure interval. The opening gesture of the glottis began at or somewhat before closure, the peak of glottal opening occurred at or near the middle of the closure interval, and the closing gesture of the glottis, which began at the peak, teminated immediately after release. The size of glottal opening during these plosives relative to that during unvoiced aspirated plosives was generally quite small. On the other hand, in the poststressed word-medial intervocalic and word-final postvocalic positions the unvoiced unaspirated plosives of Hindi in the present study were mostly produced with an approximated and quiescent glottis.

Comparison of the glottographic results of this study on unvoiced unaspirated plosives of Hindi with those from other Hindi studies and the studies from other languages seems to be complex, because of considerable variability during their production in the same language and across languages. They have to be compared separately for the three word positions.

The results of the present study on word-initial unvoiced unaspirated plosives are in agreement with those reported by Kagaya and Hirose (1975), and Benguerel and Bhatia (1980) in the sense that they all showed an open glottis during these plosives. Like Hindi, the word-initial unvoiced unaspirated plosives in Dutch (Yoshioka et al., 1982), French (Benguerel et al., 1978), Japanese (Sawashima, 1968a, 1968b; Sawashima and Miyazaki, 1973; Sawashima and Niimi, 1974; Sawashima et al., 1975; Niimi and Sawashima, 1974), Korean (Kim, 1967, 1970; Kagaya, 1971, 1974; Hirose et al., 1974), and Maithili (Yadav, 1984) were also produced with an open glottis. However, the patterns of glottal dynamics during these plosives differed depending on whether they occurred in absolute initial position in isolated words and utterances or in initial position in those words that were embedded in a carrier sentence where they were preceded by a phonetically voiced segment. When the unvoiced unaspirated plosives occurred in non-absolute initial position as they did in the present study and the studies reported by Benguerel et al.

(1978), Kagaya and Hirose (1975), Kim (1967, 1970), Sawashima (1968a, 1968b), Sawashima and Miyazaki (1973), Sawashima and Niimi (1974), Niimi and Sawashima (1974), and Yoshioka et al. (1982) they were produced with the opening-closing gestures of the glottis, but when the unvoiced unaspirated plosives occurred in absolute initial position as in Benguerel and Bhatia (1980), Benguerel et al. (1978), Hirose et al. (1974), Kagaya (1971, 1974), Sawashima et al. (1975), and Yadav (1984) the gesture of the glottis for them was not always the same. Although in general it showed a simple closing gesture from a respiratory open position, sometimes the glottis opened further from a respiratory open position and then closed and sometimes it partly closed, then opened and then closed again (Dixit, Baer and Honda, 1984; Kagaya, 1974; Yadav, 1984). Furthermore, there was considerable variability in the degree of glottal opening and the positioning of glottal peak during the initial unvoiced unaspirated plosives across languages; however, the closing gesture of the glottis for these plosives terminated at or closely around the articulatory release in all these languages. It should, nevertheless, be noted that a number of these plosives (particularly velars) in Japanese were somewhat aspirated, hence the termination of the glottal closing gesture in such cases occurred somewhat later with respect to the articulatory release.

The unvoiced unaspirated plosives of Korean, which are usually described as "tense unaspirated" (section 4), differ from the unvoiced unaspirated plosives of Hindi and other languages in that they show tight adduction of the vocal processes and folds appreciably prior to the articulatory release (Kagaya, 1974) and an EMG peak in the activity of VOC and LCA muscles (Hirose et al., 1974) during the period immediately preceding the oral release. Thus, the tense unaspirated plosives of Korean (which are unvoiced unaspirated), unlike the unvoiced unaspirated plosives of Hindi and other languages discussed above, are in fact (pre-)glottalized. This should dispel Kim's belief that the tense unaspirated plosives of Korean are not (pre-)glottalized (see Kim, 1970: footnote 7, p. 109) and should strengthen the opposite belief that they are (Martin, 1951; Ogura, 1953; Abramson and Lisker, 1972; Ladefoged, 1973).

Unlike Hindi and the languages cited above, the syllable- or word-initial unvoiced unaspirated plosives (within a carrier sentence) in Cantonese (Iwata et al., 1981), Fukienese (Iwata et al., 1979), and in Mandarin (Iwata and Hirose, 1976) were generally produced with the vocal processes closed and a spindle-shaped slight opening in the membranous part of the glottis. Similar results on these plosives from

Tibetan were reported by Kjellin (1977). According to Kjellin in Tibetan, the epiglottis covered the glottis during almost all cases of the unvoiced unaspirated plosives, but in one case where the glottis could be observed, it was found to be closed.

As indicated before, almost all word-medial intervocalic unvoiced unaspirated plosives of Hindi in the present study were produced with the closed and quiescent glottis. These results are at variance with those reported by Kagaya and Hirose (1975), and Benguerel and Bhatia (1980) on the same Hindi plosives occurring in the same word position. Their subjects invariably produced these plosives with a moderately open glottis. Our glottographic data on word-medial plosives of Hindi are also at variance with those from Dutch (Slis and Damsté, 1967), English (Lisker et al., 1969, 1970; Sawashima, 1970), French (Benguerel et al., 1978), Japanese (Sawashima, 1968a, 1968b; Sawashima et al., 1968 Sawashima and Miyazaki, 1973; Sawahsima and Niimi 1974; Niimi and Sawashima, 1974), Maithili (Yadav, 1984), and Swedish (Lindqvist, 1972). In all these studies the intervocalic unvoiced unaspirated plosives were produced with the opening-closing gesture of the glottis. The glottis opened at or near the articulatory closure and closed at or near the articulatory release. However, the position and size of the peak glottal opening differed from one language to another. It should, nevertheless, be observed that some tokens of word-medial and intervocalic unvoiced unaspirated plosives in Dutch (Slis and Damsté, 1967), English (Lisker et al., 1969, 1970; Sawashima, 1970), and Japanese (Sawashima, 1968a, 1968b; Sawashima and Miyazaki, 1973) exhibited little or no glottal opening, while a few of these plosives in English "showed no interruption of glottal vibration" (Lisker et al., 1970; Sawashima, 1970).

Most of the word-final unvoiced unaspirated plosives of Hindi in this study were also produced with an adducted and non-vibrating glottis in the present study. However, most tokens of these plosives in English (Lisker et al., 1969, 1970; Sawashima, 1970; Sawashima et al., 1970) and Swedish (Lindqvist, 1972), and all tokens of these plosives in Korean (Sawashima and Park, 1979; Sawashima et al., 1980), Maithili (Yadav, 1984) and another study of Hindi (Benguerel and Bhatia, 1980) were produced with varying degrees of glottal opening. In English and Swedish, the word-final unvoiced unaspirated plosives occurred in a vocalic environment within regular sentences, whereas in Hindi (Benguerel and Bhatia) and Maithili they occurred in absolute final position in isolated words. On the other hand, in Korean they occurred in absolute final position at the end of the carrier sentence. The temporal course of glottal dynamics during unvoiced unaspirated word-final plosives in English and Swedish was similar to the one observed during the same plosives when they occurred between vowels, that is, the glottis opened near articulatory contact and closed near articulatory release. Whereas, in Hindi (Benguerel and Bhatia), Maithili and Korean, for these plosives the glottis did not close after opening, it rather continued dilating toward a respiratory position.

The glottographic data from Cantonese (Iwata et al., 1981) and Fukienese (Iwata et al., 1979), on the other hand, appear to provide some support for the results of this study on the word-final unvoiced unaspirated plosives of Hindi. In Cantonese and Fukienese also, these plosives in word-final position were produced with a closed glottis; they were, however, glottalized in this position in Cantonese and Fukienese.

Recall that the word-final unvoiced unaspirated plosives in the present study were unexploded and that they shared a common closure interval with a voiced unaspirated plosive which occurred initially in the following word belonging to the carrier sentence. In such phonetic contexts as above, it is not unusual to find a closed glottis during the unvoiced unaspirated plosives. Thus, Fujimura and Sawashima (1971) observed "no clear glottal opening ... for any portion of the t-closure" in "fat-d's". Further, a number of unaspirated variants of voiceless plosives of English in the word-final position but before a word beginning with a vowel were produced with a closed glottis (Sawashima et al., 1970; Lisker and Baer, 1984). In the last study, the glottis was found to be closed during the unvoiced unaspirated plosives although PCA showed some contraction and IA some relaxation. Dixit (1975) reported similar suppression of IA activity for the unvoiced unaspirated plosives of Hindi irrespective of whether they were produced with an open glottis or a closed glottis, indicating that the suppression pattern of the IA muscle favored glottal opening. Unfortunately, PCA recordings made for that were unacceptable.

Thus, from the glottographic data discussed above it appears that the glottal behavior for the unvoiced unaspirated plosives is quite variable in the same subject, between subjects, within a language and across languages.

4.3 Voiced aspirated plosives

Irrespective of their phonetic environment or position in a word, all tokens of the voiced aspirated plosives of Hindi, as expected, were produced with a moderately open glottis (approximately half of that found during unvoiced aspirated plosives). The size of glottal opening for these plosives was relatively greater in the word-initial than in the word-medial or word-final positions; in the latter two positions it was larger for the word-medial than word-final plosives. Also, in the prestressed position the glottal opening for these plosives was larger than in the poststressed or unstressed positions. The opening gesture of the glottis always started during the closure interval, usually well after the articulatory contact; the peak of glottal opening invariably occurred during the noise interval usually well after the articulatory release; and the closing gesture of the glottis, which began at the peak, was regularly completed appreciably after the termination of aspiration during the initial portion of the vowel or voiced segment following these plosives. The vocal folds continued vibrating throughout the closure and noise intervals of the voiced aspirated plosives of Hindi.

These results on the voiced aspirated plosives of Hindi are, generally, in agreement with those reported by Kagaya and Hirose (1975), and Benguerel and Bhatia (1980) on the same Hindi plosives. They are also in agreement with the Maithili (Yadav, 1984) results on similar plosives. These plosives in the present study and Kagaya and Hirose's study occurred in carrier sentences and in vocalic environments, whereas in Benguerel and Bhatia's study and Yadav's study, they occurred word-medially in intervocalic position and word-initially and finally in absolute initial and final positions, respectively, in isolated words. Thus, the direct comparison of glottal dynamics for the voiced aspirated plosives in these studies can be made only for the vocalic environment. In this environment, the main difference in these studies was in the timing of glottal opening gesture with respect to the articulatory release. Benguerel and Bhatia's Hindi data and Yadav's Maithili data show that the opening gesture of the glottis for the voiced aspirated plosives began at or near the articulatory release, whereas in the data of this study and Kagaya and Hirose's Hindi data the opening gesture of the glottis for these plosives usually started close to the middle of the closure interval; only sporadically did it start at the oral release. Thus, it appears that the opening-closing gestures of the glottis during voiced aspirated plosives were executed rather quickly by the subjects of Benguerel and Bhatia's study and Yadav's study. In absolute initial position for the voiced aspirated plosives only the closing gesture of the glottis was observed, which began from a respiratory glottal position and was completed at about articulatory closure, whereas in absolute final position the glottis showed only the opening gesture which started near articulatory release and terminated sometime later in an open position suitable for quiet respiration.

It may be of interest to note that the size and timing of opening-closing gestures of the glottis with respect to oral release for the /b/ + /h/ sequences occurring between vowels in Danish (Frøkjaer-Jensen et al 1971) and English (Lisker and Baer, 1984), were strikingly similar to those observed for the voiced aspirated plosives of Hindi (present study; Kagaya and Hirose, 1975; Benguerel and Bhatia, 1980) and Maithili

(Yadav, 1984). In English /b/ + /h/, like Hindi and Maithili /bh/, the vocal folds continued vibrating throughout the entire closure and noise intervals, but in Danish /b/ (= [β]) + /h/, unlike Hindi and Maithili /bh/, the glottal pulsing was observed only during part of /h/ duration. This is not unexpected, since word-final /b/ is [p] in Danish.

4.4 Unvoiced aspirated plosives

As expected, all unvoiced aspirated plosives of Hindi, irrespective of their phonetic context and position in a word, were produced with a wide open glottis (approximately twice as wide as was observed during voiced aspirated plosives). Glottal opening for these plosives was relatively larger in the word-initial position than in the word-medial position, and in the word-medial position it was larger than in the word-final position. Also, the prestressed plosives showed relatively larger opening than the poststressed or unstressed ones. For these plosives, the opening gesture of the glottis was usually initiated somewhat prior to the articulatory contact and usually peaked at or closely around the articulatory release. The closing gesture of the glottis, which began at the peak, was always completed well after the articulatory release during the early part of the following vowel or voiced segment. Interruption of vocal fold vibration during the unvoiced aspirated plosives always occurred after the glottis had somewhat opened and oral closure had been established, whereas resumption of vocal fold vibration occurred slightly before the termination of aspiration but appreciably before the completion of the closing gesture of the glottis.

Recall that all of the test words in this study were preceded by a vowel and followed by a voiced segment within a carrier sentence. Thus, the results of this study are directly comparable to those from other studies in which unvoiced aspirated plosives occurred in a similar environment. A comparison of glottographic data on the unvoiced aspirated plosives shows that the results of this study on these plosives are consistent with those reported by Kagaya and Hirose (1975) on word-initial and medial, and Benguerel and Bhatia (1980) on word-medial unvoiced aspirated plosives of Hindi. In their studies, these plosives occurred between vowels. Similar results have been reported on word-initial unvoiced aspirated plosives of Danish (Frøkjaer-Jensen et al., 1971; Fukui and Hirose, 1983; Hutters, 1985), Icelandic (Pétursson, 1976; Löfqvist and Yoshioka, 1981), Cantonese (Iwata et al., 1981), Fukienese (Iwata et al., 1979), Mandarin (Iwata and Hirose, 1976), Tibetan (Kjellin, 1977), and on tense heavily aspirated plosives of Korean (Kim, 1967, 1970). Glottographic findings on the unvoiced aspirated plosives in word-initial and word-medial prestressed positions in English (Lisker et al., 1969, 1970; Lisker and Baer, 1984; Sawashima, 1970; Sawashima et al., 1970; Cooper et al., 1971) and

Swedish (Lindqvist, 1972), and in word-medial position in Maithili (Yadav, 1984) and Korean (Kagaya, 1971, 1974) were also similar to our Hindi results on these plosives. As in this study, in all the above studies the unvoiced aspirated plosives occurred in a vocalic environment. Thus, the temporal course of the opening and closing gestures of the glottis for these plosives in relation to the articulatory contact and release was quite comparable across languages.

On the other hand, the unvoiced aspirated plosives in Benguerel and Bhatia's Hindi study (1980) and Yadav's Maithili study (1984) were also produced in absolute word-initial and absolute word-final positions, while in Kagaya's Korean study (1971, 1974), the tense heavily aspirated plosives were also produced in absolute word-initial position. Benguerel and Bhatia's data for these plosives show a simple closing gesture of the glottis from a respiratory glottal position in absolute initial position and a simple opening gesture of the glottis to a respiratory glottal position in absolute final position. However, the data from Maithili and Korean demonstrate that in these positions the glottis does not always behave in the same manner. In absolute initial position, for the unvoiced aspirated plosives the glottis may simply close from a position of respiration, or it may somewhat open and then close, or it may somewhat close then somewhat open and then close again. But the closing gesture of the glottis is always completed considerably after oral release in the three studies, as expected for the aspirated plosives. Likewise, in absolute final position, for the unvoiced aspirated plosives the glottis may simply abduct from the preceding closed position (suitable for a vowel) toward a respiratory position, or it may open and then partially close, or it may open then partially close and then open again. However, the abduction gesture of the glottis always begins near the oral closure in Hindi and Maithili studies, as expected for the voiceless plosives.

Further, Löfqvist and Yoshioka (1981) have reported intervocalic unvoiced preaspirated plosives in Icelandic. Intervocalic voiceless plosives in Swedish are sometimes preaspirated (Lindqvist, 1972). For these plosives, the glottis begins to abduct during the preceding vowel appreciably before the articulatory closure is established, the peak of glottal opening is achieved soon after the closure and the adduction of the glottis, which starts at the peak, is completed during the closure interval near the oral release. Relative to postaspirated plosives the glottal opening is smaller during preaspirated plosives. Also, the timing of opening and closing gestures of the glottis is quite different in the two types of the aspirated plosives.

5. Mechanisms for the Control of Voicing and Aspiration in Plosive Production

How voicing and aspiration associated with certain speech sounds are produced has been a subject of considerable interest for over 2500 years. It appears from the discussion of phonetic observations and concepts of ancient Indian phoneticians and grammarians in Allen (1953) that they recognized the glottis as the source of phonation and that they observed that when the glottis was open breath was produced; when it was closed voice was produced; and when it was in an intermediate position, that is when it was half-open, both breath and voice were produced. Further, breath was discharged for the voiceless consonants; voice was discharged for the voiced consonants and vowels; and both breath and voice were discharged for the voiced glottal fricative and voiced aspirates. Also, more breath was discharged in aspirates than in inaspirates. These observations clearly indicate that voicing and aspiration are glotto-aerodynamic phenomena in that they are generated at the glottis, if the glottis is suitably adjusted and adequate transglottal air flow is present.

Similar observations have been made by many phoneticians of our time about voicing and aspiration and the underlying glottal adjustments used in their control during the production of various categories of plosives. These observations appear to be true to a large extent in a large number of known languages of the world. However, such observation about the plosives occurring in various languages are mostly speculative and inferential since only a small number of languages has been, so far, glottographically examined. It is quite clear from the glottographic results discussed above that all aspirated plosives, whether voiced or unvoiced, are produced with an open glottis; but there are many instances of unaspirated plosives both voiced and unvoiced during which the glottis does not behave the way it was supposed to behave. For example, in some cases of the so-called voiced plosives of English (Lisker et al., 1969, 1970; Sawashima, 1970; Sawashima et al., 1971; Cooper et al., 1971) and almost all cases of the so-called voiced plosives of Danish (Fukui and Hirose, 1983; Hutters, 1985) the glottis revealed some abduction of the arytenoids; on the other hand, many unvoiced unaspirated plosives of Hindi (present study) and almost all unvoiced unaspirated plosives of Cantonese (Iwata et al., 1981), Fukienese (Iwata et al., 1979), Mandarin (Iwata and Hirose, 1976) and Tibetan (Kjellin, 1977) showed a closed or virtually closed glottis. Moreover, the glottis did not vibrate for the most part of the closure interval in all such unvoiced unaspirated plosives and almost all such so-called voiced plosives. These anomalies appear to suggest that the approximated glottis is a necessary but not a sufficient condition for the initiation and/or maintenance of voicing during articulatory closure of the voiced (unaspirated) plosives. Some other mechanism or mechanisms have to be brought into play to maintain adequate transglottal air flow to accomplish this end. On the other hand, abduction of the glottis seems to be a sufficient condition for devoicing or voicelessness of the unvoiced plosives both aspirated and unaspirated. However, for those voiceless (unaspirated) plosives which are produced with an approximated glottis, there must be some mechanism or mechanisms that facilitate devoicing and prevent voicing during their closure intervals.

5.1. Control of voicing in plosives

At least since van den Berg's work on the myoelastic-aerodynamic theory of voice production (1958), it is well known among phoneticians, linguists and speech scientists, that if the vocal folds are appropriately adducted and tensed and the supraglottal vocal tract is relatively unobstructed, an adequate transglottal flow will result in voicing. Such a flow through the glottis can be generated if the magnitude of the pressure drop across the glottis is of the order of 2 to 3 cm H_2O (Ladefoged, 1964; Ishizaka and Matsudaira, 1972; Baer, 1975; Catford, 1977; see also Lindqvist, 1972, according to which voicing could be sustained with only 1 cm H_2O of transglottal pressure drop). On the other hand, voicing will be rapidly interrupted if the supraglottal vocal tract is blocked (as in plosive productions) and the pressure drop across the glottis is eliminated (Rothenberg, 1968; Lindqvist, 1972; Hutters, 1978). Usually, the pressure drop across the glottis is eliminated in about 2 to 4 glottal pulses at normal fundamental frequency during oral occlusion (Rothenberg, 1968; Ohala, 1974, 1975; Catford, 1977). Thus it appears that even when a suitable glottal adjustment for voicing has been made, voicing cannot be initiated and/or maintained during articulatory closure of a voiced stop. Obviously, some glottal or extraglottal mechanism or mechanisms have to be invoked to provide for closure voicing. The importance of such mechanisms for Hindi voiced plosives, irrespective of whether they were aspirated or unaspirated, cannot be overstated, since the closure interval of all phonologically voiced plosives was fully voiced in the present study and the other two Hindi studies (Kagaya and Hirose, 1975; Benguerel and Bhatia, 1980).

Halle and Stevens (1967, 1971) have suggested two glottal mechanisms for prolonging voicing during articulatory closure of the voiced plosives. One of these mechanisms is a slight abduction of the arytenoids which will adjust the position of the vocal folds in such a way as to facilitate their vibration during oral occlusion with relatively low transglottal flow, and the other is slackening of the vocal folds, that is reduction in the stiffness of the vocal folds which will make them more susceptible to oscillation under similar supraglottal and aerodynamic conditions. There is some evidence of arytenoid separation during articulatory closure of the so-called voiced plosives of English (Lisker et al., 1969, 1970; Sawashima, 1970), but it does not support the view advanced by Halle and Stevens, because those tokens of the so-called voiced plosives of English that showed separation of the arytenoids also showed interruption of glottal vibration. In Danish where all phonologically voiced plosives were phonetically realized as voiceless, Hutters (1985) noticed a slight arytenoid separation. A few cases of the voiced plosives of Swedish (Lindqvist, 1972) and Hindi (present study) as indicated before were produced with slight separation of the vocal folds; however, such a separation did not result from arytenoid adjustment in these languages; it rather was a consequence of a rise in supraglottal pressure in Swedish and of an increase in (stress-related) subglottal pressure in Hindi (Dixit and Shipp, 1985). These plosives were voiced throughout. Arytenoid separation observed during the so-called voiced plosives of English and Danish thus seems to have the opposite effect than surmised by Halle and Stevens.

The other mechanism for prolonging closure voicing during voiced plosives suggested by Halle and Stevens (1971) and Stevens (1977) is slackening of the vocal folds, that is reduction in the stiffness of the vocal folds. It is assumed here that Halle and Stevens have used the term "vocal cords" or "vocal folds" in the usual sense that is inclusive of the cover , the transition and the body of the vocal folds. Thus reduction in the stiffness of the vocal folds, that is slackening of the vocal folds, should result from some reduction in the activity of the vocalis muscle and/or lowered larynx position in the neck during voiced plosives. Contrarily, Hirano (1977) has suggested that given the same external tension (resulting from the cricothyroid muscle activity) contraction of the vocalis muscle for smaller loads will slacken the cover (the mucosa and the superficial layer of the lamina propria); conversely, the reduction in the vocalis will stiffen the cover. (Let us asume that, during the production of the voiced plosives, load on the vocalis muscle is relatively smaller.) Following Hirano's suggestion about the increase/decrease in the vocalis activity and its effect on the vocal fold slackening/stiffening, Fujimura (see "Discussion," p. 28, following Hirano's 1977 paper) has argued against Halle and Steven's position and has asserted that given the same external force (longitudinal tension) contraction of the vocalis muscle will slacken the cover of the vocal folds which will "facilitate or trigger" voicing; on the other hand, reduction in the vocalis activity will increase the stiffness of the cover which will be detrimental to voicing. However, the EMG data reported from a number of languages, such as Hindi (Dixit, 1975; Kagaya and Hirose, 1975), English (Hirose and Gay, 1972), Japanese (Hirose and Ushijima, 1978), and Dutch (Collier et al., 1979) revealed similar levels in the activity of the vocalis muscle for voiced and voiceless plosives, which was somewhat less than for the preceding and/or following vowels. This does not provide support for any of the above views on facilitating closure voicing. However, if vocal fold slackening is a matter of larynx lowering, then Steven's (1977) view finds some support in the data reported by Kent and Moll (1969), Perkell (1969), Ewan (1976), Riordan (1978, 1979, 1980) and Westbury (1979, 1983) where the larynx was found to be lower for voiced plosives than for voiceless plosives.

Rothenberg (1968) has suggested three extra-glottal mechanisms which may be used for the maintenance of glottal pulsing during articulatory closure of the voiced plosives: (1) pressure-actuated passive expansion of the supra-glottal cavity, (2) muscularly activated enlargement of the supraglottal cavity and (3) nasal air flow through an incomplete velopharyngeal closure. A fourth extraglottal mechanism for sustaining voicing during oral occlusion of the voiced plosives has been suggested by Westbury (1979, 1983). According to him "the requisite pressure gradient" in favor of transglottal air flow "might be maintained if subglottal pressure were increased rapidly and in concert with the supraglottal pressure rise which naturally accompanies vocal tract occlusion." As Westbury has himself acknowledged, there is no support for this mechanism in the EMG data from the expiratory muscles (Draper et al., 1959, 1960; Ladefoged, 1962, 1968), nor is there any support in the subglottal air pressure data (Netsell, 1969: McGlone and Shipp, 1972; Ohala and Ohala, 1972; Löfqvist, 1974; Nihalani, 1974; Westbury and Niimi, 1979; Dixit and Shipp, 1985). In fact the levels of subglottal air pressure during voiced and voiceless plosives have been found to be generally similar.

The three extraglottal mechanisms proposed by Rothenberg (1968) have been found to play an important role in maintaining the transglottal air flow necessary for vocal fold vibration to occur during articulatory closure of voiced plosives. Perkell (1969), Kent and Moll (1969), and Westbury (1979, 1983) have reported cinefluorographic data which revealed a comparatively larger supraglottal cavity for the voiced plosives than for the unvoiced plosives. The observed difference in the size of the supraglottal cavity between voiced and voiceless plosives has been attributed by these authors to passive expansion as well as active enlargement of that cavity. In the same vein, on the basis of their electromyographic data Bell-Berti and Hirose (1971) observed that increase in the size of supraglottal cavity during voiced plosives as against unvoiced plosives was neither exclusively passive nor active, both passive and active modes of supraglottal cavity expansion and enlargement were used by all subjects with varying preference in their experiments (see also Bell-Berti, 1975, and Bell-Berti and Hirose, 1975). Further, the larynx (Perkell, 1969; Kent and Moll, 1969;

Ewan and Krones, 1974; Ewan, 1976; Riordan, 1978, 1979, 1980; Westbury, 1979, 1983), the velum (Bell-Berti, 1975; Bell-Berti and Hirose, 1971, 1975), the tongue root (Perkell, 1969; Kent and Moll, 1969; Westbury, 1979, 1983) and the tongue dorsum (Perkell, 1969; Westbury, 1979, 1983) have been shown to be involved in cavity-enlarging maneuvers. With the exception of Ewan's thyroumbrometric data on larynx height during various consonants of English and a number of other languages including Hindi, all other data in the cited studies are from English. To the author's knowledge, there are no cinefluorographic data reported or published on Hindi. Hence it cannot be said with any degree of certainty what supraglottal cavity-enlarging maneuvers are used by speakers of Hindi for maintaining voicing during voiced plosives. However, since all voiced plosives were fully voiced in the three Hindi studies and since the duration of the closure interval for these plosives was much longer than the one that could be kept voiced unaided, it is suspected that an increase in the size of supraglottal cavity should play an important role as a voicing mechanism by preventing equalization of subglottal and supraglottal pressures during articulatory closure of the voiced plosives of Hindi, and that such an increase in the size of the supraglottal cavity cannot simply be passive. Besides Ewan's thyroumbrometric data on laryngeal height, there are some EMG data from the levator palatini muscle (Dixit, 1981) and some supraglottal pressure data (Dixit and Brown, 1978) obtained during Hindi plosives which can be exploited to assess the nature of supraglottal cavity expansion/enlargement during voiced plosives of Hindi.

As indicated by Rothenberg (1968), Ishizaka et al. (1975), Fant et al. (1976), and Westbury (1979, 1983) the walls of the supraglottal cavity are compliant in varying degrees depending on the nature of the tissue that covers these walls and the state (active/inactive) of the underlying musculature. Thus, the walls of the supraglottal cavity may be more or less yielding to the build-up of supraglottal pressure during articulatory closure of a plosive.

Simulated supraglottal pressure profiles during oral closure of a hypothetical bilabial plosive reported by Westbury (1983) are very interesting. With different degrees of wall compliance, which were mechanically analogous to rigid wall, neck wall, tensed cheeks, and relaxed cheeks, his model (a modified version of Rothenberg's model) generated four pressure profiles that were quite comparable to those observed in actual supraglottal pressure data on plosives reported from a few languages including Hindi. Similar simulated pressure profiles were earlier reported by Müller and Brown (1980). The pressure profiles for the voiced plosives of Hindi, both aspirated and unaspirated, displayed in Dixit and Brown (1978) appear to be a combination of the neck wall to the relaxed cheeks pressure profiles illustrated in Westbury (1983). In the first 20 to 45 ms of the closure interval the pressure profiles for the voiced plosives were similar to those simulated for the neck wall and tensed cheeks conditions, that is they were somewhat convex, and then during the rest of the closure interval they were like those simulated for the relaxed cheeks condition; that is they were linear and gradual.

The values of closure voicing for the first two conditions reported in the simulation studies (Westbury, 1983; Keating, 1984) vary from 25 to 60 ms and for the last condition from 80 ms (Westbury, 1983) to 145 ms (Keating, 1984). These values are for the bilabial plosives. Values of 20 to 30 ms of closure voicing, resulting from the passive expansion of the supraglottal cavity during bilabial and retroflexed closures, have been reported by Rothenberg (1968). Ohala and Riordan (1979) carried out an experiment in which they controlled intra-oral pressure during articulatory closure of various voiced plosives through a nasal catheter which could be opened or closed at unpredictable moments. Using this technique they estimated the duration of closure voicing resulting from the passive expansion of the supraglottal cavity. Their values for closure voicing of intervocalic /b/ varied from 69.73 to 90.91 ms depending on its vocalic environment. These values are close to those reported by Westbury. Thus the values for closure voicing resulting from passive expansion of the supraglottal cavity reported by Rothenberg appear to be too conservative, while those reported by Keating are considerably inflated. If the values of closure voicing resulting from passive expansion of the supraglottal cavity could be as high as 145 ms for /b/ then there is no need to invoke active cavity-enlarging maneuvers for the maintenance of closure voicing during most instances of /b/ in most languages. In almost all cases of Hindi /b/, duration of the closure interval was less than 145 ms.

There are some EMG data from the muscles of the velopharyngeal region (Minifie et al., 1974; Bell-Berti and Hirose, 1971, 1975; Bell-Berti, 1975), such as constrictor muscles of the pharynx, palatoglossus and palatopharyngeus. These data show less activity in these muscles for the voiced plosives than for the unvoiced plosives indicating that the pharyngeal walls were more compliant during voiced plosives than unvoiced plosives; but it is not known in absolute terms how compliant the walls of the supraglottal cavity really are. It may be that the muscles that underlie the surfaces of certain vocal tract parts are never completely relaxed during speech. Thus, estimates of passive closure voicing, particularly those for the "relaxed cheeks" condition should be treated with caution.

In the light of the different types of data discussed above it seems safe to say that all phonetically voiced plosives, particularly those with robust closure voicing, must involve some active enlargement of the supraglottal cavity. Following oral closure of intervocalic voiced plosives, larynx lowering (Perkell, 1969; Kent and Moll, 1969; Ewan and Krones, 1974; Ewan, 1976; Riordan, 1978, 1979, 1980; Westbury, 1979, 1983), tongue root advancing (Perkell, 1969; Kent and Moll, 1969; Westbury, 1979; 1983), tongue dorsum depressing (Perkell, 1969; Westbury, 1979, 1983), and soft palate raising (Bell-Berti and Hirose, 1971, 1975; Bell-Berti, 1975) have been observed in English. These maneuvers actively enlarge the size of the supraglottal cavity and thus assist in the maintenance of adequate transglottal pressure drop necessary for sustaining voicing during oral occlusion of the voiced plosives. It should, however, be mentioned that Bell-Berti and Hirose's EMG results on soft palate raising for the maintenance of closure voicing during voiced plosives have not been confirmed (see Lubker et al., 1970; Dixit, 1981; Westbury, 1979, 1983). Except for Ewan's thyroumbrometric data on larynx height, where the larynx has been shown to be somewhat lower during voiced plosives as compared to voiceless plosives of Hindi (besides other languages), there is no other direct evidence of active enlargement of supraglottal cavity during voiced plosives of Hindi. However, Ewan's data on Hindi, though limited, clearly shows that the assumed active enlargement of the supraglottal cavity during the production of the voiced plosives of Hindi does occur.

Nasal escape of air during articulatory closure of plosives as a voicing mechanism was investigated by Lubker (1973) on English subjects. He observed "that nasal air flow is present for some non-nasal consonants, produced by some speakers, under certain conditions." But because it was not observed under all conditions and for all speakers, he concluded "that nasal air flow during non-nasal consonant production is most probably not the result of a velopharyngeal leak and is not a major factor in voicing." A similar conclusion was reached by Westbury (1979, 1983).

Rothenberg (1968, and personal communication) studied this phenomenon in English as well as Hindi speakers. The tracings of nasal air flow that he sent clearly showed that oscillations during the articulatory closure of English /b/, which was produced without velopharyngeal leakage, were reduced to minimal amplitude because of the absence of nasal air flow. On the other hand, Hindi /d/ and /g/, which were produced with a slight velopharyngeal opening, exhibited relatively higher amplitude oscillations during oral occlusion. However, the nasalization decreased during closure to terminate just before Hindi /d/ and /g/ were released. According to Rothenberg, this was necessary in order to allow pressure build-up in the supraglottal cavity for the release of /d/ and /g/. Similar results on nasal air flow during closure of voiced plosives of Sindhi were reported by Nihalani (1975). However, Dixit and MacNeilage (1972) reported that the oscillations picked up by a nasal probe microphone and an oral microphone continued throughout the closure interval of the voiced plosives, both aspirated and unaspirated, produced by three Hindi speakers. Recent preliminary nasal air flow recordings from a Hindi speaker made at the UCLA Phonetics Laboratory reveal similar trends as to the nasalization of the voiced plosives during their closure interval.

A few years ago (1981) at Haskins Laboratories during examination and filming of the velopharyngeal port, while the author produced Hindi words containing voiced and voiceless consonants, slight lateral openings for the voiced plosives (in descending order: b, bh > d, dh > g, gh) were consistently observed by the experimenter (Dr. Kiyoshi Honda of the Kanazawa Institute of Technology, Kanazawa, Japan). Accordingly, the level of EMG activity in the levator palatini muscle (the prime mover of the soft palate) for the voiced plosives of Hindi was regularly lower than for the unvoiced plosives as reported by Dixit (1981). Since nasality in Hindi is contrastive in vowels as well as consonants, it is surprising to see such a widespread use of nasality in non-nasal consonants. This appears to suggest that the velopharyngeal leak in Hindi voiced plosives is an active and precise maneuver and therefore a learned behavior to create the necessary pressure drop across the glottis to sustain glottal pulsing through oral closure. The extent to which the velopharyngeal port is permitted to open during articulatory closure of the voiced plosives in Hindi should have a certain limit beyond which it will not be possible to produce a perceptually non-nasal plosive.

5.2. Control of devoicing in plosives

Voicing/voicelessness in plosives has been, generally, attributed to the adduction/abduction dimension of glottal adjustment. If the glottis is adducted, vocal fold vibration can be maintained for about 60 to 80 ms during articulatory occlusion of a voiced plosive simply by the compliance of the tissue surrounding the supraglottal cavity, which prevents equalization of the subglottal and supraglottal pressures and allows vocal fold pulsing (Westbury, 1979, 1983; Ohala and Riordan, 1979; Keating, 1984). In the absence of tissue compliance, the vocal fold vibration ceases in about 2 to 4 pitch periods (Rothenberg, 1968; Westbury, 1979, 1983). On the other hand, if the glottis is abducted during articulatory closure of a plosive, voicing in most cases ceases quickly in a few pitch periods (2 to 4) because of the rapid build-up of supraglottal pressure which eliminates the pressure gradient in favor of transglottal air flow. This has been found to be true for all voiceless aspirated plosives in all languages and for most voiceless unaspirated plosives in Dutch, English, French, Hindi, Maithili and Swedish (see studies cited in section 4). However, there is an equally large number of voiceless unaspirated plosives during which the glottis was either entirely approximated or showed a small spindle-shaped opening in the membranous portion. This has been observed in the present study on Hindi and the studies on Cantonese, Danish, Fukienese, Icelandic, Mandarin and Tibetan (cited in section 4). In such unvoiced unaspirated plosives, the voicing should be expected to continue at least for 60 to 80 ms during their closure interval by virtue of the observed glottal state. However, in these plosives like other voiceless plosives the vocal fold vibration ceases in a few pitch periods after articulatory closure is established. Thus there must be some mechanism or mechanisms that prevent vocal fold vibration during articulatory occlusion of such plosives although the observed glottal state favors fold pulsing.

A glottal stop or glottalization gesture occurs quite commonly with the voiceless plosives of English in certain contexts. Fujimura and Sawashima (1971), and Westbury (1979, 1983) have reported the occurrence of a glottalization gesture or glottal stop with the first member of certain medial consonant clusters or sequences when it was the unaspirated variant of voiceless plosives in English. In such cases the adducted false vocal folds covered the closed glottis. Voicelessness in these plosives was ascribed to the glottalization gesture or the glottal stop. Glottalization as a devoicing mechanism is, perhaps, used in the unvoiced unaspirated plosives (at least in the syllable-final position where they are glottalized) in Cantonese and Fukienese also. However, glottalization does not seem to be used as a devoicing mechanism in Hindi, Danish and Icelandic, where unvoiced unaspirated plosives are produced with a nearly closed glottis.

Another devoicing mechanism, a rather controversial one, has been suggested by Hutters (1984) for the phonetically unvoiced unaspirated plosives of Danish. According to her "... the activity in the PCA and INT muscles is a devoicing mechanism in unaspirated stops rather than a means to open and close the glottis. This implies that the slight abduction of the arytenoid cartilages that is seen in the Danish /b, d, g/ ([p, t, k]) should be considered a by-product of the vocal fold adjustment that causes the vocal fold vibration to die out, rather than a goal in itself." A bit of supporting evidence for Hutters' hypothesis comes from the physiological data reported by Lisker and Baer (1984) where a word-final unvoiced unaspirated bilabial plosive before a vowel was produced without glottal opening during oral closure, although a reciprocal activity pattern in the PCA and INT favored glottal opening. Whether this mechanism was used for devoicing the closure of similarly produced bilabial plosives of Hindi cannot be said in certain terms since PCA activity for Hindi plosives is unavailable, although some reduction in the INT activity was observed (Dixit, 1975).

Halle and Stevens (1971) proposed stiffness of the vocal folds as a devoicing mechanism for the unvoiced unaspirated plosives which are produced with approximated glottis, that is unspread, unconstricted glottis. EMG data reported by Dixit (1975) and Dixit and MacNeilage (1980) revealed high levels in the CT activity for the voiceless plosives of Hindi, while for the voiced plosives the CT activity was more or less suppressed. The high levels in the CT activity for these (voiceless) plosives were thought to facilitate closure voicelessness. Kagaya and Hirose (1975) reported a small peak in the CT activity for the unvoiced unaspirated plosives of Hindi, partially confirming the results reported by Dixit (1975), and Dixit and MacNeilage (1980). These data provide some support for Halle and Stevens' hypothesis of stiff vocal folds as a devoicing mechanism.

It is of interest to note that all those unvoiced unaspirated plosives of Hindi that were produced with an approximated glottis, were also produced with a high level of supraglottal air pressure build-up, like their voiceless aspirated counterparts, which were produced with a wide open glottis (Dixit and Brown, 1978). For a similarly produced unvoiced unaspirated bilabial plosive of English, Lisker and Baer (1984) reported elevated supraglottal pressure. The pressure profile in their data reveals a somewhat gradual elevation of supraglottal pressure, presumably indicating glottal leak. However in the Hindi data supraglottal pressure elevation during unvoiced unaspirated plosives was as rapid as during voiceless aspirated plosives, which points to yet another mechanism of closure devoicing, that is the reduction in the vocal tract volume by increasing the level of activity in the muscles underlying surface tissue in certain parts of the vocal tract, particularly the pharynx, and by raising the larynx in the neck. Kent and Moll (1969), Perkell (1969), Westbury (1979, 1983), Ewan (1976), Riordan (1978, 1979, 1980) have all reported a relatively higher position of the larynx in the neck for the voiceless consonants than for the voiced consonants. Furthermore, Minifie et al., (1974) reported a relatively higher level of activity in the superior and medial constrictor muscles of the pharynx. Similar results on these and the palatoglossus and palatopharyngeus muscles were reported by Bell-Berti and Hirose (1971, 1975) and Bell-Berti (1975). These data suggest an increase in the tension and decrease in the overall volume of the vocal tract. The effect of such maneuvers on the speed of supraglottal pressure build-up during voiceless aspirated plosives may be negligible since the glottis is wide open during their production, but such maneuvers, in my judgment, will play a crucial role in the rapid build-up of supraglottal pressure during unvoiced unaspirated plosives if the glottis is adducted during their production. It appears that the rapid rise in the supraglottal pressure during unvoiced unaspirated plosives of Hindi that were produced with the approximated glottis was primarily the result of rapid reduction in the supraglottal volume which in turn compressed the air trapped in the vocal tract at the point of articulatory occlusion and thus raised the pressure. However, some slight glottal leak contributing to pressure rise in these plosives cannot be entirely precluded, since the glottis was not tightly closed.

5.3. Control of aspiration in plosives

As was indicated earlier, the behavior of the glottis with respect to supraglottal articulatory events is largely conditioned by the phonetic-acoustic nature of various speech sounds, the aerodynamic requirements for which may differ from one sound to the other. For example, considerably greater air flow rate is required for the production of aspirated plosives than unaspirated plosives. The most important factor in the production of aspiration is the higher than normal rate of air flow through the glottis. The glottis must be open during those sounds that require greater than normal air flow rate, for instance aspirated plosives. In fact, the open glottis is said to be a necessary condition for the production of aspiration (Chomsky and Halle, 1968). However, it is not a sufficient condition; in addition the supraglottal vocal tract must be unobstructed. Unless the supraglottal vocal tract is unobstructed while the glottis is open, higher than normal air flow rate cannot be generated, which is so important in the generation of aspiration.

Thus, the three physiological conditions that control generation of aspiration appear to be: (1) the open glottis, (2) the unobstructed supraglottal vocal tract, and (3) the greater than normal air flow rate through the glottis relative to the size of glottal aperture (Dixit, 1979, 1983). In the production of aspirated plosives if the first two conditions are met, then the third condition is automatically met, assuming that the function of the respiratory system during speech is to provide an undifferentiated uniform air stream. The variations that occur in the air flow rate seem to result from the variations in the glottal and supraglottal resistances (Ohala, 1974, 1975; Ohala et al., 1979; Ohala, 1979; Dixit and Shipp, 1985; Dixit and Brown, 1985).

As indicated above, the aspirated plosives are produced with an open glottis, so that the glottal resistance during their production is quite low (in the voiced aspirated plosives during their plosive phase, while in the voiceless aspirated plosives during their plosive as well as occlusive phase). The articulatory resistance, which is infinite during their occlusive phase, abruptly drops to a minimum level immediately following the release of the articulators. Clearly, during the plosive phase of the aspirated plosives there is low resistance at the glottis and little or no resistance elsewhere in the vocal tract to the pulmonary air. Hence, pulmonary air which is under pressure during the occlusive phase rushes out through the glottis in large volume adequate for the generation of turbulence or frictional noise at the glottis, which is traditionally called aspiration.

From the above discussion it is quite obvious that the required (greater than normal) breath flow in the production of aspirated plosives results from the temporal coordination between the glottal and the supraglottal articulatory events. It should, however, be recalled that neither the size of glottal opening nor the timing of the glottal gesture with respect to the supraglottal articulatory events are in any sense similar in the voiced aspirated and the unvoiced aspirated plosives of Hindi. They cannot be, if the identity of the former as voiced and that of the latter as voiceless has to be preserved. On the other hand, for either one of these plosives individually, the temporal coordination in the glottal and supraglottal articulatory events is not (see Figure 5) and need not be precise. As long as their identity as voiced and voiceless and the perceptually required minimal degree of aspiration are preserved, the speaker can afford to be sloppy within certain limits in the temporal coordination of the glottal and supraglottal events during the production of aspirated (voiced and voiceless) plosives.

From the foregoing discussion it should be clear that the turbulence or frictional noise that is heard as aspiration is generated at the glottis. Thus it can be safely said that the aspiration associated with the plosives and certain other consonantal sounds is laryngeally controlled. However, the function of the larynx in the production of aspiration is constrained by the condition of the supraglottal vocal tract. If the supraglottal vocal tract is obstructed or occluded, the function of the larynx, to be specific the glottis, is rendered ineffectual. For the production of aspiration the supraglottal vocal tract must be unobstructed while the glottis is open to a certain degree.

It appears from Kim's (1970) widely quoted paper on aspiration that he also believes in the laryngeal control of aspiration. According to him "aspiration is nothing but a function of the glottal opening at the time of release." Thus, he assumes that there is a "direct correlation between the degree of glottal opening at the time of release and the degree of aspiration." Further, "what is controlled by the laryngeal muscles in the case of aspiration is not the timing of closing ... but the size of the glottal opening ..." "... the instruction to close the glottis is assumed to be simultaneous for all voiceless stops (i.e., it occurs at the time of release) ..."

There are some basic differences in our and Kim's description and discussion of laryngeal control of aspiration. Kim's assumptions are quite interesting and, to the unguarded reader, seem also quite convincingly supported by his data. Nevertheless the glottographic data presented in this study and the other cited studies cast serious doubts on some of them. There is no doubt that aspiration is laryngeally controlled and that it is a function of glottal opening. However, there seems to be no direct correlation between the degree of glottal opening and the degree of aspiration. In our data we found many examples of plosives which differed in the degree of glottal opening at the time of articulatory release but had about the same degree of aspiration and vice-versa. This is even reflected in the averaged superimposed glottographic curves in Figure 4 where unvoiced aspirated plosives in the word-final position had a longer period of aspiration than the same plosives in word-medial or word-initial positions; however the degree of glottal opening at the time of oral release for these plosives in the word-final position was relatively smaller than in the other two word positions. Also, the degree of glottal opening at the time of oral release for the voiced aspirated plosives was much smaller than for the voiceless aspirated plosives, but the degree of aspiration for the two types was about the same. Similarly, Sawashima (1968a) observed that "the voicing lag is not completely determined by the glottal opening at the articulatory release." (It should be noted that the author does not consider "voicing lag" to be aspiration.)

Kim's assumption that the instruction to close the glottis is given simultaneously at the moment of articulatory release for voiceless plosives implies a uniform rate of glottal closing. Neither the assumption nor its implication finds any support in the glottographic data. The closing gesture of the glottis which begins at the peak during voiceless and/or aspirated plosives has been observed to be variously timed. The peaks of glottal opening do not occur at the same time. Moreover the rate of glottal closing gesture is not constant. Kagaya (1971, 1974) and Hirose et al. (1973) have shown that the glottis closes more rapidly for the tense aspirate (after release) and unaspirated than for the lax slightly aspirated plosives of Korean. Thus, Kim's major claims about how the larynx controls aspiration remain unsupported in the glottographic data.

Furthermore, Kim has contended that since there is no glottal constriction during aspiration, "turbulence is created not at the glottis but at the point of constriction for the following vowel whose configuration is formed, through coarticulation, ..." Thus, he asserts that the traditional definition of aspiration "as a glottal friction" is "completely misleading." Obviously, for Kim aspiration is not glottal friction but rather cavity friction--friction generated at some point in the supraglottal cavity. However, acoustic data reported by Fischer-Jørgensen (1954), Halle et al. (1957), Fant (1973), Kagaya (1974), Ladefoged (1975), and Dixit (1979, 1983) do not support Kim's contention; they rather show a resonance pattern of the aspiratory turbulence, which would be possible only if the source of such turbulence were at the glottis, not at some point in the supraglottal vocal tract. Moreover, aspiration is considered to be a type of phonation (Ladefoged, 1971; Catford, 1977; Dixit and Ladefoged, forthcoming) and phonation by definition is generated at the glottis.

In their discussion of "source features," Chomsky and Halle (1968) stated that "....tense sounds are produced with greater subglottal pressure and this fact accounts for the well-known presence of aspiration in the tense voiceless stops of many languages. Since, however, the tenseness of the supraglottal muscles is evidently controlled by a different mechanism <u>than is tenseness in the subglottal cavities</u> [author's emphasis], these two properties cannot be combined into a single phonetic feature. Instead we must set up in addition to tenseness a feature of 'heightened subglottal pressure'." They further noted that "heightened subglottal pressure is a necessary but not a sufficient condition for aspiration. Aspiration requires, in addition, that there be no constriction at the glottis. If there is a glottal constriction aspiration will not occur." Apparently, Chomsky and Halle believe in the joint control of aspiration by the laryngeal and sublaryngeal (respiratory) systems. As far as the laryngeal control of aspiration is concerned there is a general agreement; however, sublaryngeal control of aspiration has been generally rejected. Ohala's (1974, 1975) model of speech aerodynamics does not support Chomsky and Halle's belief that heightened subglottal pressure, if it occurs during aspirated plosive production, is of respiratory origin. In discussing linguistic aspects of respiratory phenomena, Ladefoged (1968) observed that "for nearly all subjects, the respiratory muscles do not cause variations of subglottal pressure that can be associated with the production of particular consonants ..." EMG data on expiratory muscles reported by Draper et al. (1959, 1960), Hoshiko (1960, 1962), Ladefoged (1962, 1967), Munro and Adams (1971), and Adams and Munro (1973) also point in the same direction. Furthermore, subglottal air pressure data from actual languages reported by Netsell (1969), Ohala and Ohala (1972), Löfqvist (1974), Nihalani (1974), and Dixit and Shipp (1985) did not reveal substantial or significant differences for the aspirated plosives vis-a-vis the unaspirated ones. Thus, Chomsky and Halle's belief that heightened subglottal pressure is a necessary condition for aspiration and that in addition to the laryngeal system the sublaryngeal system also plays a role in the control of aspiration seems to be unfounded.

On the other hand, Halle, Hughes and Radley (1957) have attributed aspiration to the higher supraglottal pressure build-up during oral closure of English plosives. They have suggested that "Although in many instances the presence or absence of voicing serves to distinguish /b/ /d/ /g/ from /p/ /t/ /k/, in English voicing is not crucial to this distinction. The essential difference between these two classes of stops lies in the fact that in the production of the latter more pressure is built up behind the closure than in the production of the former. This difference in pressure results in higher intensity bursts and accounts for the well-known fact that /p/ /t/ /k/ bursts are often followed by an aspiration (author's emphasis), which is not present in the case of /b/ /d/ /g/." The supraglottal air pressure data on Hindi plosives reported by Dixit and Brown (1978) showed similar pressure profiles and peaks for the aspirated and unaspirated voiceless plosives, whereas pressure profiles and peaks were quite different for the voiced aspirated and voiceless aspirated plosives which had about the same degree of aspiration. Obviously, the magnitude of supraglottal pressure build-up during the articulatory closure of a plosive has nothing to do with its being aspirated. A similar observation was made earlier by Lisker (1970) that "the amount of aspiration, in so far as it is a matter of duration, is not determined entirely by peak pressure..."

6. Conclusions

From the data presented and discussed in this study the following conclusions can be arrived at concerning Hindi stops. Our survey of the literature indicates that they are also probably valid for similarly labeled stops in other languages.

(1) The voiced unaspirated plosives are, generally, produced with an approximated glottis and vibrating vocal folds throughout their closure interval.

(2) The unvoiced unaspirated plosives may be produced either with an open or an approximated glottis but non-vibrating vocal folds. When they are produced with an open glottis, the glottis generally begins to open somewhat prior to the articulatory contact or closure, peaks at or near the middle of the closure interval and closes at or immediately after the articulatory release.

(3) The voiced aspirated plosives are produced with an approximated glottis during part of the closure interval and with an open glottis during part of the closure interval and all of the noise interval. The glottis, generally, begins to open around the middle of the closure interval, peaks around the middle of the noise interval and closes much after the articulatory release during the early part of the following voiced segment. The mode of vocal fold vibration remains normal during most of the closure interval then changes to what Ladefoged calls murmur, particularly after the articulatory release. (This should not be misconstrued to suggest that murmur can occur during the closure interval; it definitely cannot.)

(4) The unvoiced aspirated plosives are produced with a wide open and quiescent glottis. The glottal opening is generally initiated somewhat earlier than the articulatory occlusion, the peak of opening occurs at or around the instant of articulatory release, and the glottal closure occurs much after the articulatory release during the initial portion of the following voiced segment.

(5) The timing of the initiation of the glottal opening gesture in relation to the articulatory closure is generally similar for the unvoiced unaspirated and the unvoiced aspirated plosives but different for the voiced aspirated plosives; the timing of the completion of the glottal closing gesture in relation to the articulatory release is generally similar for the voiced aspirated and the unvoiced aspirated plosives but different for the unvoiced unaspirated plosives; and the timing of the peak glottal opening with respect to the articulatory release is different for all three types of plosives.

(6) The degree of glottal opening systematically decreases in descending order from the unvoiced aspirated plosives to the voiced aspirated plosives to the unvoiced unaspirated plosives. The glottis is wide open during the unvoiced aspirated plosives, moderately open during the voiced aspirated plosives and narrowly open during the unvoiced unaspirated plosives.

(7) For those plosives that are produced with an open glottis, the degree of glottal opening was greater in the word-initial prestressed position than in the word-medial postressed and word-final poststressed positions; in the latter two positions, the word-medial plosives showed a somewhat greater glottal opening than word-final plosives.

(8) The degree of glottal opening does not seem to be systematically related to

either the extent of the duration of closure interval or the extent of the duration of noise interval.

(9) No apparent direct correlation between the degree of glottal opening at the instant of articulatory release and the degree (duration) of aspiration (noise interval) is observed.

(10) Voicing and aspiration are laryngeally controlled in the sense that they both are generated at the glottis and that the glottis is adducted for voicing and abducted for aspiration.

(11) An adducted glottis is a necessary condition for (normal) voicing but not a sufficient condition. In addition an adequate transglottal pressure drop should be created and maintained for the voicing to occur and to continue.

(12) An abducted glottis is a necessary condition for aspiration but not a sufficient condition. In addition the supraglottal vocal tract must be unobstructed and the breath flow through the glottis must be greater than normal relative to the size of glottal aperture.

(13) In general, the glottal adduction/abduction plays the most important role in the control of closure voicing/devoicing but they may be rendered ineffectual in the absence of certain extraglottal mechanisms.

(14) Aspiration associated with the production of plosives is not controlled by the sublaryngeal (respiratory) system.

*On leave from Louisiana State University, Baton Rouge, Louisiana 70803-2606

Acknowledgements

Special thanks to Peter MacNeilage and Peter Ladefoged. The data for this study were collected at the Phonetics Laboratory of the University of Texas at Austin. I am very grateful to Peter MacNeilage for providing the facilities and making the arrangements for carrying out this experiment. This paper was written while I was visiting UCLA Phonetics Laboratory. I am deeply indebted to Peter Ladefoged for the invitation to UCLA and for his insightful comments on the earlier draft of this paper. This work was supported in part by NSF and NIH grants.

References

Abramson, A.S., and L. Lisker (1972). Voice timing in Korean stops. *Proceedings of the Seventh International Congress of Phonetic Sciences*, pp. 439-46. The Hague: Mouton.

Abramson, A.S., L. Lisker and F.S. Cooper (1965). Laryngeal activity in stop consonants. *Haskins Labortories Status Report on Speech Research* SR-4, 6.1-6.12.

Adams, C., and R.R. Munro (1973). The relationship between intercostal muscle activity and pause placement in the connected utterances of native and non-native speakers of English. *Phonetica* 28: 227-50.

Allen, W.S. (1953). *Phonetics in Ancient India*. London: Oxford University Press.

Armstrong, L.E. (1959). *The Phonetics of French*. London: G. Bell and Sons, Ltd.

Baer, T. (1975). Investigation of Phonation Using Excised Larynxes. Doctoral dissertation, M.I.T.

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *Journal of the Acoustical Society of America* 57: 456-64.

Bell-Berti, F., and H. Hirose (1971). Velopharyngeal function in oral/nasal articulation and voicing gesture. *Haskins Laboratories Status Report on Speech Research* SR-28: 143-56.

Bell-Berti, F., and H. Hirose (1975). Palatal activity in voicing distinction: a simultaneous fiberoptic and electromyographic study. *Journal of Phonetics* 3: 69-74.

Benguerel, A.-P., and T.K. Bhatia (1980). Hindi stop consonants: an acoustic and fiberscopic study. *Phonetica* 37: 134-48.

Benguerel, A.-P., H. Hirose, M. Sawashima and T. Ushijima (1978). Laryngeal control in French stop production: a fiberscopic, acoustic and electromyographic study. *Folia Phoniat.* 30: 175-98.

Bloch, B. (1950). Studies in colloquial Japanese IV: phonemics. Language 26: 85-125.

Bronstein, A.J. (1960). *The Pronunciation of American English: An Introduction to Phonetics*. New York: Appleton-Century-Crofts.

Catford, J.C. (1977). *Fundamental Problems in Phonetics*. Bloomington: Indiana University Press.

Chao, Y.R. (1934). Zhōng-guó fāng-yán dāng-zhōng bào-fā-īn de zhǒng-lèi (Classification of the stops in Chinese dialects). *Bulletin of the Institute of History and Philology, Academica Sinica* 515-520.

Chao, Y.R. (1947). *Cantonese Primer*. Cambridge (Massachusetts): Harvard

University Press.

Chao, Y.R. (1949). *Mandarin Primer.* Cambridge (Massachusetts): Harvard University Press.

Chomsky, N., and M. Halle (1968). *The Sound Pattern of English.* New York: Harper and Row.

Cohen, A., C.L. Ebeling, K. Fokemma and A.G.F. von Holk (1961). *Fonologie van het Nederlands en het Fries.* The Hague: s'-Gravenhage.

Collier R., L. Lisker, H. Hirose, and T. Ushijima (1979). Voicing in intervocalic stops and fricatives in Dutch. *Journal of Phonetics* 7: 357-73.

Cooper, F.S., M. Sawashima, A. S. Abramson and L. Lisker (1971). Looking at the larynx during running speech. *Annals of Otology, Rhinology and Laryngology* 80: 678-82.

Delattre, P. (1966). *Studies in French and Comparative Phonetics, Selected Papers in French and English.* The Hague: Mouton.

Dixit, R.P. (1963). The segmental phonemes of contemporary standard Hindi. Master's thesis, University of Texas at Austin.

Dixit, R.P. (1975). Neuromuscular Aspects of Laryngeal Control: with Special Reference to Hindi. Doctoral dissertation, University of Texas at Austin.

Dixit, R.P. (1979). Aspiration: what is it and how is it produced? In J. Wolf and D. Klatt, *Speech Communication Papers*, 83-8. New York: Acoustical Society of America.

Dixit, R.P. (1981). Velar activity during speech: an electromyographic study. *Journal of the Acoustical Society of America* Supplement 1, Vol. 69, S54.

Dixit, R.P. (1983). On defining aspiration. *Proceedings of the XIIIth International Congress of Linguistics*, 606-10, Tokyo.

Dixit, R.P., T. Baer and K. Honda (1984). Transillumination of the glottis during Hindi obstruents. *Journal of the Acoustical Society of America* Supplement 1, Vol. 75, S7.

Dixit, R.P., and W.S. Brown (1978). Peak magnitudes of supraglottal air pressure associated with affricated and nonaffricated stop consonant productions in Hindi. *Journal of Phonetics* 6: 353-65.

Dixit, R.P., and W.S. Brown (1985). Peak magnitudes of oral air flow during Hindi stops (plosives and affricates). *Journal of Phonetics* 13: 219-34.

Dixit, R.P., and P. Ladefoged (forthcoming). An experimental study of Hindi phonation types.

Dixit, R.P., and P.F. MacNeilage (1972). Coarticulation of nasality: evidence from Hindi. Paper presented at the 83rd Meeting of the Acoustical Society of America, Buffalo, New York.

Dixit, R.P., and P.F. MacNeilage (1974). Glottal dynamics during Hindi bilabial plosives and the glottal fricative. *Journal of the Acoustical Society of America* Suppl.1, Vol.55: S80.

Dixit, R.P., and P.F. MacNeilage (1980). Cricothyroid activity and control of voicing in Hindi stops and affrictes. *Phonetica* 37: 397-406.

Dixit, R.P., and T. Shipp (1985). Study of subglottal air pressure during Hindi stop consonants. *Phonetica* 42: 53-77.

Draper, M.H., P. Ladefoged and D. Whitteridge (1959). Respiratory muscles in speech. *Journal of Speech and Hearing Research* 2: 16-27.

Draper,M.H., P. Ladefoged and D. Whitteridge (1960). Expiratory pressure and airflow during speech. *British Medical Journal* 18: 1837-1843.

Ewan, W.G. (1976). Laryngeal Behavior in Speech. Doctoral dissertation, University of California at Berkeley.

Ewan, W.G., and R. Krones (1974). Measuring larynx movement using the thyroumbrometer. *Journal of Phonetics* 2: 327-35.

Fant, G. (1973). *Speech Sounds and Features.* (Chapter 7: Stops in CV-Syllables). Cambridge (Massachusetts): M.I.T. Press.

Fant, G., L. Nord and P. Brenderud (1976). A note on the vocal tract wall impedance. *Speech Transmission Laboratory Quarterly Prog. Status Report* 4: 13-20.

Fischer-Jørgensen, E. (1954). Acoustic analysis of stop consonants. *Miscellanea Phonet.* 2: 42-59.

Fischer-Jørgensen, E. (1980). Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *Annual Report Inst. Phonetics, University of Copenhagen* 14: 207-61.

Fischer-Jørgensen, E. and H. Hirose (1974). A preliminary electromyographic study of labial and laryngeal muscles in Danish. *Haskins Laboratories Status Report on Speech Research* SR-39/40: 231-54.

Frøkjaer-Jensen, B. (1967). A photo-electric glottograph. *Annual Report Inst. Phonet., University of Copenhagen* 2: 5-19.

Frøkjaer-Jensen, B., C. Ludvigsen and J. Rischel (1971). A glottographic study of some Danish consonants. In L.L. Hammerich, R. Jakobson and E. Zwirner, *Form and Substance*, 123-40. Copenhagen: Akademisk Forlag.

Fujimura, O. (1977). Discussion, pp. 28-30, following M. Hirano, Structure and vibratory behavior of the vocal folds. In M. Sawashima and F.S. Cooper, *Dynamic Aspects of Speech Production*, pp. 13-27. Tokyo: University of Todyo Press.

Fujimura, O., and M. Sawashima (1971). Consonantal sequences and laryngeal control. *Annual Bulletin RILP* 5: 1-6.

Fukui, N., and H. Hirose (1983). Laryngeal adjustments in Danish voiceless obstruent production. *Annual Bulletin RILP* 17: 61-71.

Goldstein, M.C., and N. Nornang (1970). *Modern Spoken Tibetan: Lhasa Dialect.* Seattle: University of Washington Press.

Halle, M., G.W. Hughes and J.-P.A. Radley (1957). Acoustic properties of stop consonants. *Journal of the Acoustical Society of America* 29: 107-16.

Halle, M., and K.N. Stevens (1967). On the mechanism of glottal vibration for vowels and consonants. *Qu. Prog. Report, Res. Lab. Electronics* 85: 267-71. Cambridge (Massachusetts): M.I.T.

Halle, M., and K.N. Stevens (1971). A note on laryngeal features. *Qu. Prog Report, Res. Lab. Electronics* 101: 192-213. Cambridge (Massachusetts): M.I. T.

Han, M.S. (1961). Japanese Phonology: An Analysis Based upon Sound Spectrograms. Doctoral dissertation, University of Texas at Austin.

Hirano, M. (1977). Structure and vibratory behavior of the vocal folds. In M. Sawashima and F.S. Cooper, *Dynamic Aspects of Speech Production*, pp. 13-27. Tokyo: University of Tokyo Press.

Hirose, H., and T. Gay (1972). The activity of the intrinsic laryngeal muscles in voicing control. *Phonetica* 25: 140-64.

Hirose, H., C.Y. Lee and T. Ushijima (1974). Laryngeal control in Korean stop production. *Journal of Phonetics* 2: 145-52.

Hirose, H., L. Lisker and A.S. Abramson (1972). Physiological aspects of certain laryngeal features in stop production. *Haskins Laboratories Status Report on Speech Research* SR-31/32: 183-91.

Hirose, H., and T. Ushijima (1978). Laryngeal control for the voicing distinction in Japanese consonants. *Phonetica* 35: 1-10.

Hoshiko, M. (1960). Sequence of action of breathing muscles during speech. *Journal of Speech and Hearing Research* 3: 291-7.

Hoshiko, M. (1962). Electromyographic investigation of the intercostal muscles during speech. *Archives Phys. Med.* 43: 115-9.

Hutters, B. (1978). The glottal gesture in some Danish consonants: preliminary observations. *Annual Rep. Inst. Phonet., University of Copenhagen* 12: S87-101.

Hutters, B. (1985). Vocal fold adjustments in aspirated and unaspirated stops in Danish. *Phonetica* 42: 1-24.

Ishizaka, K., and M. Matsudaira (1972). Fluid mechanical considerations of vocal cord vibration. *Speech Commun.Res. Lab. Monograph No.8.*

Ishizaka, K., J.C. French and J.L. Flanagan (1975). Direct determination of vocal tract wall impedance. *IEEE Trans. Acoust., Speech Signal Process.* ASSP-23 (4): 370-73.

Iwata, R., and H. Hirose (1976). Fiberoptic acoustic studies of Mandarin stops and affricates. *Ann. Bull. RILP* 10: 47-60.

Iwata, R., M. Sawashima and H. Hirose (1981). Laryngeal adjustments for syllable-final stops in Cantonese. *Ann. Bull. RILP* 15: 45-54.

Iwata, R., M. Sawashima, H. Hirose and S. Niimi (1979). Laryngeal adjustments of Fukienese stops. *Ann. Bull. RILP* 13: 61-81.

Kagaya, R. (1971). Laryngeal gesture in Korean stop consonants. *Annual Bull. RILP* 5: 15-23.

Kagaya, R. (1974). A fiberoptic and acoustic study of Korean stops, affricates and fricatives. *Journal of Phonetics* 2: 161-80.

Kagaya, R., and H. Hirose (1975). Fiberoptic, electromyographic and acoustic analyses of Hindi stop consonants. *Ann. Bull. RILP* 9: 27-46.

Keating, P.A. (1984). Physiological effects on stop consonant voicing. *Working Papers in Phonetics* 59: 29-34. Los Angeles: University of California Phonetics Laboratory.

Kelkar, A.R. (1968). *Studies in Hindi-Urdu I, Introduction and Word Phonology.* Poona: Deccan College.

Kent, R.D., and K.L. Moll (1969). Vocal tract characteristics of the stop cognates. *Journal of the Acoustical Society of America* 46: 1549-55.

Kim, C.-W. (1965). On the autonomy of the tensity feature in stop classification. *Word* 21: 339-59.

Kim, C.-W. (1967). Cineradiographic study of Korean stops and a note on aspiration. *Qu. Prog. Report, Res. Lab. Electronics* 86: 259-72. Cambridge (Massachusetts): M.I.T.

Kim, C.-W. (1970). A theory of aspiration. *Phonetica* 21: 107-16.

Kjellin, O. (1976). A phonetic description of Tibetan; with a review of the literature. *Ann. Bull. RILP* 10: 127-43.

Kjellin, O. (1977). Observations on consonant types and "tone" in Tibetan. *Journal of Phonetics* 5: 317-38.

Kuraishi, T. (1963). *Iwanami Chugokugo Jiten* (Iwanami's Chinese Dictionary). Tokyo: Iwanami Shoten.

Ladefoged, P. (1962). Subglottal activity during speech. *Proc. 4th Int. Congress Phonet. Sci.* 73-91. The Hague: Mouton.

Ladefoged, P. (1964). Comment on 'Evaluation of methods of estimating subglottal air pressure.' *Journal of Speech and Hearing Research* 7: 291-2.

Ladefoged, P. (1967). *Three Areas of Experimental Phonetics.* London: Oxford University Press.

Ladefoged, P. (1968). Linguitic aspects of respiratory phenomena. In A. Bouhuys, *Sound Production in Man.* Annals N.Y. Acad. Sci. 155: 141-50.

Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics.* Chicago: The University of Chicago Press.

Ladefoged, P. (1973). The features of the larynx. *Journal of Phonetics* 1: 73-83.

Ladefoged, P. (1975). *A Course in Phonetics.* New York: Harcourt Brace Jovanovich, Inc.

Lindqvist, J. (1972). Laryngeal articulation studied on Swedish subjects. *Speech Trans. Lab Q. Prog. Status Rep.* 2-3: 10-27.

Lisker, L. (1970). Supraglottal air pressure in the production of English stops. *Language and Speech* 13: 215-230.

Lisker, L., and A. S. Abramson (1964). A cross-language study of voicing in initial stops: acoustic measurements. *Word* 20: 384-422.

Lisker, L., and A.S. Abramson (1972). Glottal modes in consonant distinctions. *Proc. 7th Int. Congress Phonet. Sci.* The Hague: Mouton.

Lisker, L., A.S. Abramson, F.S. Cooper and M.H. Schvey (1969). Transillumination of the larynx in running speech. *Journal of the Acoustical Society of America* 45: 1544-6.

Lisker, L., and T. Baer (1984). Laryngeal management at utterance-internal word boundary in American English. *Language and Speech* 27: 163-71.

Lisker, L., M. Sawashima, A.S. Abramson and F.S. Cooper (1970). Cinegraphic observations of the larynx during voiced and voiceless stops. *Haskins Laboratories Status Rep. Speech Res.* SR-21/22: 201-10.

Löfqvist, A. (1974). Variations in subglottal pressure during stop production. *Speech Commun. Sem.* 2: 99-104.

Löfqvist, A., and H. Yoshioka (1980). Laryngeal activity in Swedish obstruent clusters. *Journal of the Acoustical Society of America* 68: 792-801.

Löfqvist, A., and H. Yoshioka (1981). Laryngeal activity in Icelandic obstruent production. *Nordic Journal of Linguistics* 4: 1-18.

Lubker, J.F. (1973). Transglottal airflow during stop consonant production. *Journal of the Acoustical Society of America* 53: 212-15.

Lubker, J.F., B. Fritzell and J. Lindqvist (1970). Velopharyngeal function: an electromyographic study. *Speech Trans. Lab. Q. Prog. Stat. Rep.* 4: 9-20.

McGlone, R.E., and T. Shipp (1972). Comparison of subglottal air pressures associated with /p/ and /b/. *Journal of the Acoustical Society of America* 51: 664-5.

Malone, K. (1923). *The Phonology of Modern Icelandic.* Menasha (Wisconsin): The Collegiate Press, George Banta Publishing Company.

Martin, S.E. (1951). Korean phonemics. *Language* 27: 519-33.

Minifie, F.D., J.H. Abbs, A. Tarlow and M. Kwaterski (1974). EMG activity within the pharynx during speech production. *Journal of Speech and Hearing Research* 17: 497-504.

Müller, E.M., and W.S. Brown (1980). Variations in the supraglottal air pressure waveform and their articulatory interpretation. In N. Lass, *Speech and Language: Advances in Basic Research and Practice* Vol. 4, pp. 317-89. New York: Academic Press.

Munro, R.R., and C. Adams (1971). Electromyography of the intercostal muscles in connected speech. *Electromyography* 11: 365-78.

Netsell, R. (1969). Subglottal and intraoral air pressure during the intervocalic contrast of /t/ and /d/. *Phonetica* 20: 68-73.

Nihalani, P. (1974). An aerodynamic study of stops in Sindhi. *Phonetica* 29: 193-224.

Nihalani,P. (1975). Velopharyngeal opening in the formation of voiced stops in Sindhi. *Phonetica* 32: 89-102.

Niimi, S., and M. Sawashima (1974). A preliminary study on the simultaneous recording of laryngeal muscle activity and the glottal shape during speech utterance. *Ann. Bull. RILP* 8: 19-22.

Ogura, S. (1953). Glottalized stops in Korean (in Japanese). *Gengo Kenkyu* 22/23: 1-20.

Ohala, J.J. (1974). A mathematical model of speech aerodynamics. *Speech Commun. Semin* 2: 65-72.

Ohala, J.J. (1975). A model of speech aerodynamics. *8th Int. Congr. Phonet. Sci*. 1-7.

Ohala, M. (1979). Phonological features of Hindi stops. *South Asian Lang. Anal.* 1: 79-88.

Ohala, M., and J.J. Ohala (1972). The problem of aspiration in Hindi. *Ann.Bull. RILP* 6: 36-46.

Ohala, J.J., and C.J. Riordan (1979). Passive vocal tract enlargement during voiced stops. In J. Wolf and D. Klatt, *Speech Communication Papers*, pp. 89-92. New York: Acoustical Society of America.

Ohala, J.J., C.J. Riordan and H. Kawasaki (1979). Investigation of pulmonic activity in speech. *9th International Congress of Phonetic Sciences* pp. 89-95.

Perkell, J.S. (1969). *Physiology of Speech Production.* Cambridge (Massachusetts): M.I.T. Press.

Pétursson, M. (1976). Aspiration et activité glottale. Examen expérimental à partir de consonnes islandaises. *Phonetica* 33: 169-98.

Riordan, C.J. (1978). Larynx height and voicing. *96th Meet. Acoust. Soc. Am., Honolulu.*

Riordan, C.J. (1979). Vertical larynx movement during stop closure. *8th Int. Congress Phonet. Sci.*, 208.

Riordan, C.J. (1980). Larynx height during English stop consonants. *Journal of Phonetics* 3: 353-60.

Rothenberg, M. (1968). The breath-stream dynamics of simple released plosive production. *Bibl. Phonetica* 6. Basal, Karger.

Sawashima, M. (1968a). Observation of glottal movements. Speech Symposium, Kyoto, C-2-1 to C-2-7.

Sawashima, M. (1968b). Movements of the larynx in articulation of Japanese consonants. *Ann. Bull. RILP* 2: 11-20.

Sawashima, M. (1970). Glottal adjustments for English obstruents. *Haskins Lab. Status Rep. Speech Res.* SR 21/22: 187-200.

Sawashima, M., A.S. Abramson, F.S. Cooper and L. Lisker (1970). Observing laryngeal adjustments during running speech by use of a fiberoptics system. *Phonetica* 22: 193-201.

Sawashima, M., H. Hirose, S. Kiritani and O. Fujimura (1968). Articulatory movements of the larynx. *Proceedings of the Sixth Int. Congr. Acoust.* 13-1 to 13-4.

Sawashima, M., H. Hirose, T. Ushijima and S. Niimi (1975). Laryngeal control in Japanese consonants, with special reference to those in utterance initial position. *Ann. Bull. RILP* 9: 21-6.

Sawashima, M., and S. Miyazaki (1973). Glottal opening for Japanese voiceless consonants. *Ann. Bull. RILP* 7: 1-9.

Sawashima, M., and S. Niimi (1974). Laryngeal conditions in articulations of Japanese voiceless consonants. *Ann. Bull. RILP* 8: 13-17.

Sawashima, M., and H.S. Park (1979). Laryngeal adjustments for syllable final stops in Korean: some preliminary results of fiberoptic observation. *Ann. Bull. RILP* 13: 83-9.

Sawashima, M., H.S. Park, K. Honda and H. Hirose (1980). Fiberscopic study on laryngeal adjustments for syllable-final applosives in Korean. *Ann. Bull. RILP* 14: 125-38.

Slis, J.H., and P.H. Damsté (1967). Transillumination of the glottis during voiced and voiceless consonants. *IPO Ann. Prog. Rep.* 2: 103-9.

Stevens, K.N. (1977). Physics of laryngeal behavior and larynx modes. Phonetica 34: 264-79.

Trager, G.L., and H.L. Smith (1951). *An Outline of English Structure*. Norman (Oklahoma): Battenburg Press.

van den Berg, Jw. (1958). Myoelastic-aerodynamic theory of voice production. *J. Speech Hear. Res.* 1: 227-44.

Westbury, J.R. (1979). *Aspects of the Temporal Control of Voicing in Consonant Clusters in English*. Austin: Texas Linguistic Forum 14: 1-304.

Westbury, J.R. (1983). Enlargement of the supraglottal cavity and its relation to stop consonant voicing. *Journal of the Acoustical Society of America* 73: 1322-36.

Westbury, J.R., and S. Niimi (1979). An effect of phonetic environment on voicing control mechanisms during stop consonants. In J. Wolf and D.Klatt, *Speech Communication Papers*. New York: Acoustical Society of America.

Yadav, R. (1979). Maithili Phonetics and Phonology. Doctoral dissertation, University of Kansas.

Yadav, R. (1984). Voicing and aspiration in Maithili: a fiberoptic and acoustic study. *Indian Linguistics* 1-30.

Yoshioka, H., A. Löfqvist and R. Collier (1982). Laryngeal adjustments in Dutch voiceless obstruent production. *Ann. Bull. RILP* 16: 27-35.

Yoshioka, H., A. Löfqvist and H. Hirose (1981). Laryngeal adjustments in the production of consonant clusters and geminates in American English. *J. Acoust. Soc. Am.* 70: 1615-23.

In defense of the phonetic adequacy of the traditional term "voiced aspirated"

R. Prakash Dixit

Abstract

This paper defends the phonetic adequacy of the traditional term "voiced aspirated" as a descriptor and classifier of a fourth category of homorganic stops such as /bh/, which has been questioned in recent phonetic literature.

Introduction

In those languages that possess four manner categories of homorganic stops such as /p/, /ph/, /b/ and /bh/, the fourth category has been, traditionally, described and classified as "voiced aspirated". This description and classification of the fourth category of stops has long been considered adequate phonologically as well as phonetically. Recently, however, the phonetic adequacy of the term "voiced aspirated" as a descriptor and classifier of the fourth category of stops has been questioned. According to Ladefoged (1971) "when one uses a term such as voiced aspirated, one is using neither the term voiced nor the term aspirated in the same way as in the description of the other stops." That is, unlike the voiced unaspirated stops which are produced with normal closure voicing, the closure voicing during the voiced aspirated stops is not normal (Catford 1977, Ladefoged 1975). Moreover, the voiced aspirated stops are also not aspirated either, since during their production the release of the oral closure is not followed by a period of voicelessness. They are thus unlike the voiceless aspirated stops where the release of the oral closure is immediately followed by a period of voicelessness (Abercrombie 1967, Catford 1977, Ladefoged 1971,1975).

There has been some confusion in the phonetic literature as to what aspiration really is. Part of this confusion can be, perhaps, attributed to Lisker and Abramson's (1964) work on voice onset time associated with stop consonant production, although to no fault of theirs, as they did not consider voiced aspirates. Their findings on voice onset time led them to regard the "noise feature of aspiration ... simply as an automatic concomitant of a large delay in voice onset." Unfortunately, "the noise feature of aspiration," which provided the phonetic basis for the description and classification of the voiced aspirated stops as aspirated, was forgotten and the "large delay in voice onset" or "voicing lag" became the equivalent of aspiration. From then on most phoneticians and linguists used these terms in the sense of aspiration. Thus, the voiced aspirated stops were considered phonetically neither voiced nor aspirated and a few new terms such as "whispery voiced," "breathy voiced," "murmured," "murmured aspirated," and "voiced phonoaspirated" were suggested as phonetically more adequate replacements of the term "voiced aspirated."

The purpose of this paper is to examine and discuss the phonetic adequacy or inadequacy of the various terms mentioned above in the light of glottographic, aerodynamic and spectrographic data from Hindi (a four-category Indo-Aryan language) and to show that the term "voiced aspirated" is a better phonetic descriptor and classifier of the fourth category of stops than its suggested replacements.

Experimental Results

The results presented here are based on the analysis of a large body of data. Although only a few illustrations are given here, they may be taken as typical of the data as a whole. Glottographic, aerodynamic and spectrographic data from one speaker of Hindi are presented in Figures 1, 2 and 3, respectively. These figures display records obtained during the nonsense words /pipi/, /phiphi/, /bibi/ and /bhibhi/ which were produced in a frame sentence /didi___bolije/ 'elder sister ___ (please) say'. We will not consider the data on the voiceless unaspirated stop /p/ since it is neither voiced nor aspirated. Photo-Electric Glottograms (PEG) in Figure 1 show that

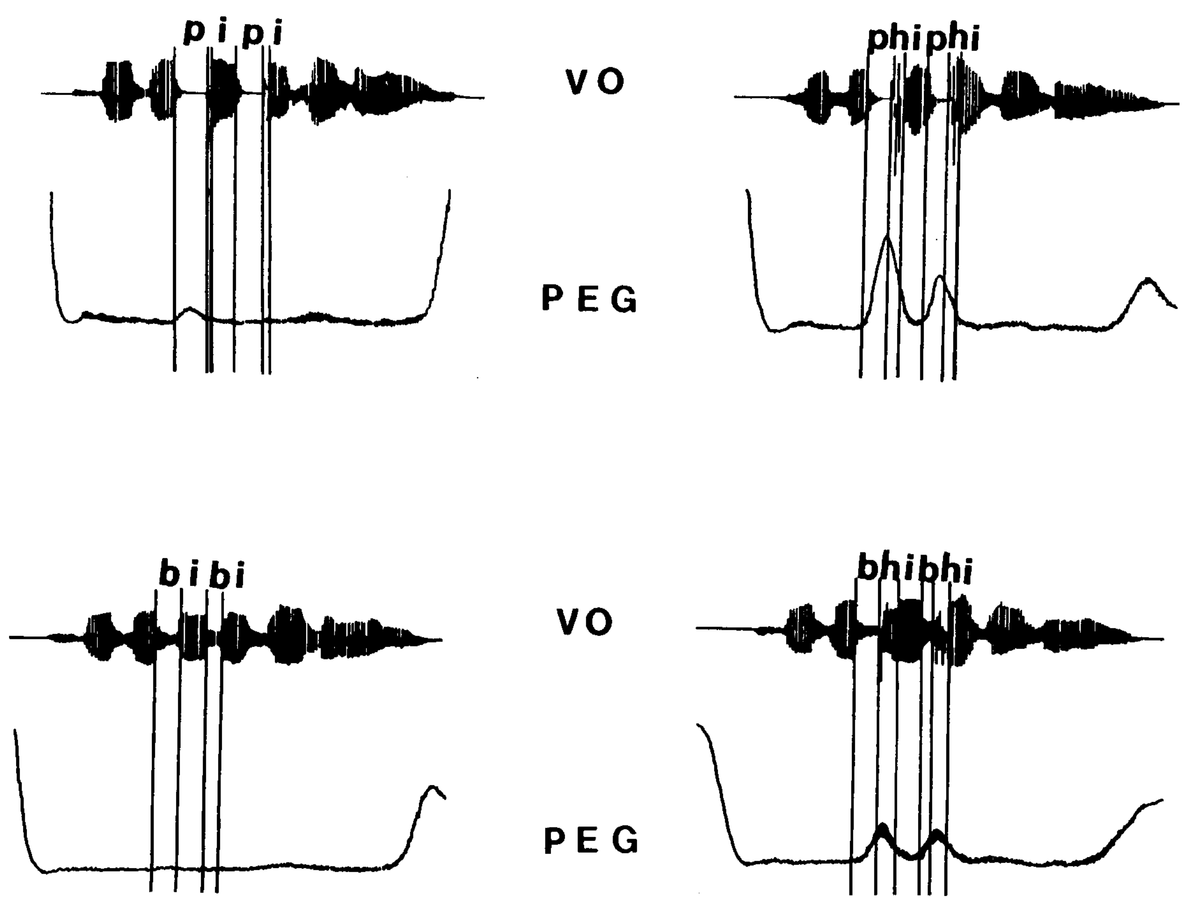


Figure 1.

/b/ is produced with an approximated glottis, while /bh/ and /ph/ are produced with a moderately and a widely open glottis, respectively. The glottal opening during /bh/ begins appreciably before the oral release, peaks around the middle of the noise interval and terminates during the initial part of the following vowel. However, during /ph/ the glottal opening starts at or slightly prior to the articulatory closure, peaks at or near the articulatory release and terminates during the early portion of the following vowel. Notice that the glottal opening during /ph/ is approximately double that during /bh/.

Oral air pressure (P_O) and oral air flow (U_O) curves in Figure 2 show that the pressure profiles and the magnitudes of pressure during the articulatory closure for /b/ and /bh/ are about the same but the magnitude of flow after the articulatory release is much greater for /bh/ than for /b/. However, the pressure profile as well as the magnitude of pressure and flow for /ph/ are different than those for /b/ and /bh/. For /ph/ the pressure rise is rapid, the pressure build up is higher and the flow rate is greater than each of these for either /b/ or /bh/.

The spectrograms in Figure 3 show that the closure interval of /ph/ is mostly voiceless, except for a few vertical striations indicating vocal fold vibration continuing from the preceding voiced environment. The closure intervals of /b/ and /bh/, on the other hand, are fully voiced. The acoustic pattens of closure voicing for /b/ and /bh/ appear to be virtually identical. Notice that a period of voicelessness occurs between the articulatory release of /ph/ and the onset of the following vowel. However, such a period during /bh/ is not voiceless; it is occupied by a fuzzy acoustic pattern of vertical striations, which appear to be quite different from the pattern observed during closure interval of /bh/ or /b/. On the other hand, during this period acoustic noise can be observed for /bh/ and /ph/ alike, in about the same frequency regions as the second, third and fourth resonances of the vowel following these stops.

Discussion

The spectrographic data presented here clearly demonstrate that the voiced aspirated stops of Hindi are both voiced and aspirated. They are voiced because they are produced with regular closure voicing like the voiced unaspirated stops and they are aspirated because, like the voiceless aspirated stops, they are produced with glottal noise--in other words aspiration--following the release of oral closure. However, the period following the release of oral closure, during which aspiration occurs, is voiceless in the voiceless aspirated stops but not in the voiced aspirated stops; in the latter category of stops the vocal folds continue to vibrate through this period. In the voiced aspirated stops the mode of vocal fold vibration during the period of aspiration is apparently different from that during the closure interval. This is reflected in the fuzzy pattern of vertical striations during the aspiration vis-a-vis the clear pattern of vertical striations during the closure interval (Figure3). The difference in the mode of vocal fold vibration within the voiced aspirated stops is further reflected in an approximated glottis and an extremely low flow rate during the closure interval versus a moderately open glottis and a high flow rate during the noise or period of aspiration (Figures 1 and 2).

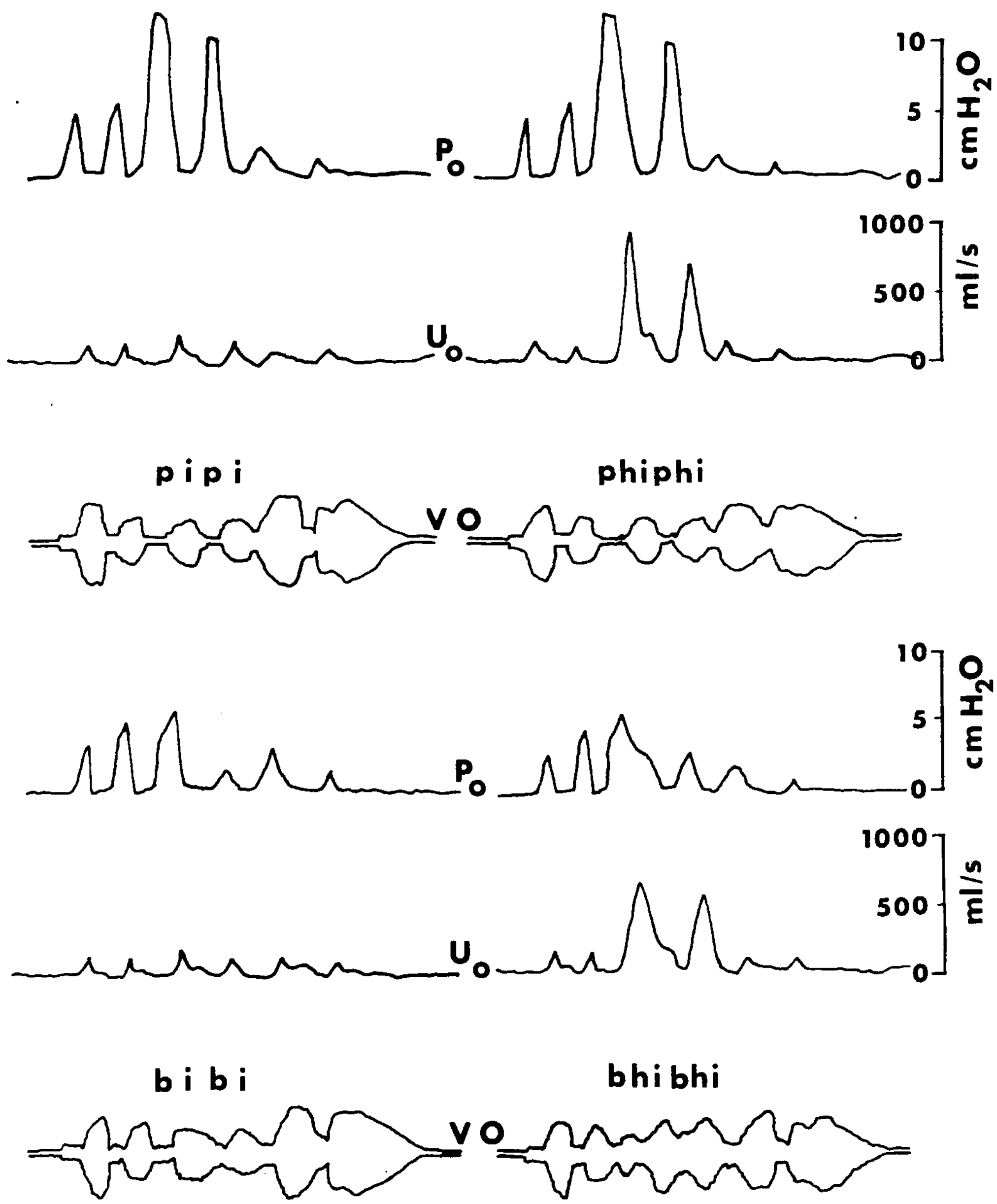


Figure 2.

During the aspiratory interval of these stops the glottis is only moderately open and the vocal folds are relatively slack; consequently they can vibrate in the absence of an articulatory obstruction to the airflow. But they do not touch one another while vibrating. On the other hand, during the aspiratory interval of the voiceless aspirated

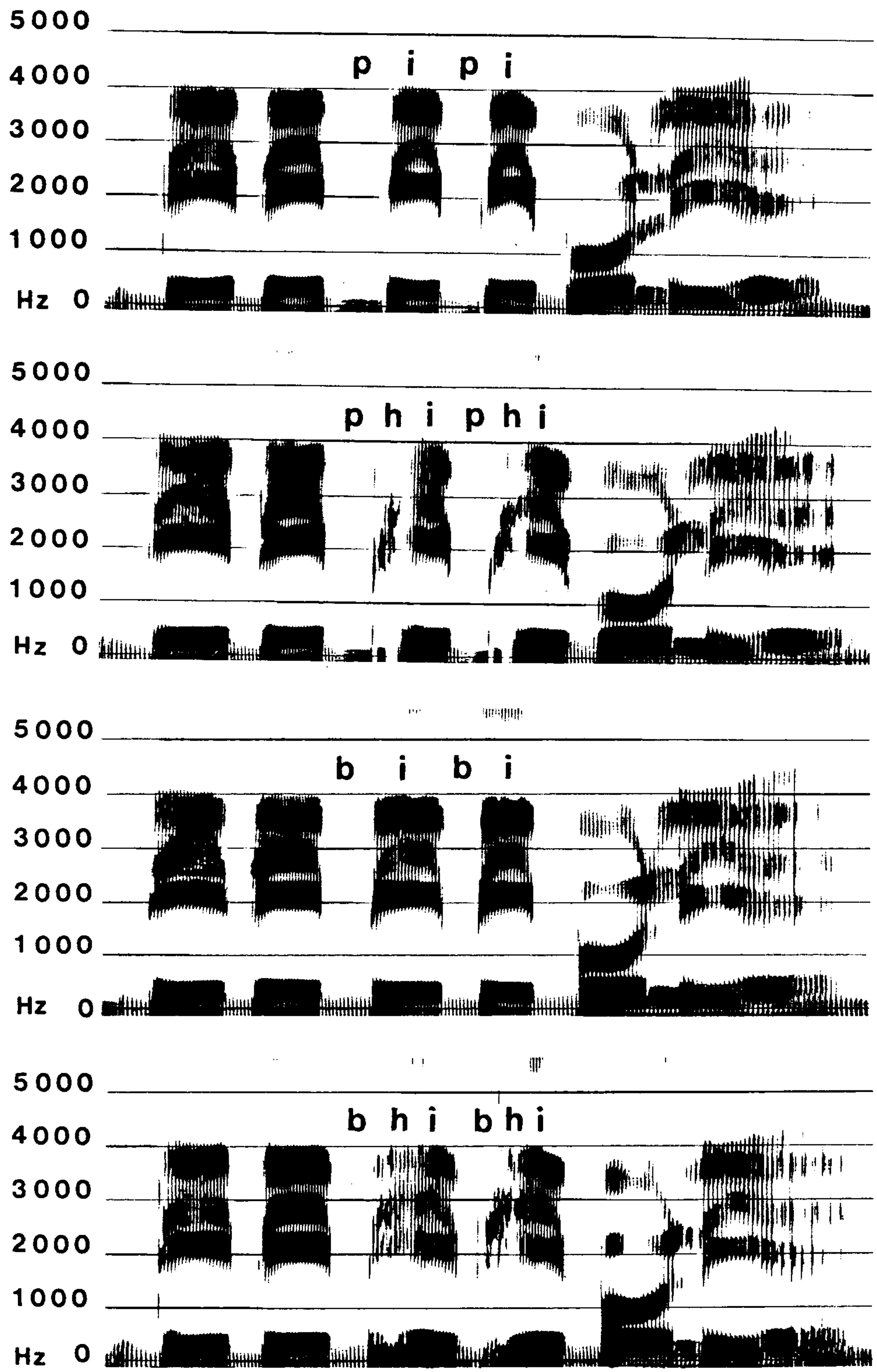


Figure 3.

stops the glottis is widely open and the vocal folds are relatively tense; they simply cannot vibrate even in the presence of high flow rate. Thus, the aspiratory interval of the voiceless stops is voiceless but that of the voiced stops is not. Further, the spectrograms in Figure 3 show that the acoustic noise during the voiceless aspirated as well as voiced aspirated stops is found in about the same frequency regions as the upper resonances of the following vowel (as pointed out earlier), which indicates that the noise source is located at the glottis. If the source were located elsewhere a different noise pattern would have resulted. Glottal noise is commonly called aspiration. Thus, the phonetic description of the voiced aspirated stops as aspirated cannot be reasonably rejected (Dixit 1983). Likewise, their phonetic description as voiced is also unquestionable. In anticipation of the forthcoming aspiratory phase the glottis begins to open appreciably before the release of the oral closure; but the vocal folds continue to vibrate and remain fairly close together almost until the articulatory release, as attested by fiberoptic observations (Benguerel and Bhatia 1980, Kagaya and Hirose 1975). Those sounds in which the vocal folds "form a closure or near closure during successive periods of the oscillation" are said to be regularly voiced (Peterson and Shoup 1966). Thus, the traditional term "voiced aspirated" is an adequate phonetic descriptor and classifier of the fourth category of homorganic stops. Further "voiced aspirated" is a better term since it produces a symmetrical matrix of classificatory terms, and is capable of capturing phonological generalizations, while its suggested replacements produce an asymmetrical and counterintuitive matrix and create problems in the description of such sound changes as Grassmann's Law, as shown by Halle (1973).

On the basis of the definitions of phonetic terms given in Peterson and Shoup (1966), Benguerel and Bhatia (1980) have proposed the term "voiced phonaspirated" as a phonetically more adequate descriptor than the term "voiced aspirated" for the fourth category of homorganic stops. There is no problem with the term "voiced" which adequately describes what happens during oral closure. However, the term "phonoaspirated" which describes what happens after the release of oral closure is problematic. It appears that "phono" in "phonoaspirated" was prefixed to "aspirated" to indicate the particular mode of vocal fold vibration which occurs during the aspiratory period and which is different from the one that occurs during the oral closure. However, "phono" is also prefixed to "constricted" in the term "phonoconstricted" where it indicates a very different mode of vocal fold vibration from the one that occurs during the aspiratory interval. As "phono" describes two entirely different modes of vocal fold vibration, it renders the term "phonoaspirated" phonetically inadequate, and thus unacceptable.

The discussion of phonation types in Catford (1977) suggests the term "whispery voiced" for the voiced aspirated stops. He says that "the fact is that in such sounds as [bh] there is whispery voice rather than voice during the stop and for a certain period after its release." The glottal stricture used in the production of whispery voice is described by Catford as "narrowed vibrating." The degree of opening according to him is less than 25% of maximal glottal opening, while during voiceless stops it is from 60 to 95% of maximal glottal opening. Let us assume that the opening for the voiceless aspirated stops is 95% of maximal glottal opening. Now recall that the degree of glottal opening for the voiced aspirated stops in the data presented here was about half of that for the voiceless aspirated stops. Thus, the voiced aspirated stops were produced with more than 45% rather than less than 25% of maximal glottal opening. Obviously, they were not produced with whispery voice, since the glottal stricture was inappropriate for the generation of whispery voice, being twice as wide as required for such a phonation. Moreover, Catford's assumption that the glottis is in the phonatory posture for whispery voice during the stop phase is also not borne out by the data. But even if such a posture were present during the stop phase it could not generate whispery voice in the presence of a supraglottal obstruction. Thus, the term "whispery voiced" instead of "voiced aspirated" is phonetically inappropriate.

Ladefoged has suggested the terms "murmured" (1971,1973,1975) and "murmured aspirated" (Ladefoged, Williamson, Elugbe and Uwalaka 1976) in place of the term "voiced aspirated" for the phonetic description of the fourth category of homorganic stops. According to him (1975) "murmured sounds are sometimes made ... with the glottis fairly open at one end. They can also be made with a narrower opening extending over nearly the whole length of the vocal cords." Ladefoged has called both these physiological possibilities the "murmur" state of the glottis and has assumed that such a state occurs during the oral closure as well as after the release of the closure in the so-called murmured or murmured aspirated stops (1973, 1976). Thus the vocal fold vibrations that occur during the oral closure are assumed to be of "the kind that would be expected from a small volume of air flowing through the glottis while it is in the position for a murmured sound" (1973); that is the vocal fold vibrations during articulatory closure are said to be different from those in normal voice vibrations. However, these assumptions do not find support in the glottographic and acoustic data presented here. The phonation that is generated after the release of a closure was earlier (1971,1973,1975) described as murmur or breathy voice. Later (1976) in relation to the somewhat different voiced aspirated stops of Owerri Igbo it was surprisingly described as aspiration, (surprisingly since "aspiration" for Ladefoged is "a period of voicelessness during and immediately after the release of an articulatory stricture" (1971)). This was the result of redefining aspiration in an attempt to accommodate the voiced aspirated and the unvoiced aspirated stops under the same phonetic category of "aspiration". The attempt, however, did not succeed since the closure voicing in the voiced aspirated stops of Owerri Igbo was still considered to be murmur, which is almost certainly contrafactual.

Lately, Ladefoged has changed his position. In the second edition of his book *A Course in Phonetics* he states that "voicing during the vowel and the closure are, as usual, the result of air flowing between the vocal cords while they are held loosely, fairly close together". In other words the closure voicing in the voiced aspirated stops is normal regular voicing. This is strongly supported by the glottographic and acoustic data presented here. Further, the vibrations that occur after the articulatory release of these stops are described as "murmured (breathy) vibrations". That is, after the articulatory release of the voiced aspirated stops "murmur" or "breathy voice" occurs. Breathy voice may as well be called "voicy aspiration".

As we have seen, in the voiced aspirated stops of Hindi aspiration is accompanied by glottal vibration which noticeably changes its quality. It sounds more like breathy voice. However, it would be inappropriate to call the voiced aspirated stops "breathy voiced," because this term describes only the state of the glottis after the release of the closure in these stops. On the other hand, the term "voiced breathy voiced" sounds strange. If the term "murmur" could be strictly limited to the meaning of "breathy voice," then perhaps a term like "voiced murmured" could be suggested to describe the fourth category of homorganic stops. The seeds of this term were already present in Lisker and Abramson's (1964) and Benguerel and Bhatia's (1980) work, but for some reason they did not suggest it. If this term is accepted then a diaeresis [..] should not be used under the stop part of the consonant, since it will give the wrong impression that the closure voicing is murmur rather than regular voicing.

To conclude, the term "voiced murmured," although phonetically adequate, will produce an asymmetrical matrix of classification that will fail to capture phonological generalizations. However, it may turn out to be a useful term in speech synthesis. On the other hand, the term "voiced aspirated" is not only phonetically adequate but also produces a symmetrical matrix of classificatory values and is capable of capturing phonological generalizations. It will thus be more attractive to the linguist. The other terms which really are both phonetically and phonologically inadequate should be discarded.

References

Abercrombie, D. (1967). *Elements of General Phonetics*. Edinburgh: Edinburgh

University Press.

Benguerel, A-P., and T.K. Bhatia (1980). Hindi stop consonants: an acoustic and fiberscopic study. *Phonetica* 37: 134-48.

Catford, J.C. (1980). *Fundamental Problems in Phonetics.* Bloomington: Indiana University Press.

Dixit, R.P. (1983). On defining aspiration. *Proceedings of the XIIIth International Congress of Linguists, CIPL, Tokyo.*

Halle, M. (1973). Review of P. Ladefoged, *Preliminaries to Linguistic Phonetics.* Language 49: 926-33.

Kagaya, R. and H. Hirose (1975). Fiberoptic electromyographic and acoustic analyses of Hindi stop consonants. *Annual Bulletin RILP* 9: 27-46.

Ladefoged, P. (1971). *Preliminaries to Linguistic Phonetics.* Chicago: University of Chicago Press.

______ (1973). The features of the larynx. *Journal of Phonetics,* 1: 73-83.

______ (1975). *A Course inPhonetics.* New York: Harcourt Brace Jovanovich.

______, K. Williamson, B. Elugbe and A.A. Uwalaka (1976). Stops of Owerri Igbo. *Studies in African Linguistics,* Supplement 6: 147-63.

Lisker, L. and A. Abramson (1964). A cross-language study of voicing in initial stops: acoustic measurements. *Word* 20: 384-422.

Peterson, G.E. and J.E. Shoup (1966). A physiological theory of phonetics. *Journal of Speech and Hearing Research* 9: 5-68.

The effect on F_0 of the linguistic use of phonation type

Ian Maddieson and Susan Hess

Paper to be presented at the Eleventh International Congress of Phonetic Sciences, Tallinn (Estonia, USSR), August 1987

Abstract

Phoneticians generally expect that laxer adjustments of the vocal cords will produce lower F_0. Hence, languages with phonological contrasts between syllables with tense (somewhat creaky) and lax (somewhat breathy) phonation would be expected to show a difference in pitch between them. We measured F_0 in several minority languages of China with contrasts that have been described as tense vs lax. Our results show that a pitch difference is only sometimes present. The patterns are, in part, explicable in terms of different phonetic realizations and different diachronic sources of the tense/lax contrast, and in terms of its phonological function.

A tendency for different phonatory settings to be associated with pitch differences has been noted by many observers. For example, Laver (1980), in his discussion of laryngeal tension settings, remarks that "there is a strong possibility that in tense voice the pitch range will be higher than in lax voice". Later he comments that "lax voice tends to be accompanied by a low pitch-range". But he goes on to note that there is nothing necessary about the association of laryngeal tension with pitch, commenting that "it is certainly possible to compensate for these tendencies."

Laver is discussing tense and lax laryngeal settings as attributes of individual voice quality. However, a number of languages use tense and lax phonation for linguistic contrast between vowels. This phenomenon is quite common among languages spoken in Southwestern China and adjoining parts of Southeast Asia. We have been conducting studies of the phonation type contrast in several of these languages, and have reported some of our results elsewhere (Maddieson & Ladefoged 1985, Maddieson &

Hess 1986). In the present paper we focus on the relation between F_0 and phonatory tension in five of the languages in question. We hypothesized that pitch would correlate with tension, following the tendency noted by Laver, in languages which did not also have tonal contrasts. In languages with tonal contrasts with a high functional load and phonological systems in which phonatory tension is not an aspect of particular tones, we anticipated that the need to maintain the separation of tonal registers would inhibit this tendency. Instead, speakers would draw on the compensatory mechanisms available to counteract it.

Our data consists of measurements of F_0 from 5 languages - Wa, Jingpho, Yi, Lahu, and Lisu. Wa is a non-tonal language of the Mon-Khmer family (Diffloth 1980, Qiu, Li & Nie 1980). The others are Sino-Tibetan languages with tonal systems with a high functional load. Yi (Liangshang dialect, Li & Ma 1983) and Jingpho (of Yunnan, Lu 1984) have similar tone systems, distinguishing high, mid and low-falling tones. In these two languages the phonatory contrast is independent of the tone system, although it is limited to particular syllabic nuclei in Yi. Lisu is usually analyzed as having a 6-tone system in which tense phonation is characteristic of two of the tones (Mu & Duan 1983). These two tones are mid-level and mid-falling, and can be matched with two of the "lax" tones, also mid-level and mid-falling. Mu and Duan transcribe the pitch height of the tense tones as 44 and 42, and the paired lax tones as 33 and 31, implying that the "tense" tones are indeed higher. Lahu has a system of seven tones, two of which are variously described as being checked by a glottal stop (Matisoff 1973), or having tense vowels (Ma 1984). These two tones, high-falling and low-falling, can be matched with two of the tones that occur non-checked or lax. Whereas Matisoff gives the same pitch values for tense and lax tones, Ma transcribes the tense tones as 54 and 21 but the lax tones as 53 and 31, indicating a smaller pitch range for the tense ones.

3 speakers of each language were recorded with the assistance of Ren Hongmo. The speakers read a wordlist containing 8-10 pairs of monosyllabic words with a minimal tense/lax contrast. Each list was read twice, giving 48-60 examples of each contrast (except for Lahu where only one repetition was recorded, giving 30 cases). F_0 was measured at the onset and offset of the vowel from narrow-band spectrograms. If a more

extreme value of F_0 occurred after the vowel onset that value was also measured.

The F_0 measurements in each language were examined in a 3-way analysis of variance, specifying speaker, word pair and tension as main effects. In Table 1 the mean onset and offset F_0 values are shown for the tense and lax vowels in each language. Significant differences (at the .0001 level) are printed **bold**. All other tense/lax differences are not significant (fall below the .05 level). Measurements of the peak F_0 value did not show a different pattern from those made at the onset, hence these measures are not reported.

Table 1. F_0 measures (in Hz) on tense and lax vowels.

	Wa		Jingpho	
	onset	offset	onset	offset
"tense"	146	112	**157**	128
"lax"	145	115	**145**	126

	Lisu		Lahu	
	onset	offset	onset	offset
"tense"	147	119	213	**195**
"lax"	148	122	214	**126**

	Yi	
	onset	offset
"tense"	**157**	153
"lax"	**152**	154

In Wa, words in citation form are spoken with a falling intonation. No pitch difference between tense and lax vowels was observed at either the onset or offset of the vowel. On the other hand, in Jingpho, a significant pitch difference at the vowel onset was observed. The Jingpho wordlist includes pairs of words with all three tones, but pairs with low-falling tone predominate (6 out of 10). Because of this, the mean offset value is low. The word pairs examined in Yi were all mid-level tone, hence onset and offset values are close. The onset F_0 differs between tense and lax syllables by a small but highly significant amount in Yi. In Lisu there is no significant difference at either onset or offset, despite Mu & Duan's indication to the contrary. Since phonatory tension is a property of particular tones in this language we had expected no effort to avoid a pitch distinction. Lahu shows a significant difference in F_0 at the vowel offset. The mean offset value in the two lax falling tones is considerably lower than in the tense tones.

Our results are thus generally counter to our hypothesis, which predicted that an F_0 difference would occur in the nontonal language Wa, and in Lisu and Lahu where phonation type is an aspect of tone, but not in Jingpho or Yi where phonation type is independent of tone.

Are there historical or synchronic facts about these particular languages which enable us to understand this result? Jingpho and Wa share a similar historical origin for the tense/lax contrast, namely, the somewhat breathy lax syllables are those which used to have initial voiced consonants. However, synchronically, the phonation type contrast is more salient in Jingpho than it is in Wa. We have used the difference in amplitude between the second harmonic and the fundamental, $H_2 - F_0$, as our measure of phonation type. This measure has a higher value for tenser phonation than for laxer phonation (Maddieson & Ladefoged 1985). In Wa the mean difference in the $H_2 - F_0$ measure between lax and tense vowels is just under 2 dB, whereas in Jingpho it is just over 7 dB. In addition, the tense/lax contrast in Wa is accompanied by some vowel quality difference: tense vowels have a higher first formant than lax ones, i.e. they are lower in the perceptual vowel space. In Jingpho, vowels in tense and lax syllables do not differ. It may therefore be the case that in Wa the small pitch difference that might have been expected from the not-very-salient phonation type contrast is counteracted by the effect of vowel lowering in tense syllables. In Jingpho on the other hand, the phonation type contrast is made salient enough so that the conditioning environment for any allotonic variation can be readily recognized.

Lisu developed tense phonation in syllables which were originally checked (i.e. stop-final). In Lisu we found that the mean difference in the $H_2 - F_0$ measure between tense and lax was about 3 dB, confirming the existence of a moderately salient phonation type difference. Since there is no pitch difference, this suggests that the system should be reinterpreted as one with four tones in which a phonation type contrast operates within two of the tones, rather than as one with six tones, two of which have a marked phonation type.

Lahu shows no reliable evidence of a phonation type difference based on the measure we have used, nor is there usually any auditory impression of one. Instead, in the historically checked syllables, a final glottal stop usually occurs and the vowel is considerably abbreviated (about 275ms shorter than in "lax" syllables). The much lower offset pitch in the two falling "lax" tones seems simply due to their much greater length; the pitch continues to fall and reaches a much lower level. In Lahu, phonation type is only marginally involved in syllabic contrasts. Duration, extent of pitch change and glottal stop are more central to the contrast which has been described as "tense" vs "lax". Matisoff's representation of the "tense" syllables as having a final glottal stop is more accurate than Ma's account, though Ma correctly indicates the greater pitch range of the "lax" (unchecked) syllables (cf Hombert 1983).

Yi is again somewhat different. Although the difference between "tense" and "lax" vowel pairs is quite distinctive, with an auditorily "harsher" quality for the tense members, the $H_2 - F_0$ measure does not distinguish them. Perhaps this measurement is simply not appropriate for detecting phonatory differences in the rather unusual range of "fricative vowel" phonetic segments found in Yi. We think that it is more likely that the tense/lax contrast is produced in a different way here. We speculate that the "tense" vowels employ a supralaryngeal mechanism like that used in the "strident" vowels found in some of the Khoisan languages, which involves a narrowing between the base of the epiglottis and the upper part of the arytenoid cartilages. The use of this mechanism in !Xóõ has been described in some detail by Traill (1985). Traill has listened to our Yi recordings and agrees that there is an auditory similarity between the strident vowels of !Xóõ and the tense vowels of Yi. However, in !Xóõ, strident vowels have somewhat lowered pitch, rather than the slightly higher pitch found in Yi "tense" vowels.

In the meantime, we find that, particularly in the data from Jingpho, we have provided a phonetic basis for a different hypothesis. This is the diachronic hypothesis that tonogenesis and splitting of tones in tone languages can arise from phonation type contrasts on vowels, as has been proposed by Pulleyblank (1978, 1984) for Chinese. Previous work has concentrated on consonantal sources for tones, and the effect of contrasts on vowels has largely been ignored. We now see that such effects can be significant. However, as data from Wa and Lisu demonstrate, phonation type may be contrastive in vowels without any accompanying pitch differences.

References

Diffloth, Gerard. 1980. The Wa Languages (Linguistics of the Tibeto-Burman Area 5.2). California State University, Fresno.

Hombert, Jean-Marie. 1983. A brief encounter with Lahu tones. Linguistics of the Tibeto-Burman Area 7.2: 109-111.

Laver, John. 1980. The Phonetic Description of Voice Quality. Cambridge University Press, Cambridge.

Li Min and Ma Ming. 1983. Liangshan Yiyu yuyin gailun [Description of the sounds of the Liangshang Yi language]. Sichuan Minzu Chubanshe, Chengdu.

Liu Lu. 1964. Jingpoyu gaikuang [Brief description of the Jingpho language]. Zhongguo Yuwen 132: 408-417.

Ma Shice. 1984. Lahuyu gaikuang [Brief description of the Lahu language]. Minzu Yuwen 1984.3: 70-80.

Maddieson, Ian and Peter Ladefoged. 1985. "Tense" and "lax" in four minority languages of China. Journal of Phonetics 13: 433-454.

Maddieson, Ian and Susan Hess. 1986. "Tense" and "lax" revisited: more on phonation type and pitch in minority languages of China. UCLA Working Papers in Phonetics 63: 103-109.

Matisoff, James. A. 1973. A Grammar of Lahu. University of California Press, Berkeley & Los Angeles.

Mu Yuzang & Duan Ling. 1983. Lisuyu gaikuang [Brief description of Lisu language]. Minzu Yuwen 1983.4: 72-80.

Pulleyblank, E. G. 1978. The nature of the Middle Chinese tones and their development in Early Mandarin. Journal of Chinese Linguistics 6: 173-203.

Pulleyblank, E. G. 1984. Middle Chinese: a Study in Historical Phonology. University of British Columbia Press, Vancouver.

Traill, Anthony. 1985. Phonetic and Phonological Studies of !Xóõ Bushman (Quellen zur Khoisan-Forschung 1). Helmut Buske, Hamburg.

Qiu Efeng, Li Daoyong & Nie Xizhen. 1980. Wayu gaikuang [Brief description of the Wa language]. Minzu Yuwen 1980.1: 58-69.

Characteristics of the voice source

Peter Ladefoged, Ian Maddieson, Michel Jackson, Marie Huffman.

[To appear in the Proceedings of the European Conference on Speech Technology, Edinburgh, 2-4 September 1987]

Abstract

A well defined voice source is essential for synthesizing different varieties of natural sounding speech. Some insight into the most appropriate specification of this source can be found by analysing phonation types in different languages. Using inverse filtering it is possible to recover the corresponding glottal pulse shapes. Virtually the only reliable measure that can be used to distinguish the waveforms of the different phonation types is the duty cycle of the glottal pulse. Measurements of the rising and falling slopes of the glottal pulse are less successful in this respect. Using spectral analyses in addition to glottal waveform measurements shows that the amplitude of the fundamental is largely independent of the spectral shape defined by the remaining harmonics. These findings suggest that it is appropriate to specify variations in voice quality by variations in the duty cycle and the amplitude of an independent low frequency resonance.

Introduction

Good speech synthesis requires a natural sounding voice source. Furthermore if we want to be able to synthesize a wide range of voice qualities, then we must be able to make systematic variations in the voice source. But it is by no means clear how we should specify the source in order to be able to make these systematic variations. We approached this problem by trying to find out what we needed to do in order to be able to characterize a wide range of voice qualities.

Before we could determine the acoustic correlates of different types of voice quality we had to have a set of exemplars of each type. In building a data set of this kind, a basic problem is to decide what counts as an exemplar of a particular voice quality. Labels such as "hoarse voice" and "rough voice" are notoriously difficult to apply consistently. Observers who have been trained in a particular tradition have learned to apply a particular classificatory system in a uniform manner (Laver 1980); but (like the cardinal vowel system that had to be used to describe vowel qualities before more quantifiable techniques became available), such voice quality systems cannot be learned entirely from written descriptions; they have to involve discussion among observers learning the system. Accordingly, we were reluctant to rely simply on the judgments of skilled observers who, however well trained, have no independent criteria to back up their judgments. Instead we decided to use samples garnered from languages in which differences in phonation type cause words to have different meanings.

In the course of other research (Ladefoged, Maddieson and Jackson, in press; Huffman 1987) we have investigated a large number of languages that use contrasts among phonation types. There are several languages that use, for example, a regular, modal, phonation in one word, and a more breathy or creaky phonation in another word. Thus Gujerati, an Indo-Aryan language spoken in India, has a word that can be transcribed phonetically as [bar] meaning 'twelve' and another word [ba̤r̤] meaning 'outside,' where the diareses (two dots) under the symbols in the phonetic transcription indicate that there is a breathy voice quality in otherwise virtually identical segments to those in the word meaning 'twelve'. As another example, Jalapa Mazatec, an American Indian (Otomanguean) language spoken in Mexico contrasts [ndæ̰] 'arse,' in which there is a creaky voice in the vowel, with [ndæ̤] 'horse,' which has a breathy voiced vowel.

There are two distinct ways in which one can choose to characterize these different voice sources: the specification can be in the temporal (waveform) domain, or in the frequency (spectral) domain. The two specifications are, of course, completely convertible one into the other; but both when synthesizing speech and when analyzing differences among utterances, it is sometimes convenient to state differences between sounds in the one way, and at other times in the other. We obtained data on the glottal pulse largely by inverse filtering the speech waveform to remove the resonances associated with the vocal tract transfer function (Javkin, Antoñanzas-Barroso and Maddieson 1987, Ladefoged, Maddieson and Jackson, in press). The spectral data used for the conclusions reported in this paper were obtained through further processing the inverse filtered glottal waveform by standard computer techniques, including FFT and chirp z transforms.

Figure 1 is an example of one kind of inverse filtered data. The Burmese words shown differ in that the one on the left has what is called a "smooth" tone, whereas the one on the right has a "creaky" tone, in which the larynx becomes increasingly tense. The first point to note is that there is a significant difference in the duration of the closed portion of the pulse. The creaky sample has a longer mean closed portion.

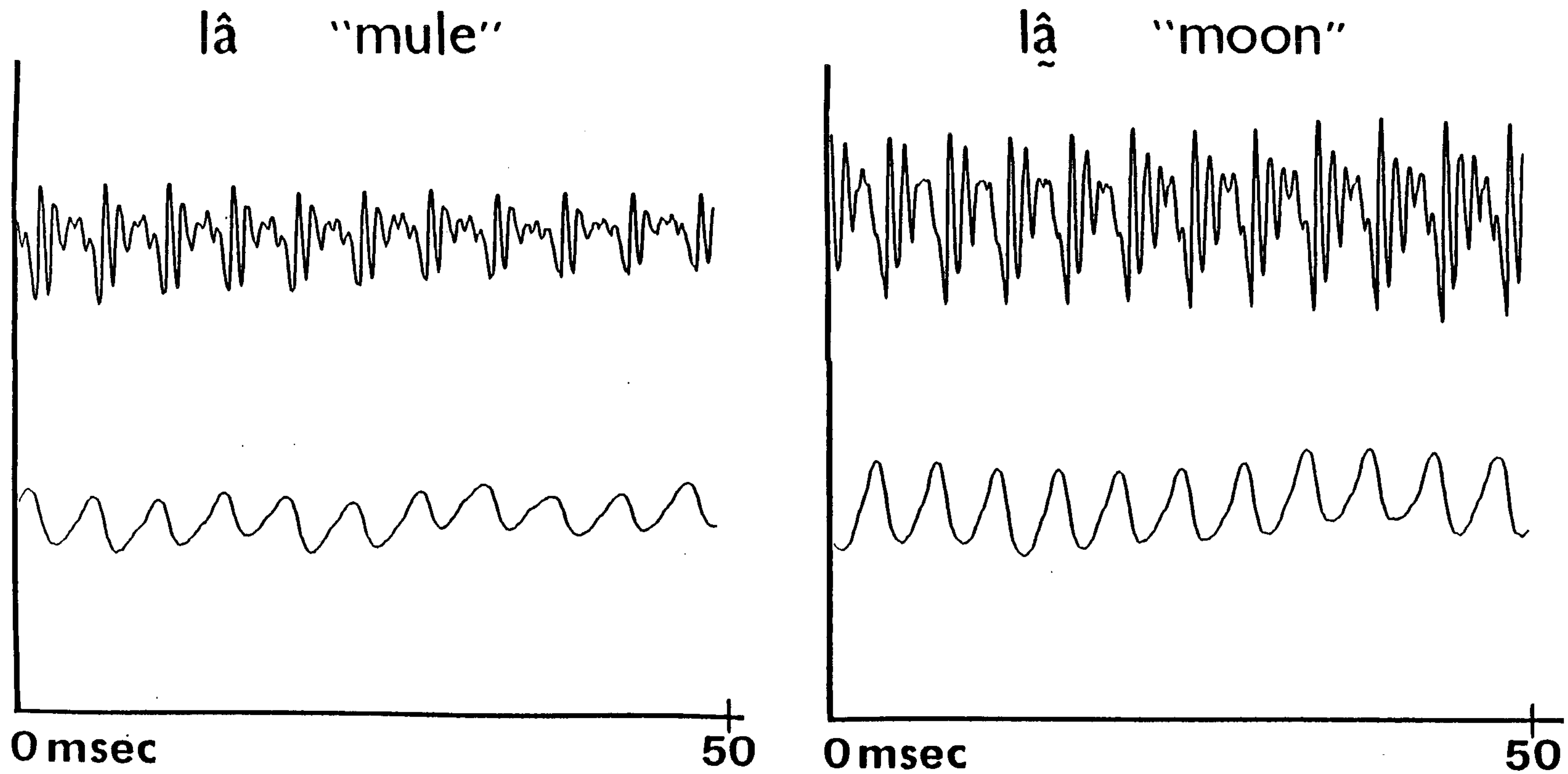


Figure 1. The original waveform and the integrated inverse filtered waveform of "smooth" and "creaky" tone words in Burmese.

Huffman (1987) found similar results for the modal versus breathy phonation contrasts in Hmong. The next point concerns the rising and falling slopes of the glottal pulses. There is no difference in these particular tokens between the smooth and creaky phonations; but this may simply be due to the overall greater amplitude of the "creaky" sample, which is an incidental feature of the particular sample being compared; it is not an inherent aspect of the linguistic contrast. We can control for irrelevant amplitude and fundamental frequency differences by calculating ratios and proportions, such as the ratio of the rising and falling slopes or the proportion of each cycle occupied by the closed phase in each of two samples, and then comparing these. In our sample analysis we found a significant difference ($p < .002$) in pulse symmetry, calculated as the derivative at the onset of the rising portion over the derivative at the offset of the falling portion. This result indicates that the vocal cords are closing more abruptly in the "creaky" tone sample in a way that is not simply related to the amplitude of the pulse.

We have not, so far, been able to make a greal deal of use of differences in glottal pulse shape for characterizing phonation types in different languages. This is partly because recovering glottal pulse shapes is a very demanding procedure. Javkin et al (1987) report that the glottal pulse shapes that they observed were affected by the low frequency noise associated with the air conditioning in the building; and Huffman (1987), working with flow recordings, found that the glottal pulse wave shape was very much affected by the small judgment calls necessary for removing the resonances associated with the formants. The difficulty in getting good comparative measures is also partly because aspects of the shape such as the ratio of the rising slope (the opening phase) to the falling slope (the closing phase) is affected by the fundamental frequencies and overall intensities of the sounds being compared. In general we have found that the only robust measurement of the waveform that can be reliably used to differentiate voice qualities in our linguistic data is the duty cycle, the ratio of the duration of the open period to the period as a whole.

We will now consider differences among phonation types in terms of the shape of the glottal spectrum. In comparison with measurements of the glottal wave shape, measures of spectral shape are less influenced by differences in fundamental frequency and amplitude, and they are less seriously affected by small errors in removing the formants. From an analytic point of view, there are considerable practical advantages in concentrating on measurements in the spectral domain. As long as we are concerned only with spectral information we do not have to worry about possible phase distortion. The inverse filtering procedure can be carried out using either FM recordings which preserve phase information, or regular AM recordings which do not.

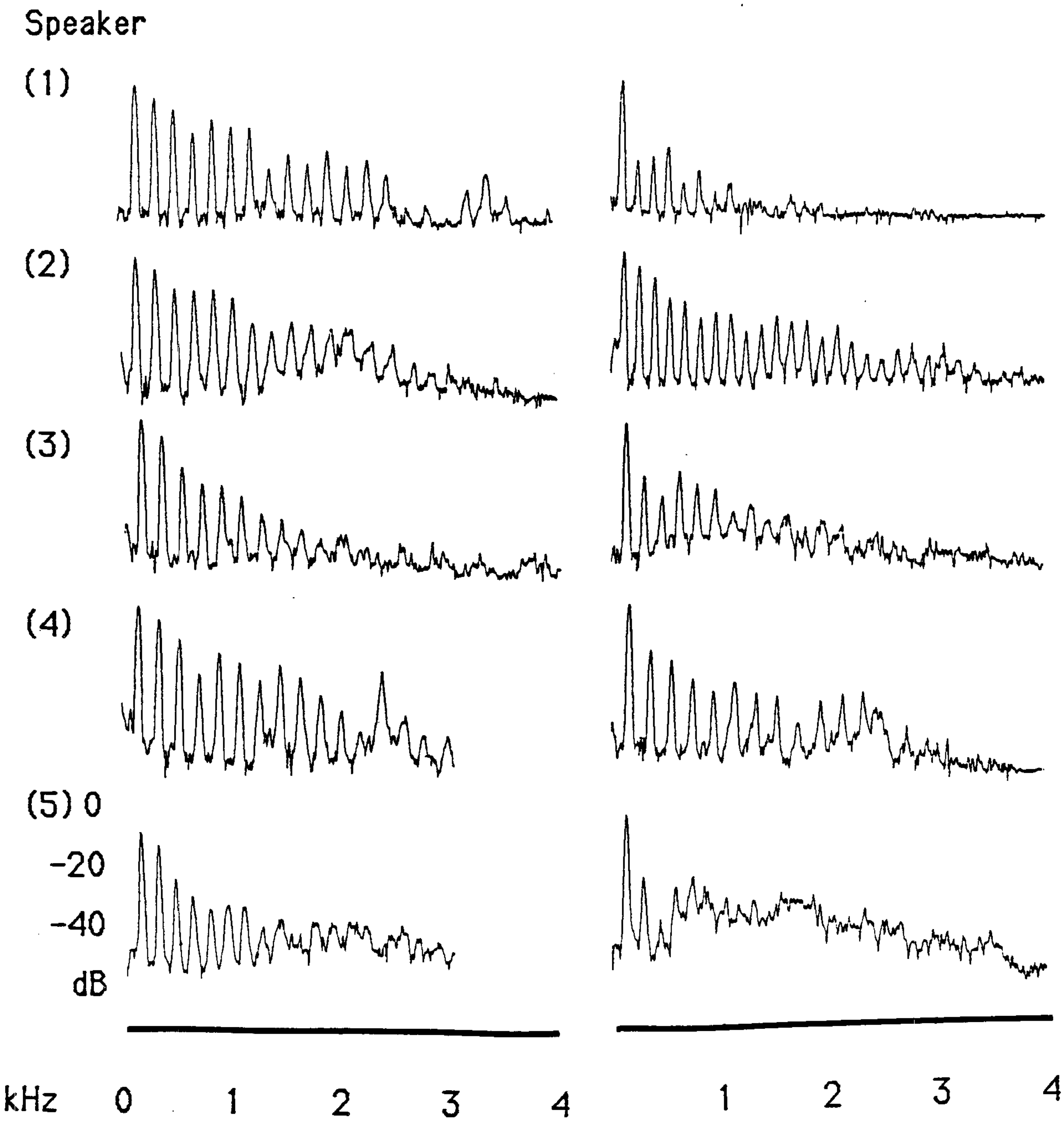


Figure 2. Spectra of inverse filtered !Xóõ modal and breathy vowels.

We have examined glottal spectra obtained through making FFTs and chirp z transforms of inverse filtered waveforms in a number of languages. We expected that the best way of describing differences among these glottal spectra would be in terms of the spectral slope. Determining the glottral spectral slope, however, is not as straightforward a task as one might assume. Figure 2 shows the glottal spectra during the modal and breathy vowels [a] and [a̤] as produced by five speakers of !Xõó, a Khoisan language spoken by Bushmen in the Kalahari desert. (We know that there is not likely to be a high demand for speech synthesis equipment among the !Xõó; but this should not deter us from using their speech to provide controlled, readily repeatable examples of types of phonation that are very similar to those that characterize different individuals in our own society.) Visual inspection of these glottal spectra is sufficient to indicate that even the modal voice sources of these five speakers of !Xõó do not have simple, smoothly falling, spectral slopes.

We measured spectral slopes by attempting to make a linear regression through the amplitudes of the first eight harmonics (H1-H8). Because of the idiosyncratic behavior of the fundamental, we also fitted slopes through these same harmonics, but omitting the fundamental (H2-H8). Linear regressions were calculated through the peaks of the harmonics in two ways, with the frequencies of the harmonics measured on a linear scale (dB/kHz) and with them measured on a log scale (dB/octave). We thus have four ways of fitting a regression line to the peaks of the harmonics in different spectra: using a linear scale between harmonics, using a log scale, and each of these using the first 8 harmonics and using only harmonics 2 through 8. Even with these four different possibilities, it is often impossible to determine the glottal spectral slope. In an analysis of a considerable body of data we found that we could determine a probably significant ($p<.05$) regression line through the peaks of the first eight harmonics 78% of the time on a linear scale, and 42% of the time on a log scale. Considering only the second through the eighth harmonics we could determine the slope 62% of the time on a linear scale, and 67% of the time on a log scale. Our general impression of these data, leads us to suspect that the energy in the low frequency components might best be modeled independently of that of the overall spectral slope.

We must now consider how these results from analyzing different phonation types could be applied to work in speech synthesis. It would seem that the most expedient thing to do is to have a pre-defined glottal pulse shape, but to be able to vary the duty cycle of this pulse. In addition, a low frequency resonance in the vicinity of F0 could be used to modify the overall spectral slope associated with the chosen pulse shape.

Interestingly, this is exactly the pair of parameters that are used in the synthesizer described by Holmes (1985). In a future project we hope to be able to discover the extent to which such synthesizers can produce the wide range of phonation types that we have found in the world's languages.

REFERENCES

Huffman, Marie (1987). "Measurements of phonation type in Hmong." *Journal of the Acoustical Society of America*. 81.2, 495-504.

Holmes, John (1985). "A parallel-formant synthesizer for machine output." in *Computer Speech Processing* (eds Frank Fallside and William A. Woods). Englewood Cliffs, N.J: Prentice Hall, 163-188.

Javkin, Hector, Norma Antoñanzas-Barroso and Ian Maddieson. (1987). "Digital inverse filtering for lingiuistic research." *Journal of Speech and Hearing Research* 30, 122-9.

Ladefoged, Peter, Ian Maddieson, and Michel Jackson (in press). "Investigating phonation types in different languages." *Proceedings of the 5th Conference on Vocal Fold Physiology* (ed. Osamu Fujimura) [to appear].

Laver, John (1987). *The Phonetic Description of Voice Quality.* Cambridge: Cambridge Univ Press.

Investigating phonation types in different languages

Peter Ladefoged, Ian Maddieson and Michel Jackson

[Abridged version of a paper presented at the Fifth International Conference on Vocal Cord Physiology, Tokyo, January 1987. The full version of this paper, which will be published in the Proceedings of this conference, contains additional material that has appeared in previous issues of *Working Papers in Phonetics*, as shown in Table 1.]

Abstract

The UCLA Phonetics Lab has been investigating different phonation types used in a wide variety of languages. Recordings of many languages spoken in South East Asia, Southern Africa, America and elsewhere have been used to demonstrate that differences in the meanings of words may depend simply on the mode of vibration of the vocal cords. The research has been aimed at both determining reliable instrumental procedures for characterizing the different phonation types, and describing the contrasts within and between languages. The techniques used include aerodynamic measures of the pressures and airflow involved; measures of the waveform designed to quantify the jitter, shimmer, and degree of noisiness present; spectrographic analyses using both commercially available instrumentation, such as the Kay Elemetrics Sona-Graph, and computer FFT, LPC, Chirp Z transforms; and inverse filtering to recover the glottal pulse which can then be characterized both in terms of its wave shape and its spectral components.

Five years ago, at a previous conference in this series, types of phonation that might appear pathological in one language were shown to convey differences in meaning in another language (Ladefoged 1983). The present paper will describe some of the research on phonation types that has been done since that time in the UCLA Phonetics Laboraory. We have used a number of different techniques to analyze a wide variety of languages, as summarized in Table 1.

This paper will focus on the techniques that we have been employing rather than on our linguistic results. Nevertheless, in order to understand why we have used some techniques rather than others, a certain amount of background concerning our general aims is necessary. In all our investigations we are trying to describe the glottal activity during certain sounds. If we want to find out what the vocal cords are doing, the best way is to look at them using fiber optic techniques; and if we want to find out how these actions are achieved we could use electromyography to record the behavior of the muscles. But unfortunately neither of these possibilities is usually appropriate in linguistic investigations. Our speakers are volunteer subjects who cannot be persuaded that intrusive techniques will be helpful to them; they are not in need of speech pathology, and do not want to change the way they speak. As a result, often all we have to work with are acoustic records, sometimes supplemented by simple physiological techniques, such as measures of the oral pressure and airflow.

Table 1. Languages investigated in recent UCLA Phonetics Laboratory work on phonation types.

Languages	Investigators
Burmese	Javkin and Maddieson 1983, Javkin, Antoñanzas-Barroso and Maddieson (1985,1987)
Korean	Dart (1984, 1987)
Jalapa Mazatec	Kirk, Ladefoged and Ladefoged (1984)
Hani, Yi, Jingpho, Wa	Maddieson and Ladefoged (1985a, 1985b)
Hmong	Huffman (1985, 1987)
Hmong,Burmese, !Xóõ	Jackson, Ladefoged, Huffman and Antoñanzas-Barroso (1985a, 1985b)
!Xóõ	Ladefoged and Antoñanzas-Barroso (1985)
Wa, Jingpho, Liangshan Yi	Maddieson and Hess (1986)
Tsonga	Traill (1987)

Apart from not being able to use intrusive techniques, linguistic investigations have two other differences from clinical studies of voice quality. Firstly they are concerned with glottal activities that take place very quickly, rather than with chronic states of the glottis. A distinctive glottal gesture that affects the meaning of a word typically lasts 100-200 ms or less. Methods of studying pathological voices that require continuous phonation for several seconds cannot be used. Secondly linguists are primarily concerned with just those aspects of the sound that are of least interest to many speech scientists. Speech pathologists, for example, need to determine how a given voice deviates from the norm. This is often fairly difficult, as we do not have good measures of what constitutes a normal voice quality. But the linguist's task is even more difficult, as it requires separating out not two but three theoretically distinct (but in practice intermingled) aspects of a sound. A linguist would like to be able to determine which aspects of a speech sound are ascribable to the individual speaker's voice quality, which are characteristic of the particular language, and which occur in all languages, thus forming a non-distinctive part of the human phonetic endowment.

This three-way abstraction from the data largely determines our research procedures. So that we can be sure that we are describing the properties of a language rather than those of an individual, we aim to record several speakers of each language; and so that we can distinguish the properties of a given language from those that are common to all languages, we need to examine data from an adequate sample of the languages of the world. In practice the first of these aims is often difficult to achieve; and we are inevitably a long way from achieving the second.

[The body of the paper then reported the studies listed in Table 1 that had not been published elsewhere. Readers are referred to the cited *Working Papers in Phonetics* for further details.]

Summary of linguistic findings

Although the primary emphasis of this report has been on the techniques we have used for investigating phonation types, it is appropriate to consider briefly some general aspects of the use of linguistic data. Firstly we should note that the study of linguistic contrasts has proved a fertile field for the investigation of different types of phonation. Individuals may differ from one another in the way that they produce a given contrast; but each individual usually pronounces contrasting pairs of words in much the same way every time, enabling us to get good, stable data. Secondly we have been able to document differences among phonation types that are clearly under voluntary control, in that they are used to differentiate words in a language. Thirdly we have shown that contrasts within a language often cannot be characterized in terms of absolute values, but operate on a scale, in much the same way as contrasts in pitch or vowel quality. Finally, we have found that languages differ greatly in the phonation types that they use. There are obvious preferences in that every language uses a voiceless state of the glottis for some sounds, and modal voicing for others (though not all languages have sounds that contrast simply through the use of these different glottal states). But among the languages that contrast more than these two glottal states, there is a wide variety of possibilities, and no clear evidence for a preference involving contrasts of one type as opposed to another. The subtle differences between stiff and slack voice in Jingpho are very different from the modal voice and breathy voice differences in !Xóõ, and these in turn are different from the differences between modal and murmured voice in Tsonga. We have learned a lot from our studies of these and other languages; but there is much more to be done before we can pull all this together into a comprehensive picture.

Acknowledgments

This research was supported by a USPHS grant. We are very grateful to our colleagues, Norma Antoñanzas-Barroso, Sarah Dart, Susan Hess, Marie Huffman, Michel Jackson, Hector Javkin, Paul Kirk, Jenny Ladefoged, and Tony Traill, for all their assistance.

References

Bickley, Corine (1982). Acoustic analysis and perception of breathy vowels. *Working Papers, Speech Communication Group, MIT, 1,* 71-81.

Dart, Sarah N. (1984). Testing an aerodynamic model with measured data from Korean. *UCLA Working Papers in Phonetics, 59,* 1-17.

Dart, Sarah N. (1987). An aerodynamic study of Korean stops; measurement and modeling. *Journal of the Acoustical Society of America, 81.*

Fant, Gunnar (1956). On the predictability of formant levels and spectrum envelopes from formant frequencies. In *For Roman Jakobson*, edited by M. Halle, pp. 109-120. The Hague: Mouton.

Fant, Gunnar (1973). *Speech sounds and features.* Cambridge: MIT Press.

Fujimura, Osamu and Lindqvist, Johan (1971). Sweeptone measurements of vocal tract characteristics. *Journal of the Acoustical Society of America, 49,* 541-557.

Huffman, Marie K. (1985). Measures of phonation in Hmong. *UCLA Working Papers in Phonetics, 61,* 1-25.

Huffman, Marie K. (1987). Measures of phonation in Hmong. *Journal of the Acoustical Society of America, 81.*

Jackson, Michel, Ladefoged, Peter, Huffman, Marie and Antoñanzas-Barroso, Norma (1985a). Measures of spectral tilt. *UCLA Working Papers in Phonetics, 61,* 72-78 .

Jackson, Michel, Ladefoged, Peter, Huffman, Marie and Antoñanzas-Barroso, Norma (1985b). Automated measures of spectral tilt. *UCLA Working Papers in Phonetics, 62,* 77-88.

Javkin, Hector, Antoñanzas-Barroso, Norma and Maddieson, Ian (1985). Digital inverse filtering for linguistic research. *UCLA Working Papers in Phonetics, 60,* 87-100.

Javkin, Hector and Ian Maddieson (1983). An inverse filtering study of creaky voice. *UCLA Working Papers in Phonetics, 57,* 115-125.

Kirk, Paul L., Ladefoged, Peter and Ladefoged, Jenny (1984). Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics, 59,* 102-113.

Ladefoged, Peter (1967). *Three areas of experimental phonetics.* London: Oxford University Press.

Ladefoged, Peter (1983). The linguistic use of different phonation types. In *Vocal fold physiology: contemporary research and clinical issues.*, edited by Diane Bless and James Abbs, pp. 351-360. San Diego: College Hill Press.

Ladefoged, Peter and Antoñanzas-Barroso, Norma (1985). Computer measures of breathy phonation. *UCLA Working Papers in Phonetics, 61,* 79-86.

Ladefoged, Peter and Traill, Anthony (1984). Instrumental phonetic fieldwork. In *Topics in linguistic phonetics in honour of E. T. Uldall,* edited by Jo-Ann Higgs and Robin Thelwall. Occasional Papers in Linguistics and Language Learning 9:1-22. The New University of Ulster.

Maddieson, Ian and Hess, Susan (1986). "Tense" and "lax" revisited: more on phonation type and pitch in minority languages in China. *UCLA Working Papers in Phonetics, 63,* 103-109 .

Maddieson, Ian and Ladefoged, Peter (1985). "Tense and "lax" in four minority languages of China. *UCLA Working Papers in Phonetics, 60,* 59-83.

Rothenberg, M.A. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *Journal of the Acoustical Society of America, 53.6,* 1632-1645.

Traill, Anthony (1985). *Phonetic and phonological studies of !Xóõ Bushman* (Quellen zur Khoisan-Forschung 1). Hamburg: Helmut Buske.

Traill, Anthony and Jackson, Michel (1987). Speaker variation and phonation type in Tsonga nasals. To appear in *UCLA Working Papers in Phonetics, 67*.

Wakita, Hisashi, and Fant, Gunnar (1978). Toward a better vocal tract model. *Quarterly Progress and Status Report, 1/79,* 9-29. Stockholm: Royal Institue of Technology, Speech Transmission Laboratory.

Yumoto, Eiji, Gould, Wilbur J., and Baer, Thomas (1982). Harmonics-to-noise ratio as an index of the degree of hoarseness. *Journal of the Acoustical Society of America, 71.6,* 1544-1550.

ANOTHER LOOK AT THE REGISTER DISTINCTION IN MON

Theraphan L.Thongkum

Chulalongkorn University, Bangkok

1. *Introduction*

Lee (1983) investigated the acoustical parameters governing the register distinction in Mon and the relative significance of these parameters. Four parameters were looked at : vowel duration, frequencies of the first two formants, fundamental frequency, and distribution of spectral energy. He concluded that for citation forms, only two of these parameters indicated that significant differences exist between the two registers. The socond register vowels had longer duration and lower pitch level. No consistent difference could be found between the two registers with respect to vewel quality or phonation type. He ended his paper by saying :

> Further, our findings suggest that the most significant parameter of the register distinction is that of pitch, in particular the onset Fo and the overall pitch level. Indeed, as Shorto (1962) suggests, Mon is a quasi-tonal language.
>
> (Lee, 1983 : 95)

Diffloth (1985) disagreed with Lee's experimental findings. He did not believe that numerical acoustic coefficient could tell us that one phonetic parameter is more important than the other. He commented :

> An answer to Lee's question does not come from acoustical measurements alone; it would have required the use of a speech synthesizer able to imitate a wide spectrum of phonation types, as well as pitches, and the computation of recognition and error responses from native speakers of Mon.
>
> (Diffloth, 1985 : 56)

He confirmed what he thought he had heard by citing Shorto (1967), Haswell (1874), Blagden (1910), Halliday (1922), Huffman (1976), and Sakamoto (1974) : "For Mon, head register is characterized by clear voice and chest register by breathy voice".

It is possible to end the argument by saying that the problem has been caused by dialect variations which Lee forgot to mention in his article. Diffloth had worked on Burmese Mon for a few years, and he was the one who prepared the wordlist and lined up Mon informants for Ladefoged when the recordings were made in a Buddhist temple in Bangkok. I witnessed the whole procedure. Later, the recordings were used by Lee for his acoustical measurements of "Mon register distinction".

From the above account, we might be able to point out the following weaknesses, if there have been any :

1. The so-called "register distinction" in Mon as has been reported in the literature is only a myth. In fact, Mon is a two-register tone language (as defined by Pike 1948).

2. Influenced by the literature of Mon language studies, Diffloth was inclined to hear what he did not really hear while collecting field data. He happened to work on Mon dialects that do have tone instead of register contrast. As a result, the wordlist prepared by him and the Mon informants used were not suitable for an acoustical study of a register language, which always involves phonation types.

3. The Mon monks who acted as informants were not familiar with the situation, i.e. three foreigners and a Thai woman directing them to say some words and sentences in order to make good tape-recordings. Thus, their speech could have been unnatural. In general, the Mon, at least Thai Mon, are ashamed of "speaking with their throat, not their mouth," as described by themselves and the Thai. Perhaps, the informants could control their laryngeal setting so well that phonation differences did not occur when the recordings were being made.

4. Different phonation types do exist, but Lee failed to prove it to us because the technique of measuring used in his study was not adequate. However, the editor of UCLA Working Papers in Phonetics 60 notes at the end of Diffloth's paper (1985 : 57) that "although some other technique might reveal a difference in laryngeal setting for the registers in Mon, careful listening by several persons with 'trained ears' in the UCLA Phonetics Laboratory does not suggest that breathiness is at all a consistent feature of 'chest register' in the tape recordings we have, whereas an observable and statistically reliable pitch height difference does occur".

I must admit that I am not satisfied with the explanation given above. I have worked on tone languages of Southeast Asia that have phonation type as a property of tone, and also on Mon-Khmer register languages. My intuition and experience tell me that something has gone wrong, and that I should help settle the matter. Let us examine carefully what the scholars of Mon-Khmer language have said about the registers in Mon.

Haswell (1901) points out that the Peguan (Mon) alphabet consists of twelve vowels, and that "the vowels are mostly in pairs, the first is light, the second is heavy sound, of what might properly be called the same vowels" (p. 1). It is very interesting that Haswell describes Mon vowels in terms of light vs. heavy. He could have heard some kinds of voice quality. At present, the Mon in Thailand also describe their language as having light words vs. heavy words. We can not tell whether Haswell really heard phonation-type differences or he was influenced by the traditional way of explaining Mon vowel sounds.

Blagden (1910) associates "glottal activity" with the initial consonants in Mon. He notices that the so-called "voiced and voiced aspirated stops" are in fact "voiceless and voiceless aspirated stops" accompanied by glottal activity which also influences the following vowel. He says :

> ....the consonants : g, gh, j, ḍh (only used in Pâli words), d, dh, b, bh called "voiced" are actually pronounced voiceless : k, kh etc. ... but their pronunciation is accompanied by glottal activity which distinguishes them fairly clearly from the consonants of the first series; this (glottal activity) profoundly modifies the vowel which follows in a way which is difficult to describe, but seems in certain cases to be a rather guttural quality pertaining to the posterior part of the oral cavity. ...
>
> (Blagden 1910 : 479)

Neither Haswell nor Blagden mention pitch differences at all.

Shorto (1962) uses the term "register" as defined by Henderson (1952) to describe the tenseness vs. laxness in Mon speech. He says :

> The quasi-tonal register distinction, ..., is inherent in all Mon words. Chest register, characterized by breathy voice quality in association with a general laxness of speech organs, and somewhat centralized articulation of vowels, ...; head register, characterized by clear voice quality, ...
>
> (Shorto 1962 : x)

Shorto (1967) gives more phonetic details on the register distinction in Mon. He points out the differences of vowel quality, consonant articulation and voice quality, but not pitch differences. The tense-lax distinction affects not only single consonants or vowels, but the whole complex of the word. He states :

> The exponents of register are diverse in character, comprising a difference of voice quality; differences of vowel quality, slight in some cases but in others considerable; and in some cases differences in consonant articulation. ... . Contrastive voice quality is always present and is probably the feature most readily perceived.
>
> Head register is characterized by clear voice throughout the word or equivalent segment. Chest register is characterized by a breathy voice with lowering of the glottis and a relatively centralized articulation of vowels, ...
>
> Pitch difference as an exponent of register is lacking. ... .
>
> A unitary formulation of the diverse exponents of Mon register — the differences of voice quality, of vowel articulation, and of consonant articulation — may be sought in terms of a tense-lax opposition affecting not merely single phonemes but the whole complex of the word or equivalent segment. Thus in chest register laxness results not only in the voicing of prevocalic consonants, but also in less vigorous movements of the tongue towards the periphery, leading

to the relative centralization of vowels noted above.

(Shorto 1967 : 246)

Huffman (1976) recognized a similar type of register phenomenon in a Thai Mon dialect spoken in Ban Bang Kradi. Briefly he says :

> The register distinction in Mon is relatively subtle; 2nd register vowels are mildly lax and breathy, and are slightly lower in quality (more open) than their 1st register counterparts. The distinction is particularly difficult to hear in the low front /ɛ/ ≠ /ɛ̀/ position.
>
> (Huffman 1976 : 585)

Diffloth (1984) uses the field data collected by himself at three Thai Mon villages : Ban Nong Du, Ban Bang Khan Mak and Ban Nakhon Chum, and at many Burmese Mon villages, together with the language data gathered from Shorto (1962) and Sakamoto (1974) for reconstructing Proto-Mon and Proto-Monic. Regarding the register distinction in Mon, he says that "Mon has a contrast between vowels pronounced with a clear voice and vowels with a breathy voice, and that there are actually important differences in the phonetic features which accompany these two registers in Mon, notably in pitch patterns" (p. 52).

It is important to note that every scholar of Mon-Khmer languages has talked about voice quality and phonation type in Mon, but only a few of them mention pitch. In his book which appeared in 1962, Shorto used the term "quasi-tonal register distinction", and later on in 1967, in his article on "The register distinctions in Mon-Khmer languages," Shorto pointed out "the lack of pitch difference" in Mon. Even though Diffloth disagrees with Lee's conclusion that Mon is "a quasi-tonal language" because the pitch differences in Mon are statistically significant, he himself mentions "pitch patterns" in his book (Diffloth 1984).

2. *Language data*

The number of Mon in Thailand is estimated at 200,000.Hundreds of Mon villages are scattered in the central region of Thailand. In 1981, I visited many Mon villages to collect language materials for a dialect survey of Mon. I noticed that the Mon inhabiting different areas spoke differently. My informants often pointed out to me that the Mon living in a nearby village or across the river spoke Mon with a different accent. This was due to the fact that the Mon population in Thailand migrated from different regions of Burma, and that they entered the country at different periods. In November 1986, I stayed about a week in a Mon village located in Nakhon Chum Sub-District, Ban Pong District, Rajaburi

Province. About 1,000 words were collected during this field trip. After examining the data obtained carefully, I selected about 116 word pairs for good quality recordings. The pronunciation of these 116 minimal or nearly minimal pairs was carefully checked. Eight Mon speakers from Ban Nakhon Chum volunteered to come to our recording studio in Bangkok. Most speakers of Thai Mon are literate in Thai; they cannot read Mon script. Writing Mon with Thai script is not an easy task either. The only thing I could do was to elicit the word pairs I wanted by means of interviewing them. During the interviews, the Thai glosses were used as clues. It took quite a long time to obtain the data from eight speakers. The tapes were edited later. However, only 16 word pairs were used for acoustic analysis, the results of which are being submitted in this paper. They are as follows :

1. a) /hərip/ "to blink"
   b) /həri̤p/ "to snatch and run away"
2. a) /ɓi/ "river"
   b) /ɓi̤/ "you (vulgar)"
3. a) /tep/ "a kind of bamboo trap"
   b) /te̤p/ "woman who has a lover"

Due to bad editing of the tapes, the word pair /tep/ and /te̤p/ had to be replaced by /cep/ "to taste" and /ce̤p/ "to seep" for S6 (speaker 6) and S8 speaker 8).

4. a) /həte/ ~ /te/ "to forge iron"
   b) /həte̤/ ~ /te̤/ "mercury"
5. a) /wɛk/ "slightly torn (of cloth)"
   b) /wɛ̤k/ "to tuck behind the ear (as a flower)"
6. a) /phɛŋ/ "split bamboo"
   b) /phɛ̤ŋ/ "marijuana"
7. a) /cək/ "to gore"
   b) /cə̤k/ "rope, cord"
8. a) /cə/ "to shield"
   b) /cə̤/ "to bump into"
9. a) /pat/ "to smooth and level off"
   b) /pa̤t/ "Mon orchestra"
10. a) /ɗaŋ/ "expensive"
    b) /ɗa̤ŋ/ "king posts (which support the ridge-pole)"
11. a) /put/ "to carve"
    b) /pṳt/ "to rub across (as when making a fire)"

12. a) /cu/ "steep"
    b) /cṳ/ "to stop (in order to rest)"
13. a) /pot/ "to polish"
    b) /po̤t/ "pot"
14. a) /əto/ "ear"
    b) /əto̤/ "jujube"
15. a) /thɔt/ "forcefully"
    b) /thɔ̤t/ "to deepfry"
16. a) /ɓɔŋ/ "to give birth"
    b) /ɓɔ̤ŋ/ "bamboo tube for smoking marijuana"

The above 16 word pairs were chosen for the acoustic analysis of Mon vowels because of the three following reasons : a) they are minimal pairs; b) they all contain pure vowels : i i̤ e e̤ ɛ ɛ̤ ə ə̤ a a̤ u ṳ o o̤ ɔ and ɔ̤; and c) they represent two types of syllable structure, namely dead or checked syllable (CVC̥) and live or ordinary syllable (CV, CVC). During my field trip, I had noticed that syllable structure had some influence on the phonetic characteristics of vowel length, vowel quality and pitch.

3. *Wideband spectrograms*

Wideband spectrograms can provide good displays of the acoustics of different phonation types. During the creaky voice vowels, the vertical striations (i.e. glottal pulses) occur at irregularly spaced intervals. The formants are fairly clear during the modal voice vowels and are less well-defined for breathy voice vowels. (Kirk, et al. 1984).

After investigating all the wideband spectrograms of the test words that I made in the Phonetics Laboratory of the Linguistics Department at UCLA, I decided to discard the labels "modal voice vowels" vs. "breathy voice vowels". The choice of the labels "tense vowels" vs. "lax vowels" as suggested by Maddieson and Ladefoged (1985) seems to be more appropriate. Although most speakers pronounced first register vowels with modal (clear or normal) voice and second register vowels with breathy voice (see Figure 1), some speakers do not make this type of phonation distinction. For example, S5 (female) makes a distinction between two different degrees of breathiness, less breathy voice for first register vowels and more breathy voice for second register vowels, as shown in Figure 2. In comparison with the other subjects, S8 (male) has a very low voice. He seems to make three types of phonation distinction : modal voice vs. breathy voice, creaky voice vs. breathy voice, and creaky voice (in the middle or at the end) vs. a

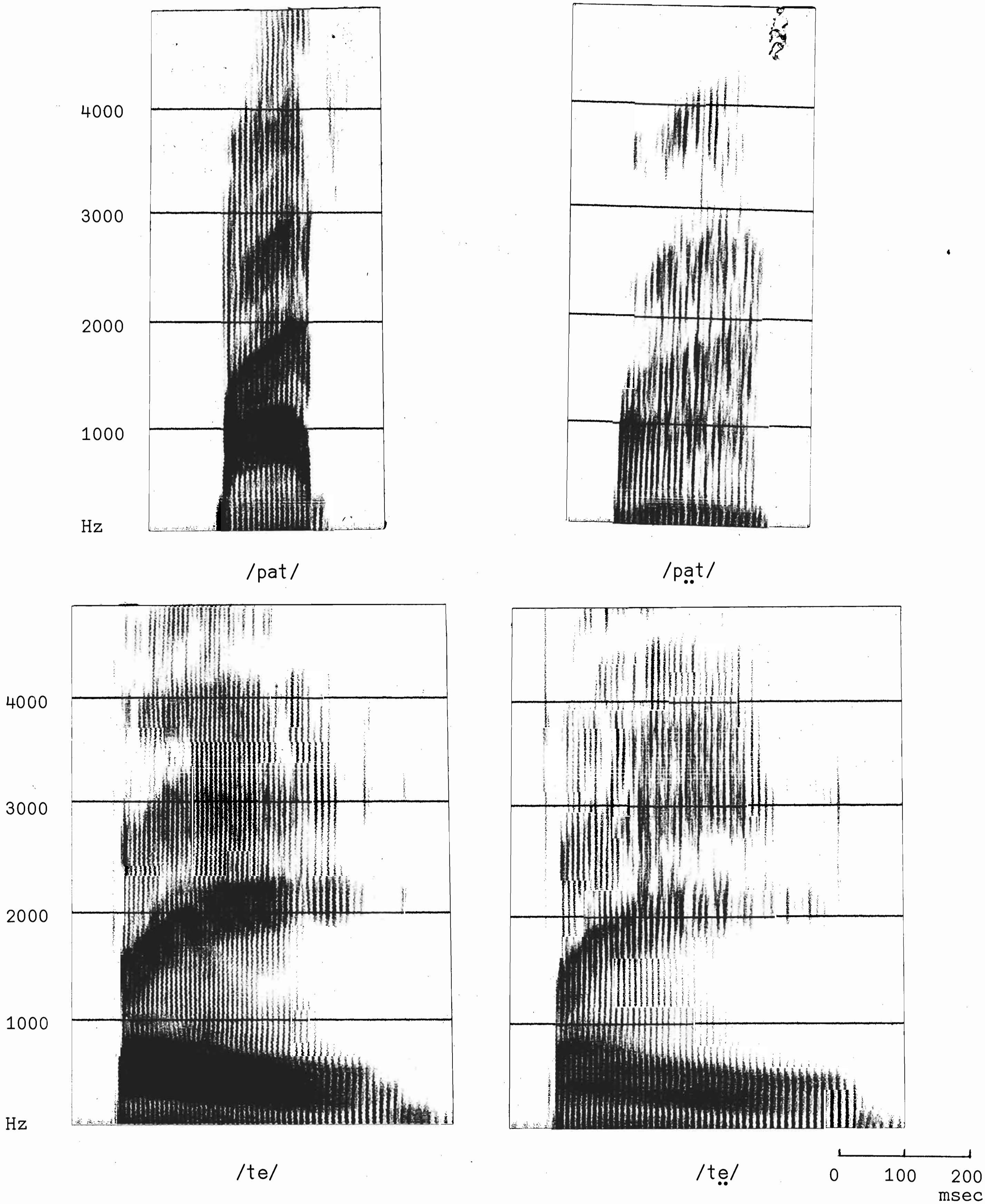


FIGURE 1 : Wideband spectrograms of vowels with modal voice and breathy voice of Speaker 7.

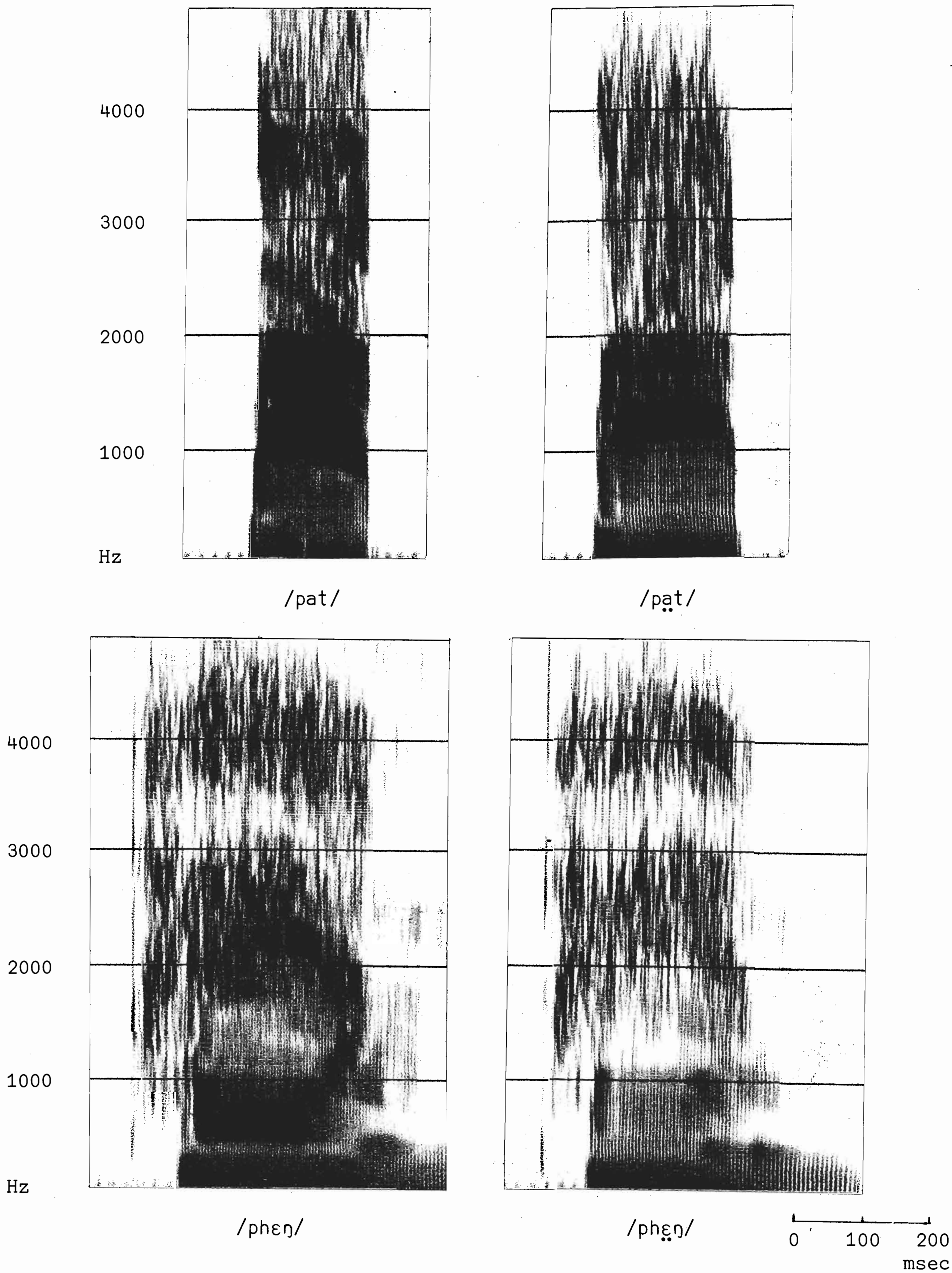


FIGURE 2 : Wideband spectrograms of vowels with less breathy and more breathy voice of Speaker 2.

combination of breathy and creaky voice (perhaps, whispery-creaky voice), as shown in Figures 3-5.

From investigating the wideband spectrograms of 128 test words pronounced by the eight Mon speakers, we may conclude that different kinds of phonation types do occur, at least, in the Mon dialect of Nakhon Chum.

4. *Power spectra*

Phonation type differences are reflected in spectral energy distribution. Narrowband power spectra offer a way of quantifying the spectral tilt. Power spectra produced by the sound spectrograph can be used to quantify the relative amount of energy in different harmonics (Fischer-Jørgensen 1977, Stevens 1981, Lee 1983, Kirk et al. 1984, Maddieson and Ladefoged 1985, Ladefoged et al. 1987, etc.). To detect phonation type differences in Mon, the difference in dB between the amplitude of the fundamental (Fo) and intensity of the second harmonic (Hz) was measured (see Figure 6). The spectra were taken at half or one-third of vowel duration. The difference between the amplitude of the fundamental and that of the second harmonic is displayed in Figures 7 and 8. For the eight speakers, the mean for tense voice (modal voice or slightly breathy voice or slightly creaky voice) is 2.834 dB with a standard deviation of 4.17 (i.e. the fundamental has 2.834 dB less amplitude than the second harmonic). The mean for lax voice (heavily breathy voice or breathy creaky voice) is -3.228 dB with a standard deviation of 5.293 (i.e. the fundamental has -3.228 dB more amplitude than the second harmonic). The difference is highly significant ($p < .0005$ t = 15.378, df = 128). See Figure 7. For each speaker the relationship of the fundamental and the second harmonic is shown in Figure 8 and Table 1. My findings indicate that an overall phonation difference in Mon does exist. This means that my study supports the claim of Mon-Khmer specialists, such as Shorto, Huffman, Diffloth, etc., that the fundamental difference between the register distinction in Mon lies in a phonation difference.

5. *Fundamental frequency*

Narrowband spectrograms were made and measured at five points starting from the onset to the end of the vowel (see Figure 9). For plotting the results of Fo measurements the 32 test words were divided into four sets based on phonation-type differences and syllable types : CVC̥ (consonant - tense vowel - voiceless stop), CV̤C̥ (consonant - lax vowel - voiceless stop), CV(C) (consonant - tense vowel - ø final or nasal or semivowel), and CV̤(C) (consonant - lax vowels - ø final, nasal or semivowel). The mean values of Fo (in Hz) of the eight speakers

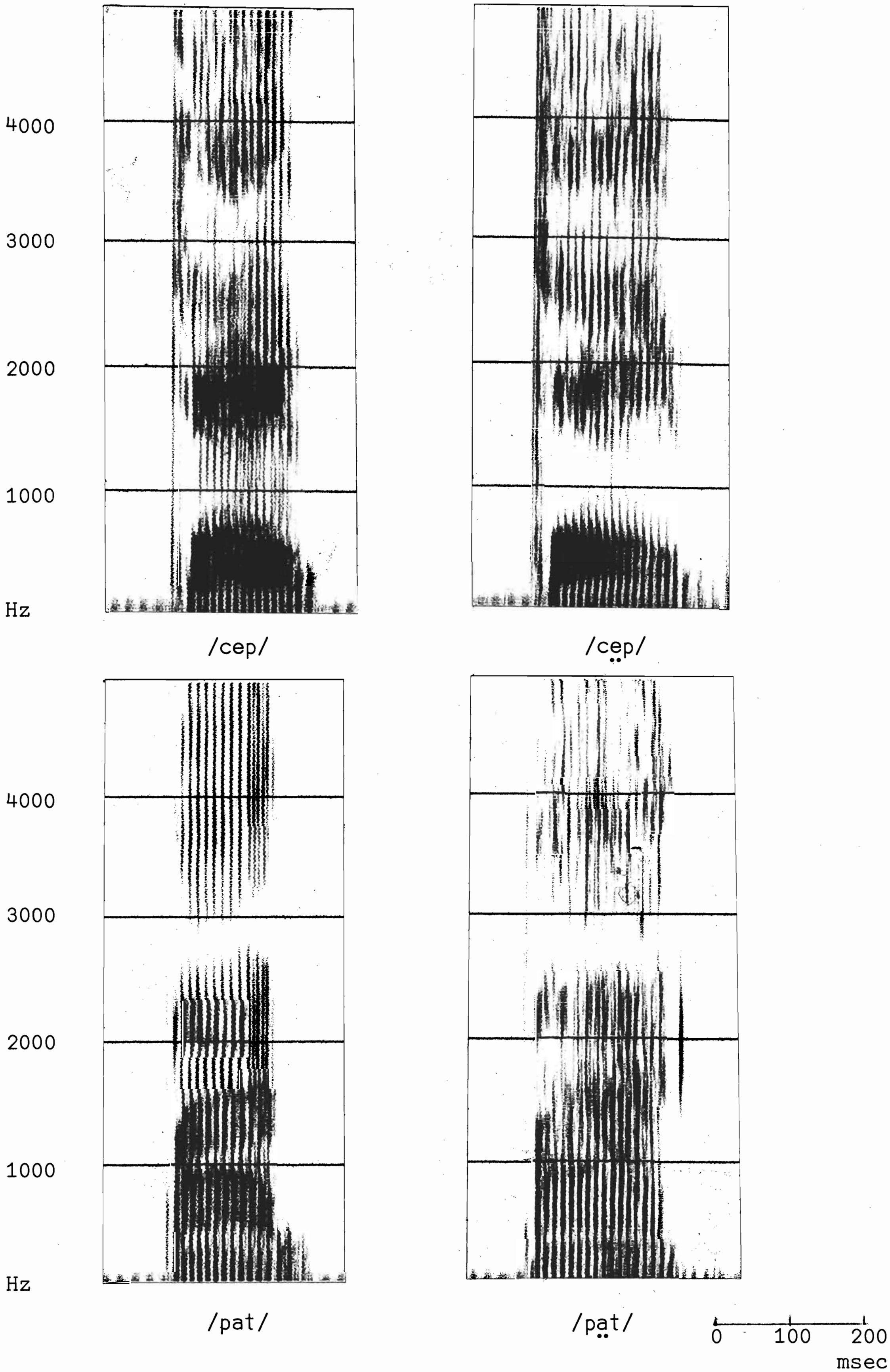


FIGURE 3 : Wideband spectrograms of vowels with modal voice and breathy voice of Speaker 8.

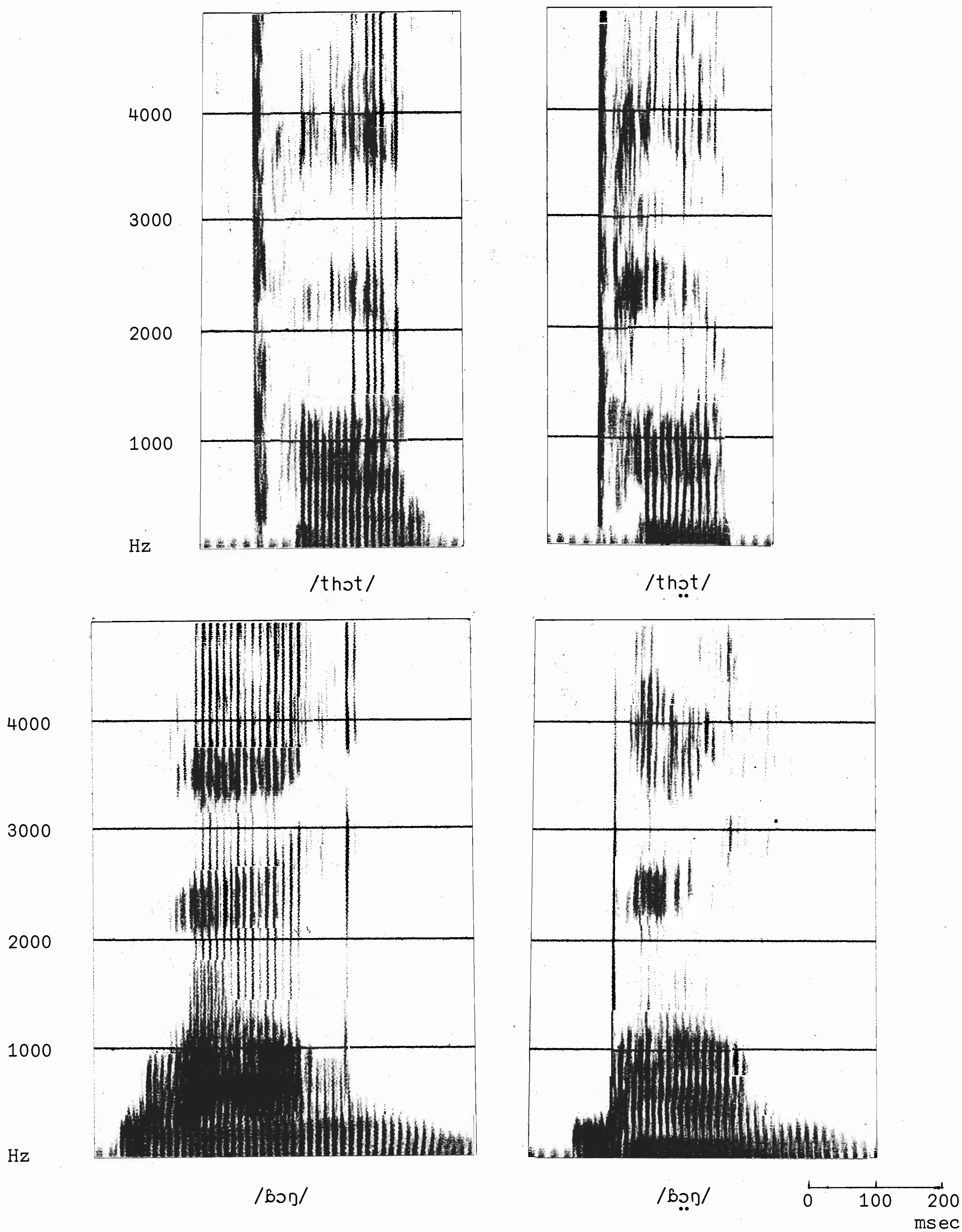


FIGURE 4 : Wideband spectrograms of vowels with creaky voice and breathy voice of Speaker 8.

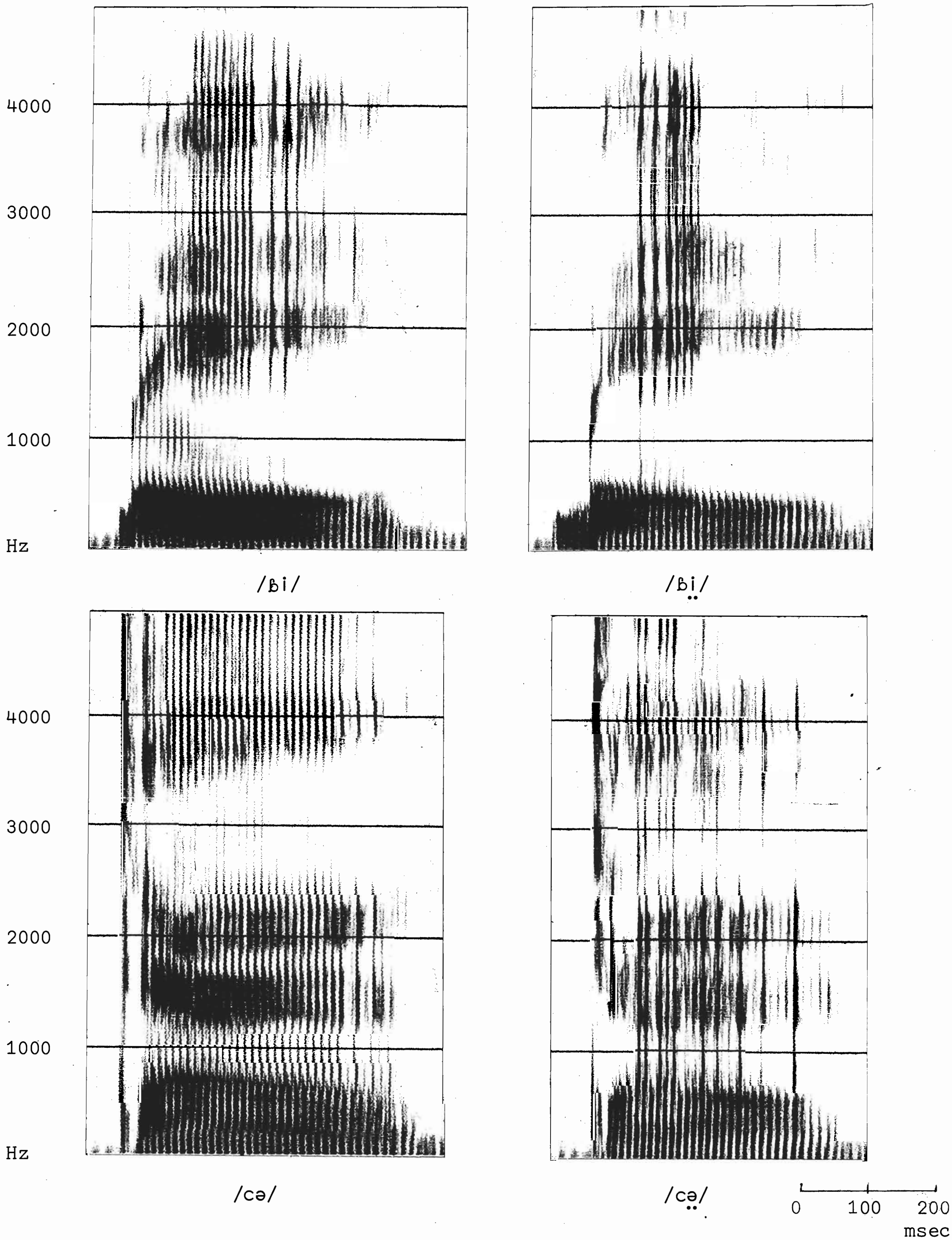


FIGURE 5 : Wideband spectrograms of vowels with creaky voice and a combination of breathy and creaky voice of Speaker 8.

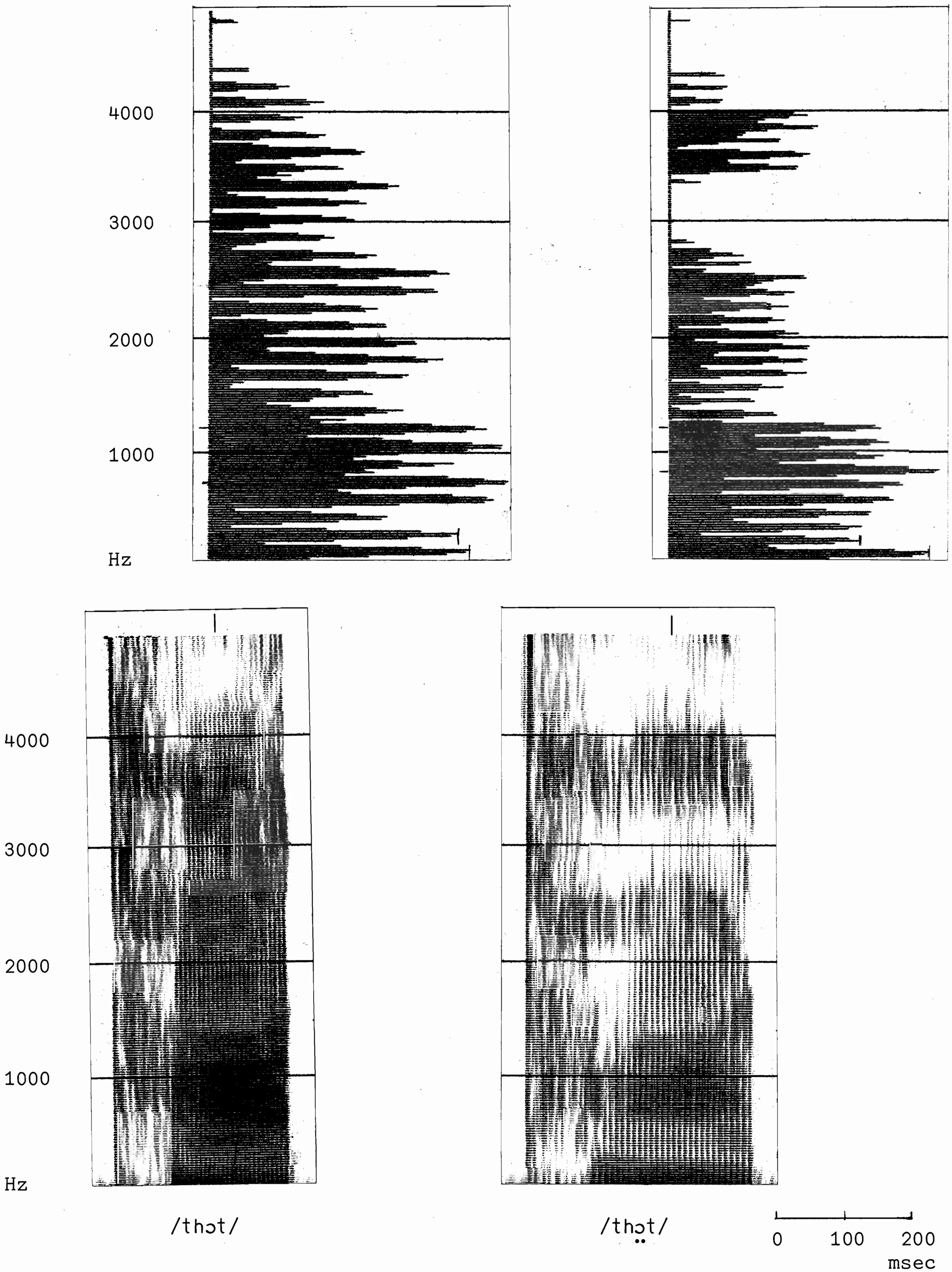


FIGURE 6 : Power spectra and wideband spectrograms of the tense and lax vowels in /thɔt/ and /thɔ̤t/ of Speaker 2.

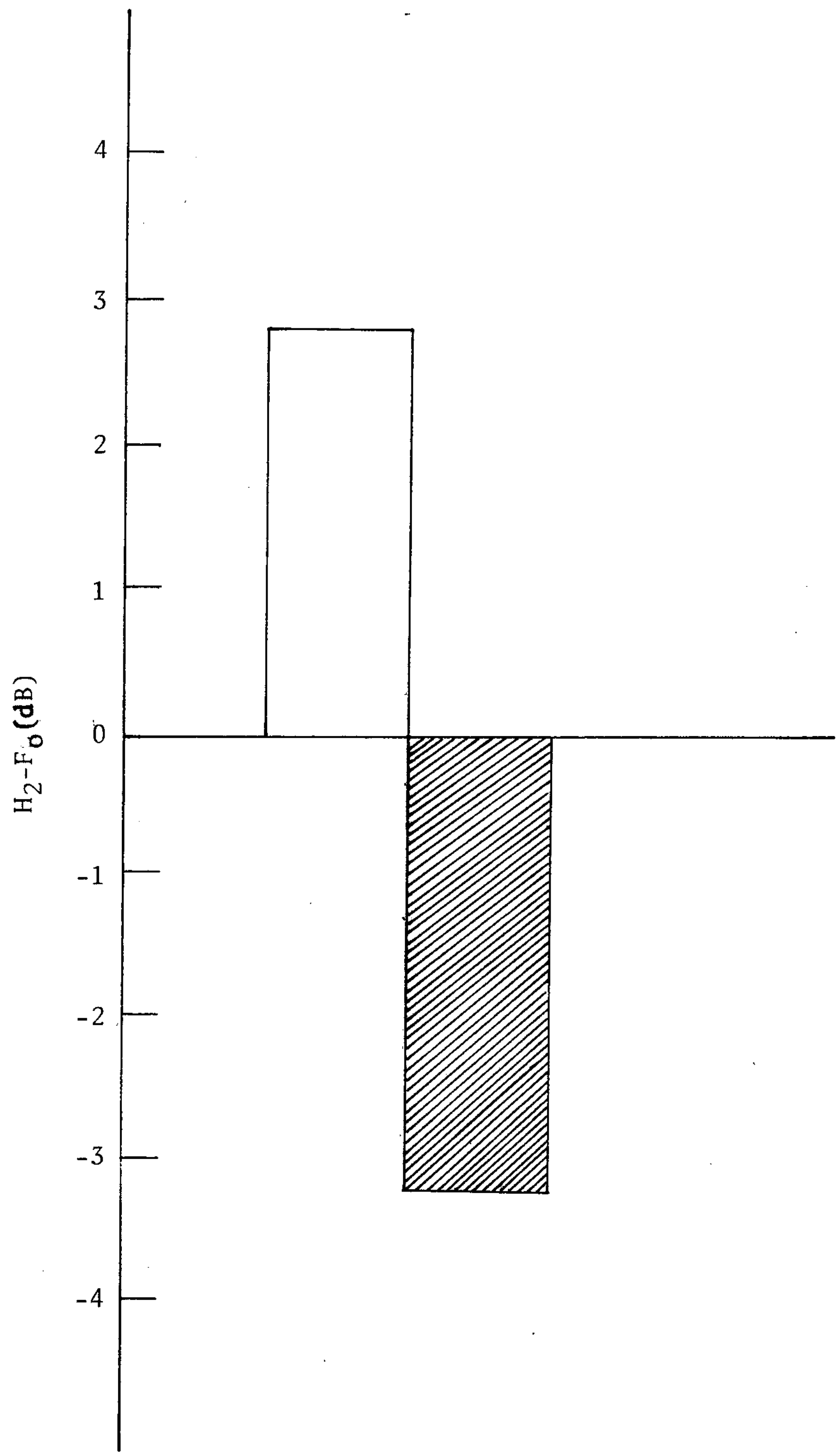
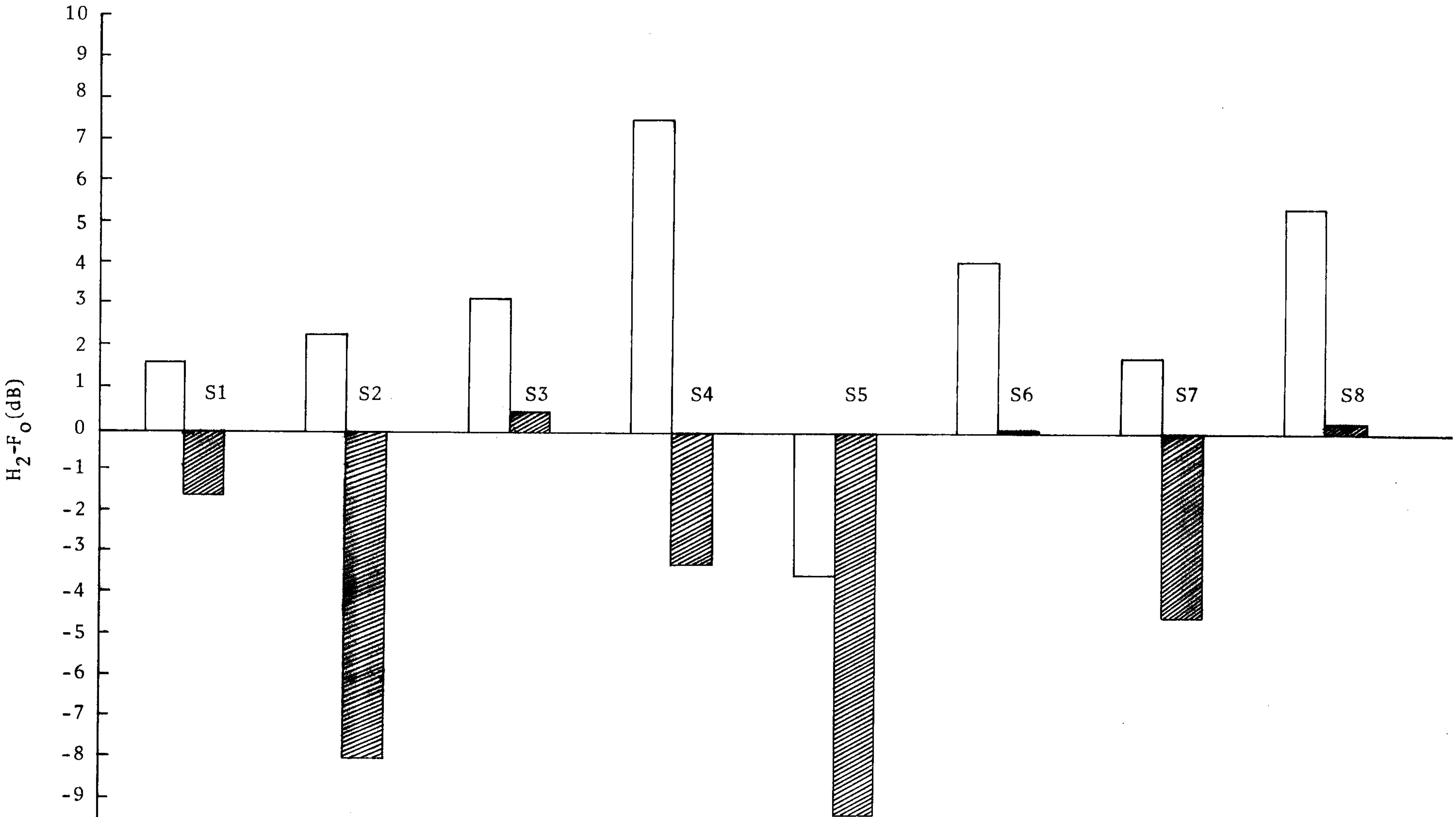


FIGURE 7 : Mean of the H_2-F_0 measure illustrating the relationship between the amplitude of the fundamental and that of the second harmonic of 8 speakers.

tense vowel lax vowel



<u>FIGURE 8</u> : H_2-F_o measure of each speaker.

tense vowel   lax vowel

TABLE 1 : Relative amplitude of the fundamental (Fo) and the second harmonic (Hz) in dB. (H_o : $\mu_1 = \mu_2$, $p < .005$, df 15 (n-1), t = 2.947)

		First register vowel	Second register vowel
		Tense voice	Lax voice
		(H_2-F_o)	(H_2-F_o)
SPEAKER 1	$\bar{x}$	1.64	-1.54
	SD	4.80	5.26
	t	4.834	
SPEAKER 2	$\bar{x}$	2.35	-7.98
	SD	1.03	2.53
	t	16.688	
SPEAKER 3	$\bar{x}$	3.21	0.49
	SD	2.24	3.01
	t	3.147	
SPEAKER 4	$\bar{x}$	7.54	-3.23
	SD	1.97	6.57
	t	7.602	
SPEAKER 5	$\bar{x}$	-3.46	-9.3
	SD	4.53	2.65
	t	6.706	
SPEAKER 6	$\bar{x}$	4.14	0.063
	SD	2.51	3.08
	t	6.885	
SPEAKER 7	$\bar{x}$	1.79	-4.49
	SD	1.55	3.94
	t	6.399	
SPEAKER 8	$\bar{x}$	5.47	0.16
	SD	2.82	3.76
	t	6.481	

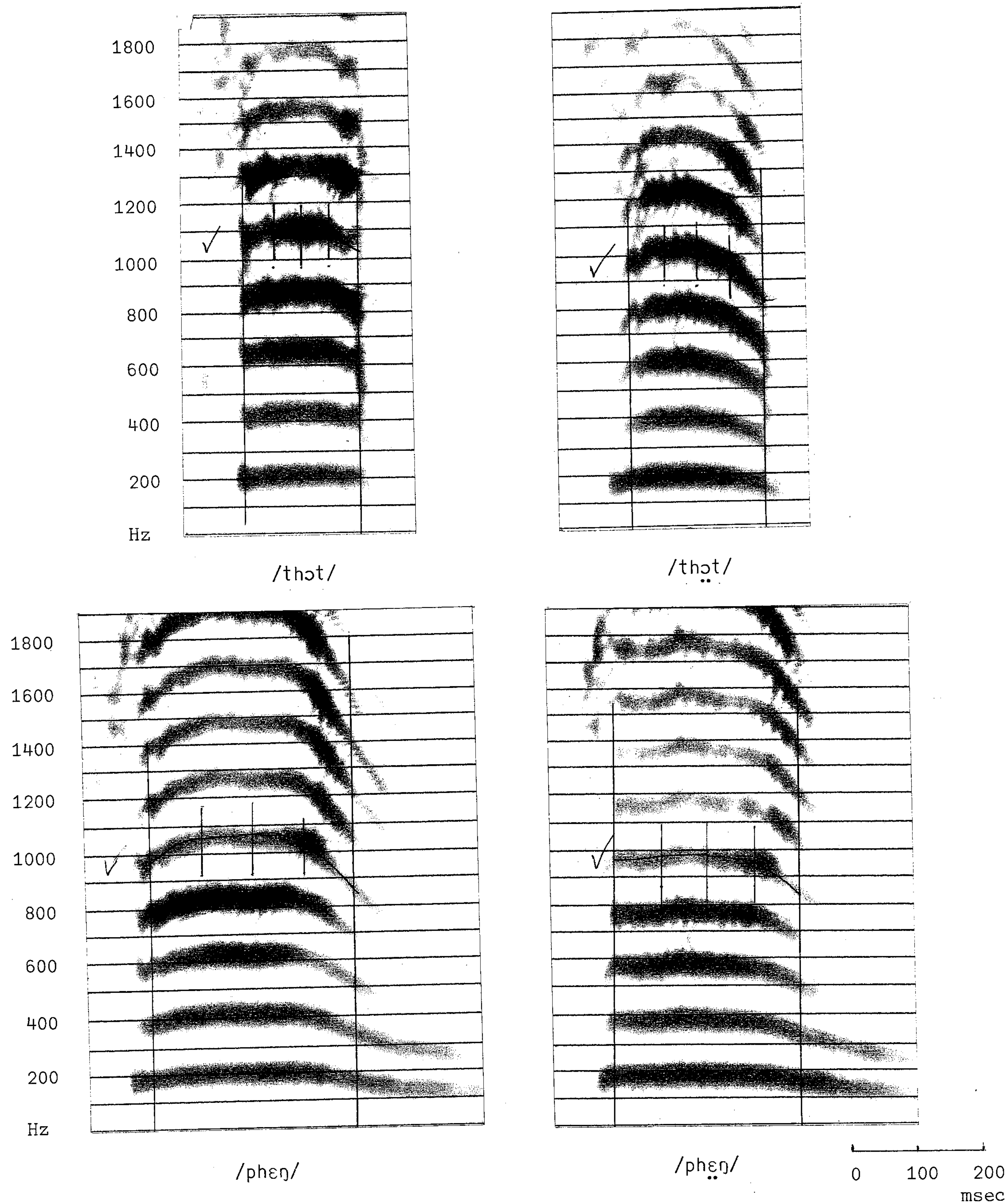


FIGURE 9 : Narrowband spectrograms of Speaker 6 displaying the five points of F_o measurement.

and of each individual speaker are displayed in Figures 10, 11 and 12. The results given in Table 2 show that the difference in fundamental frequency at 0% (the beginning), 25%, 50% (the mid-point), 75% and 100% (the end point) of the harmonics is statistically significant. My findings here agree with those of Lee (1983).

It should be pointed out that a) vowels in CVC̥ and CV̤C̥ syllable types carry higher pitch and less fall pitch contour than vowels in CV(C) and CV̤(C) syllable types; b) tense vowels have higher pitch than lax vowels; and c) vowels in both registers have fall or rise-fall pitch contour. Most speakers seem to use similar pitch patterns, except S5 (female) and S8 (male). Figure 12 indicates that in both registers, S5 shows a more sharp fall contour, and that S8 does not seem to have pitch difference between the two registers, especially in CV(C) and CV̤(C) syllable types.

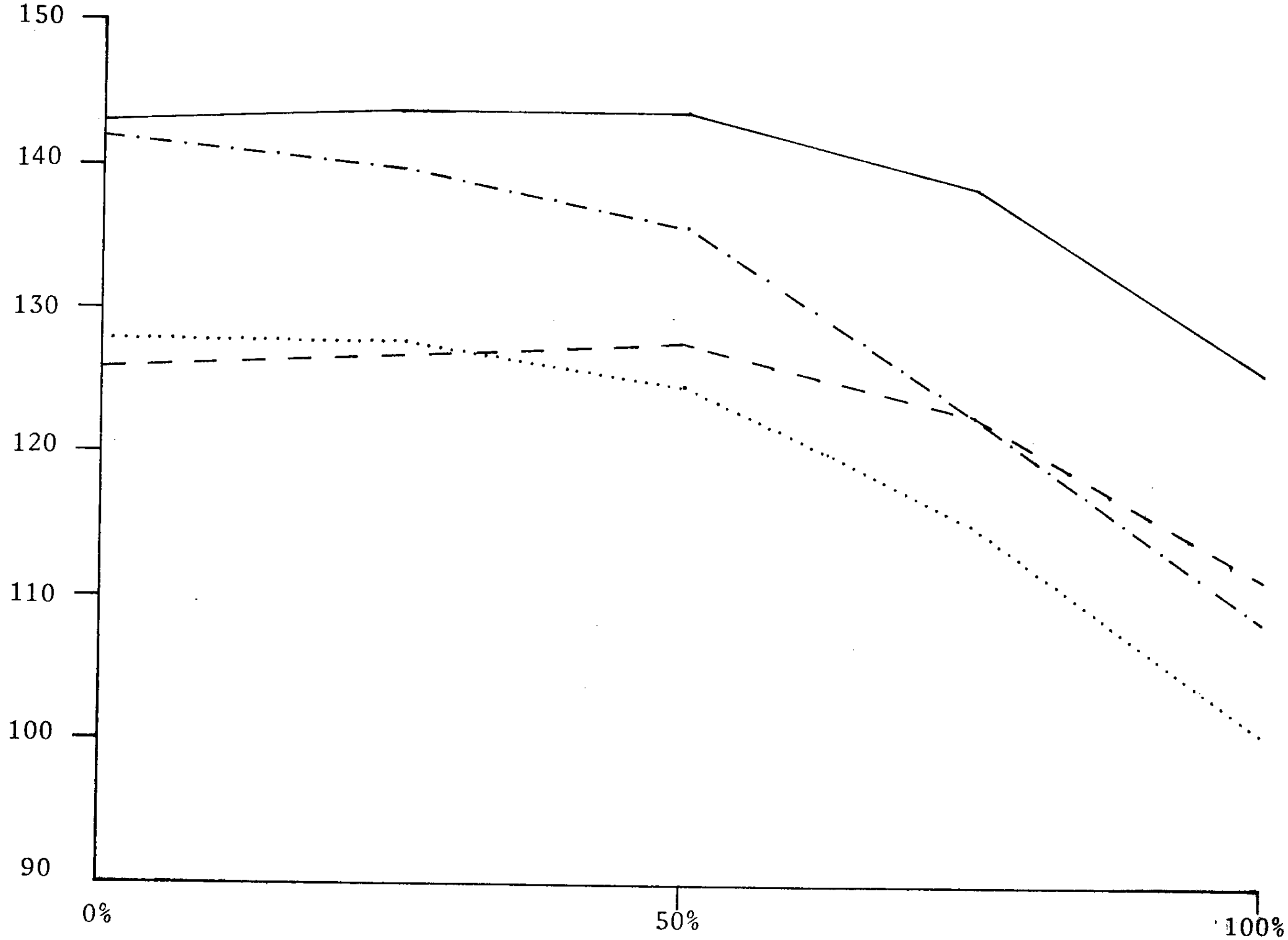
6. *Vowel duration*

Although vowel length in Mon is not linguistically significant, both short and long vowels can be heard. In general, vowels in checked syllables (CVC̥ and CV̤C̥) are shorter than those in other types of syllables.

The mean duration of vowels in CVC̥ and CV̤C̥ syllable types of the eight speakers is 155.14 msec with a standard deviation of 23.22 and 193.08 msec with a standard deviation of 36.61, respectively. The duration difference between tense and lax vowels is statistically significant ($p < .0005$, $df = 64$, $t = 10.824$). In CV(C) and CV̤(C) syllable types, the mean duration of tense and lax vowels are 328.84 msec and 331.70 msec with a standard deviation of 92.76 and 93.14, respectively. The different is so minute that it is not statistically significant ($t = 0.841$, $df = 64$). See Figure 13. The duration of vowels of each speaker is shown in Figures 14-15 and Tables 3-4. Lee (1983 : 82) points out that the second register vowels are longer in duration than the first register counterparts, and that the difference is statistically highly significant ($p < .001$). His finding about vowel duration in Mon is somewhat true.

7. *Formant frequencies*

The mean formant frequencies of F_1 and F_2 at the steady state of the vowel duration of the eight speakers are shown in Figures 16 and 17. In every type of syllable, the lax vowels i̤, e̤, ɛ̤, a̤ and o̤ are higher than the tense vowels i, e, ɛ, a and o, but the tense vowels ə, u and ɔ are higher than the lax vowels ə̤, ṳ



<u>FIGURE 10</u> : Mean Fo values of tense and lax vowels in 4 types of syllable structure of 8 speakers.

——— CVC̥   —·— CV(C)

– – – CV̈C̥   ······ CV̈(C)

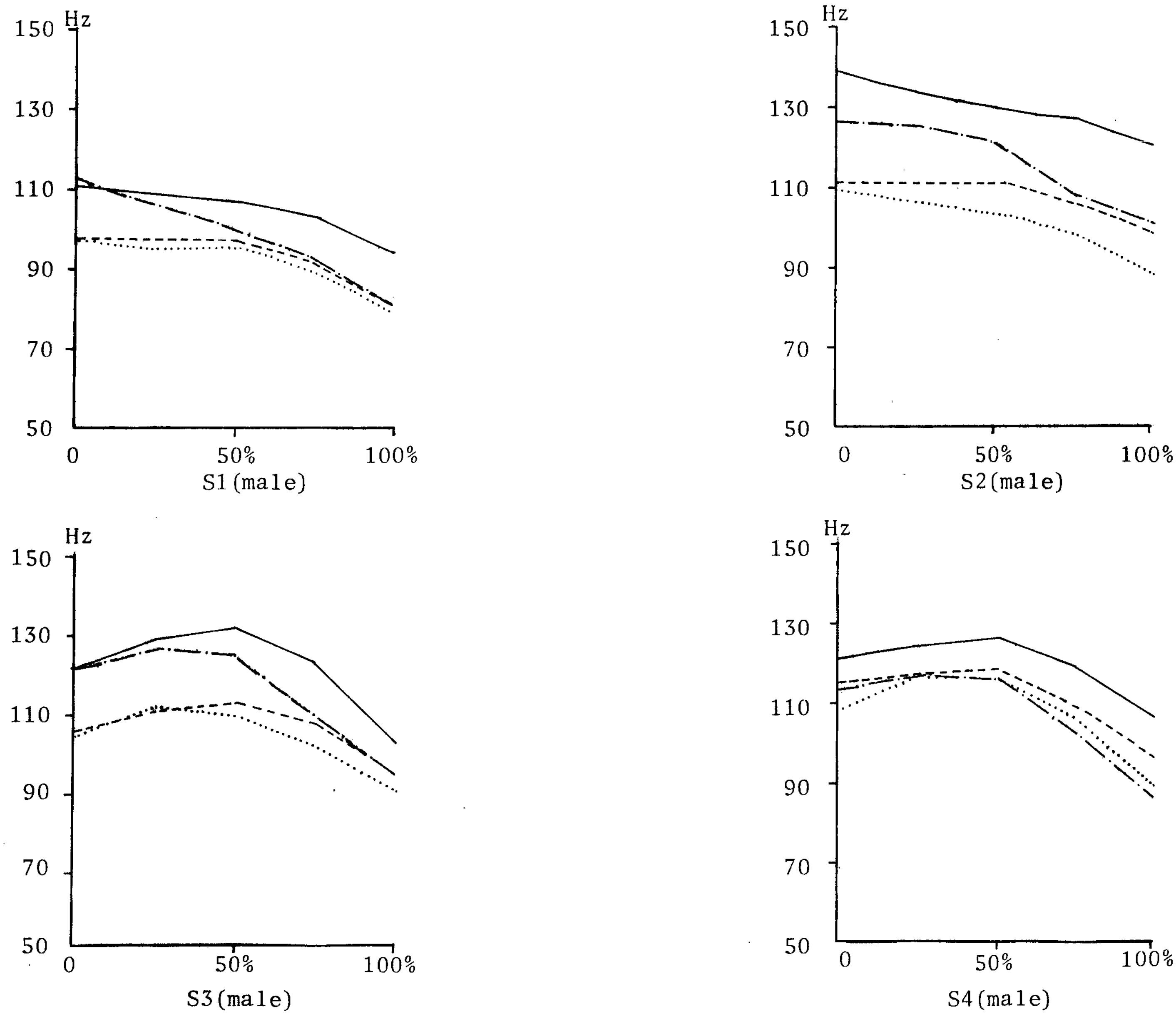


FIGURE 11 : Fo contours of tense and lax vowels in 4 syllable types of S1,S2,S3 and S4.

———— CVC̥ —·— CV(C)

- - - - CV̤C̥ ········ CV̤(C)

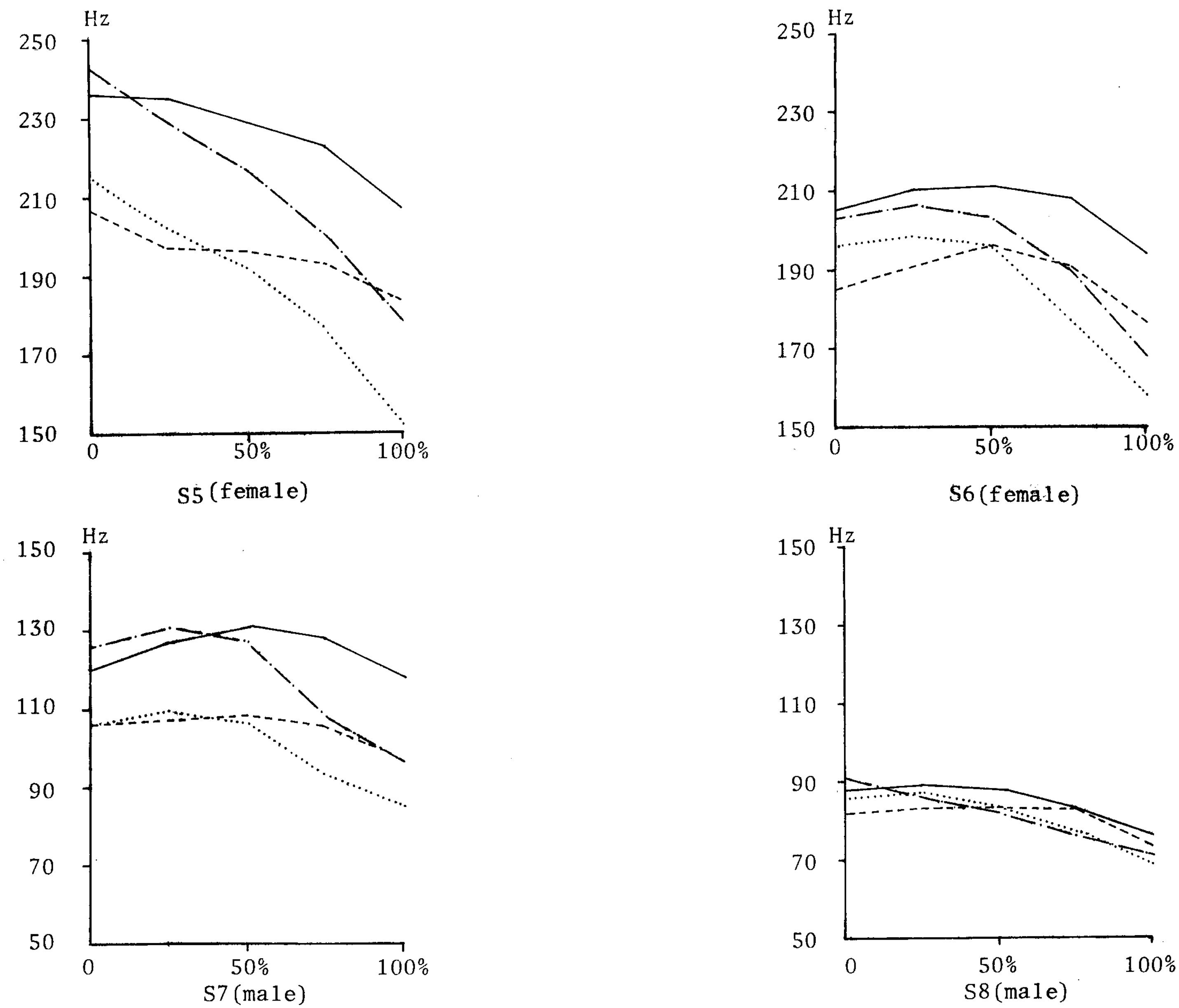


FIGURE 12 : Fo contours of tense and lax vowels in 4 syllable types of S5,S6,S7 and S8.

——— CVC̣   —·— CV(C)

----- CV̤C̤   ········ CV̤(C)

TABLE 2 : Mean Fo values in Hz of tense and lax vowels of 8 speakers measured at 5 points. ($H_o : \mu_1 = \mu_2$, $p < .0005$, df = 64, t = 3.460)

SYLLABLE TYPE : CVC̥ vs. CV̤C̥

	0%		25%		50%	
	tense	lax	tense	lax	tense	lax
x̄	142.73	126.09	144.45	126.72	144.06	127.50
SD	49.61	43.43	48.40	40.72	47.73	41.43
t	9.578		11.240		10.684	

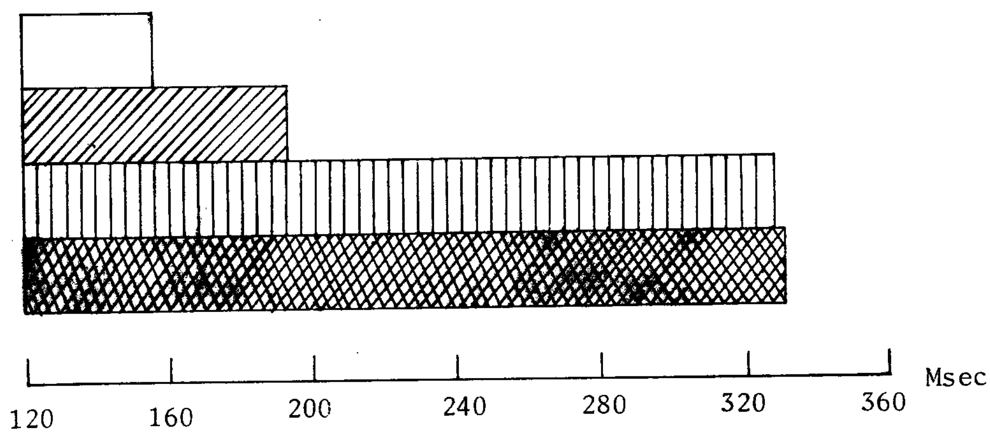
	75%		100%	
	tense	lax	tense	lax
x̄	139.22	123.20	127.27	112.42
SD	47.60	41.51	46.15	40.97
t	11.428		9.987	

SYLLABLE TYPE : CV(C) vs. CV̤(C)

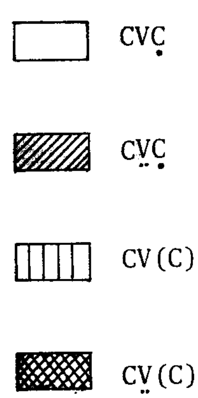
	0%		25%		50%	
	tense	lax	tense	lax	tense	lax
x̄	142.27	127.89	140.70	128.44	136.09	125.23
SD	49.76	46.78	47.29	42.86	45.99	41.72
t	9.278		7.662		6.160	

	75%		100%	
	tense	lax	tense	lax
x̄	123.20	114.84	109.30	101.25
SD	44.25	38.72	39.95	34.30
t	4.814		6.201	

VOWEL DURATION



<u>FIGURE 13</u> : Mean values of the duration of tense and lax vowels in 4 syllable types of 8 speakers.



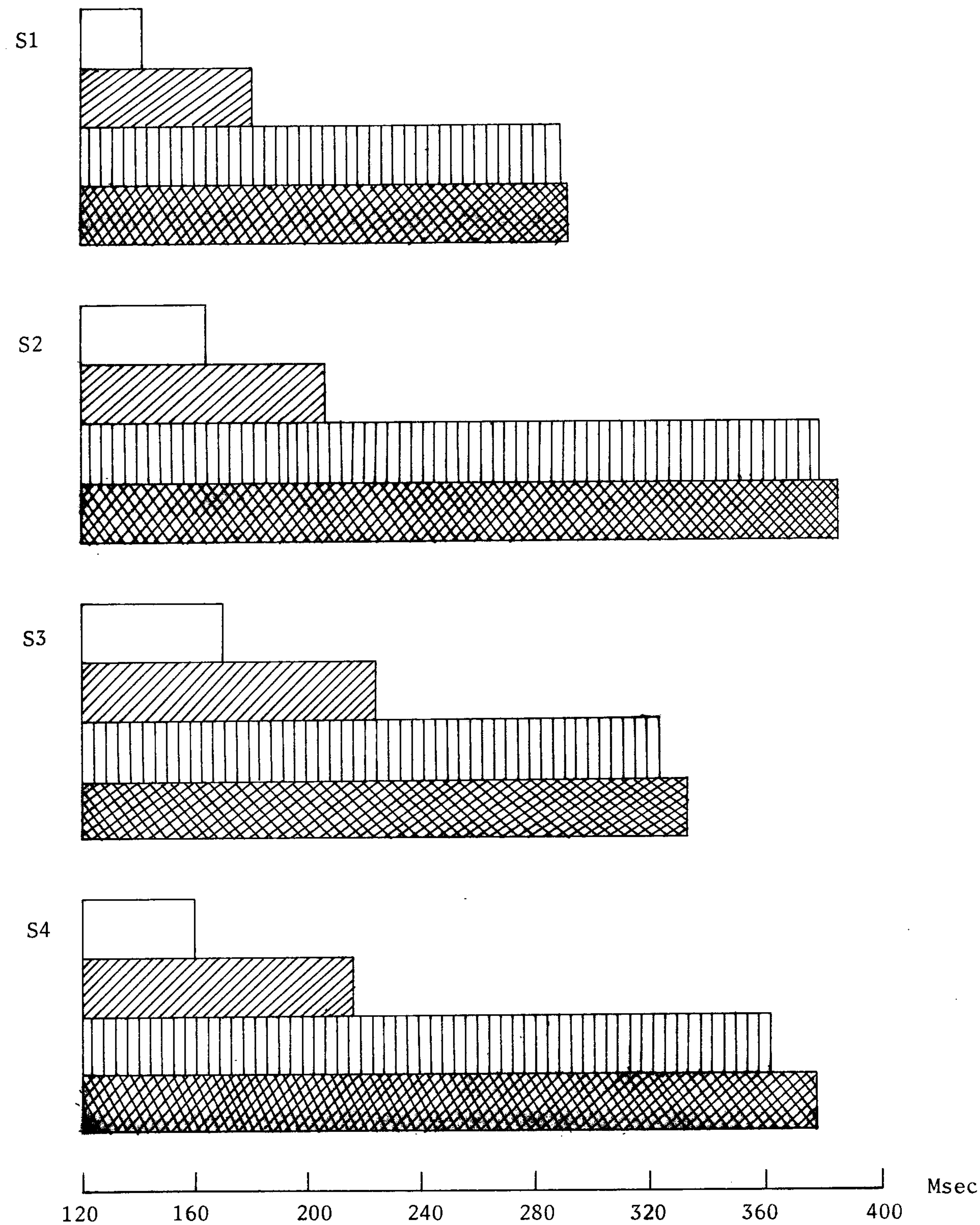


FIGURE 14 : Duration of tense and lax vowels in 4 syllable types of S1,S2,S3 and S4.

VOWEL DURATION

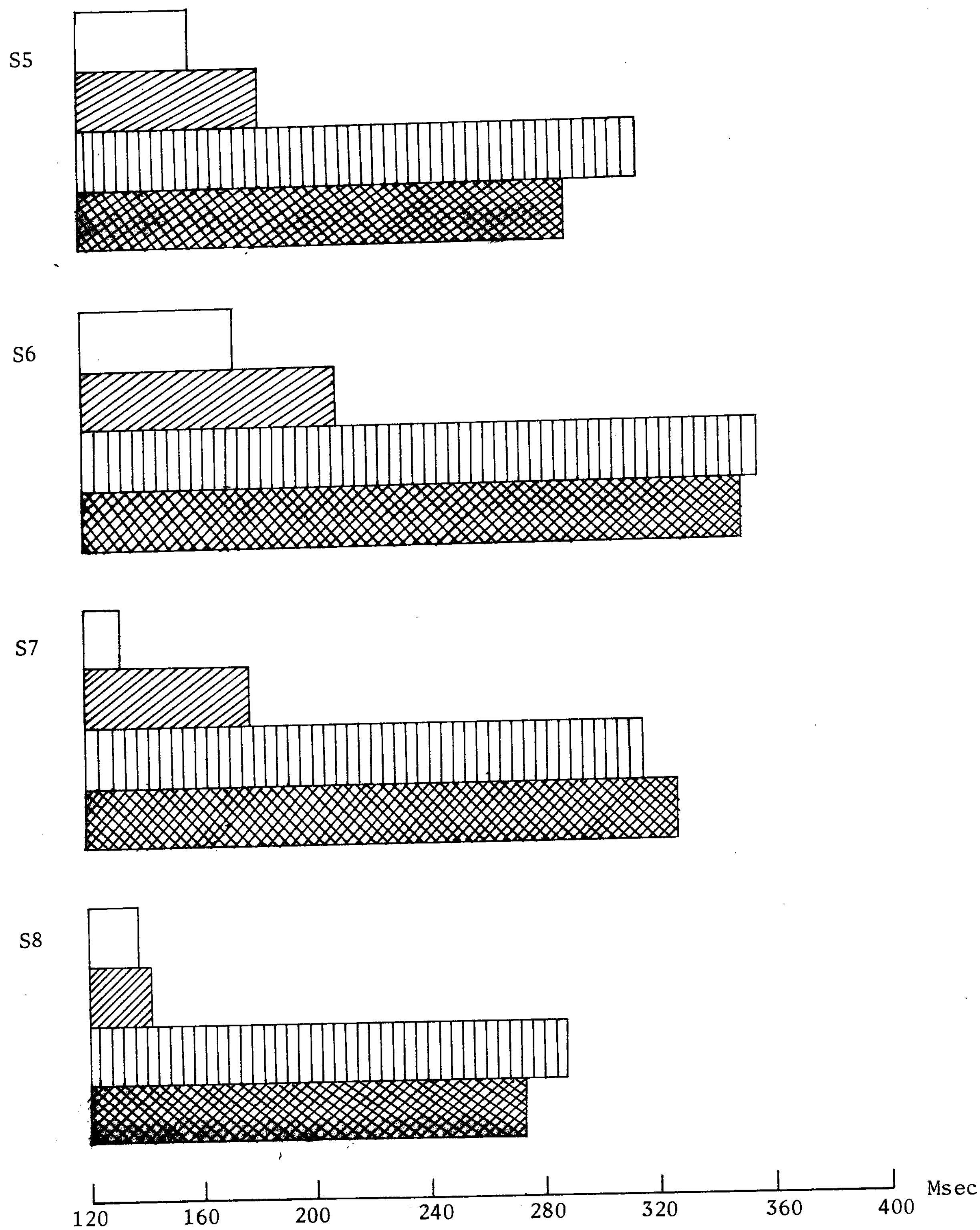


FIGURE 15 : Duration of tense and lax vowels in 4 syllable types of S5,S6,S7 and S8.

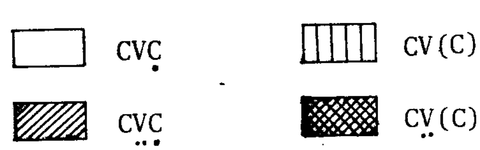


TABLE 3 : Duration of tense and lax vowels in CVC̥ and CV̤C̥ syllable types in msec.
($H_o : \mu_1 = \mu_2$, $p < .005$, df 7 (n-1), $t = 3.499$)

		First register vowel (Tense voice)	Second register vowel (Lax voice)
SPEAKER 1	$\bar{x}$	140.88	180.63
	SD	18.73	31.93
	t	4.352	
SPEAKER 2	$\bar{x}$	164.25	207.50
	SD	21.63	32.61
	t	6.023	
SPEAKER 3	$\bar{x}$	170.00	224.00
	SD	26.33	30.91
	t	5.881	
SPEAKER 4	$\bar{x}$	159.75	217.00
	SD	21.07	42.71
	t	4.217	
SPEAKER 5	$\bar{x}$	160.00	184.25
	SD	14.42	17.87
	t	4.476	
SPEAKER 6	$\bar{x}$	174.50	210.25
	SD	25.16	26.74
	t	6.300	
SPEAKER 7	$\bar{x}$	133.75	178.5
	SD	10.66	17.03
	t	5.741	
SPEAKER 8	$\bar{x}$	138.00	142.50
	SD	10.69	17.62
	t	0.590	

TABLE 4 : Duration of tense and lax vowels in CV(C) and CV̤(C) syllable types in msec. ($H_0 : \mu_1 = \mu_2$, $p < .005$, df = 7 (n-1) t = 3.499)

		First register vowel (Tense voice)	Second register vowel (Lax voice)
SPEAKER 1	$\bar{x}$	288.25	291.13
	SD	81.71	75.50
	t	0.406	
SPEAKER 2	$\bar{x}$	377.75	384.75
	SD	119.92	115.66
	t	0.478	
SPEAKER 3	$\bar{x}$	323.50	333.50
	SD	64.42	64.32
	t	1.104	
SPEAKER 4	$\bar{x}$	362.50	378.00
	SD	100.31	104.26
	t	2.026	
SPEAKER 5	$\bar{x}$	316.50	311.50
	SD	83.36	81.95
	t	0.802	
SPEAKER 6	$\bar{x}$	357.25	351.75
	SD	96.25	85.91
	t	0.633	
SPEAKER 7	$\bar{x}$	316.00	328.75
	SD	102.31	115.48
	t	1.233	
SPEAKER 8	$\bar{x}$	289.00	274.25
	SD	87.57	74.46
	t	1.391	

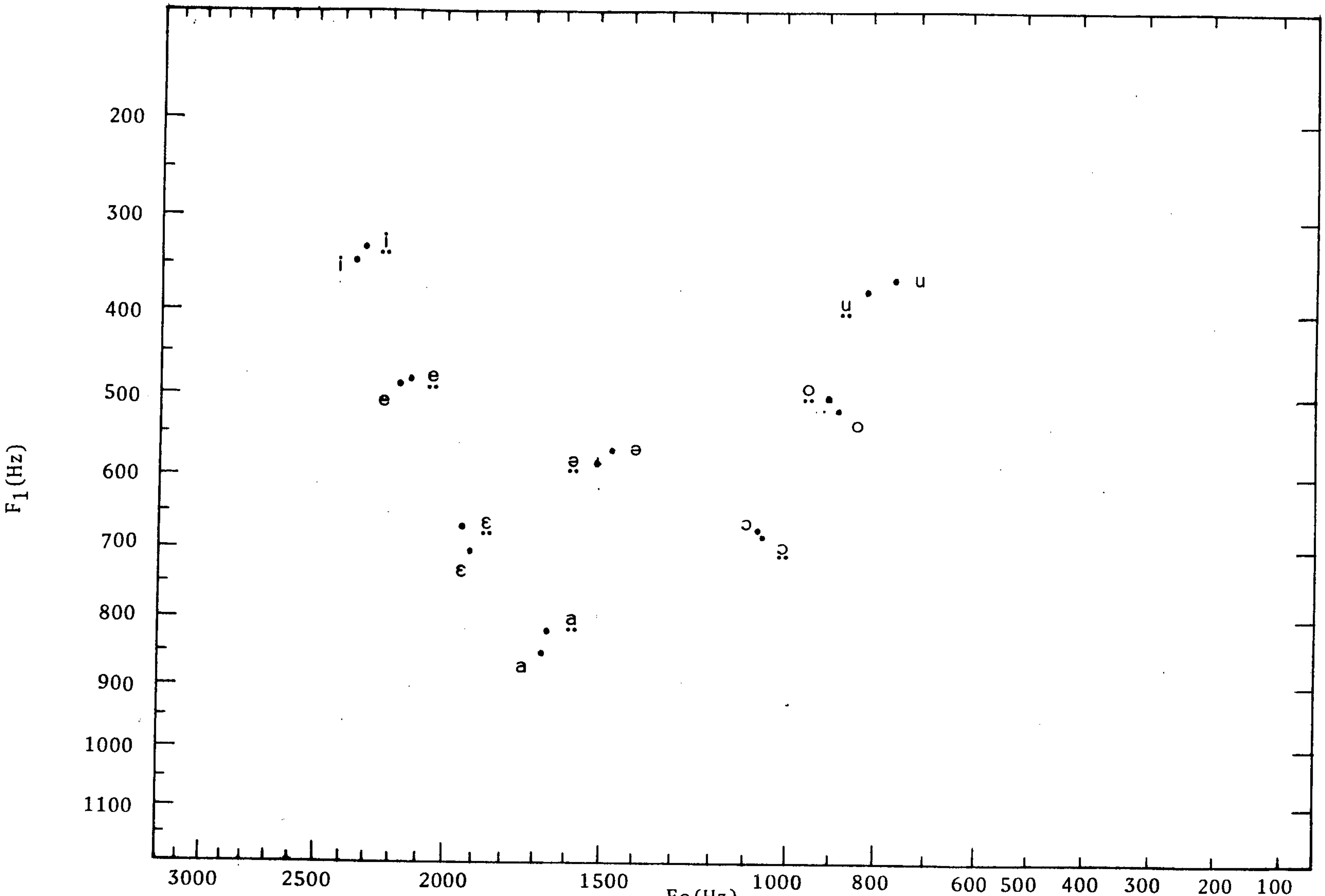


FIGURE 16 : Mean values of the formant frequencies (F1andF2) of tense and lax vowels in CVC̣ and CVC syllable types of 8 speakers.

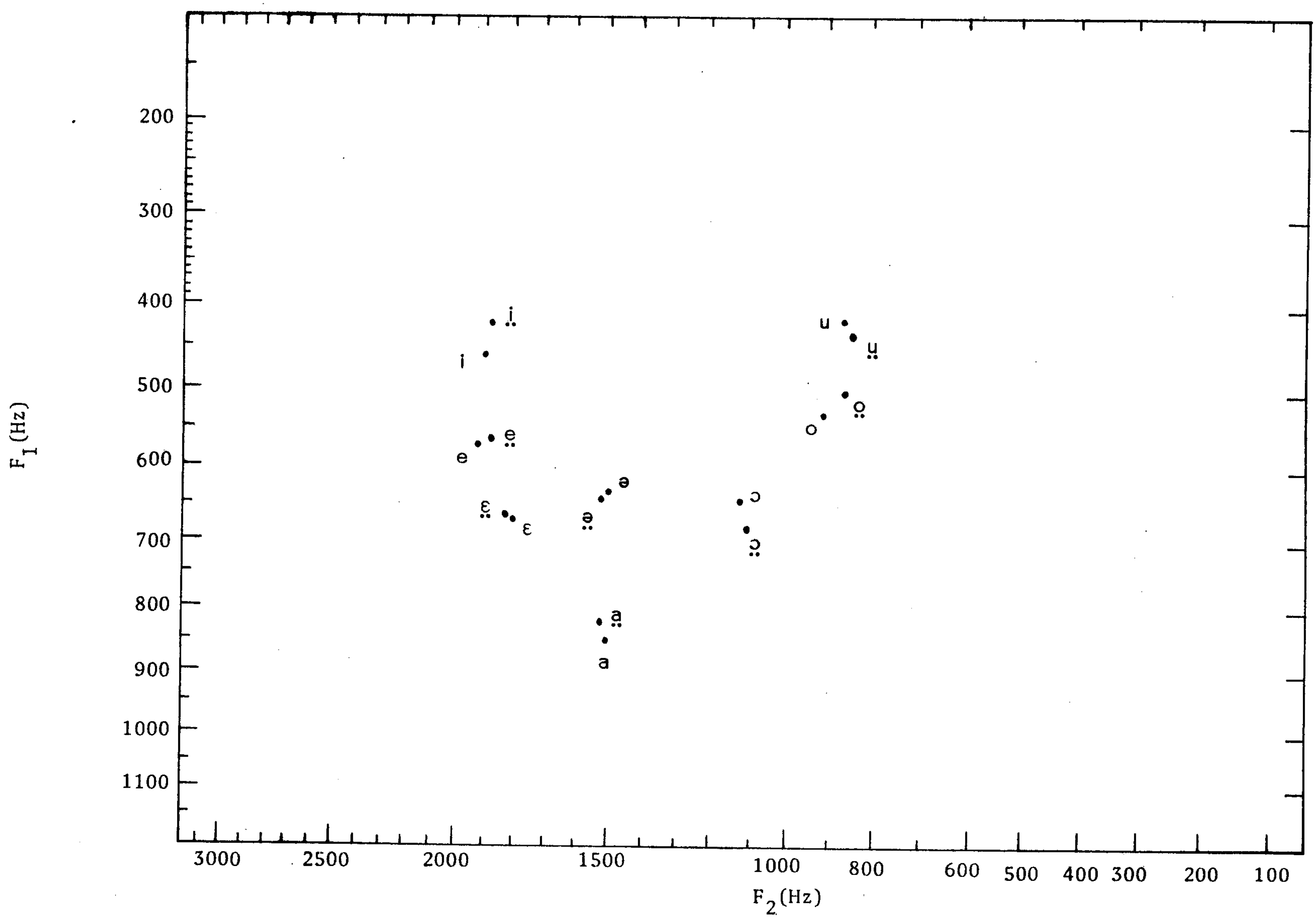


FIGURE 17 : Mean values of the formant frequencies (F_1 and F_2) of tense and lax vowels in CV(C) and CV̤(C) syllable types of 8 speakers.

and ɔ̤. The F_1 difference between the two registers is statistically significant ($p < .05$) only for i/i̤, o/o̤ (in CVÇ and CV̤Ç syllable type), ɛ/ɛ̤ and o/o̤ (in CV(C) and CV̤(C) syllable type). With respect to the F_2 difference between the two registers, it is significant for ɛ/ɛ̤ (in CVÇ and CV̤Ç syllable type), e/e̤ and u/ṳ (in CV(C) and CV̤(C) syllable type). See Tables 5 and 6. Thus, it is not true that first register vowels are peripheral and second register vowels are centralized, as said by Shorto (1967). My findings agree very closely with those of Lee's (1983).

Conclusion

In this study of the register distinction in Mon, four phonetic parameters were examined : distribution of spectral energy, fundamental frequency, vowel duration, and frequencies of F_1 and F_2. The results of the acoustical measurements indicated that the significant differences between the first register (tense) vowels and second register (lax) vowels are : 1) power spectra, 2) fundamental frequency, and 3) vowel duration (only in CVÇ and CV̤Ç syllable types). It can be concluded that vowel length in checked syllables and phonation-type differences are as significant as pitch differences. These findings agree very well with the intuition of Mon speakers. My informants sometimes criticized my pronunciation when trying to imitate them as too high vs. too low, too light vs. too heavy, or too short vs. too long. With respect to the register distinction in Mon, these may be regarded as good hints that Mon people listen to different phonetic cues. However, a definite answer can be found only by means of doing perception testing.

Perhaps, we could end the arguments by saying that in the previous studies of Mon registers, none of the investigators were completely right or completely wrong.

Acknowledgements

I would like to express my gratitude to Chulalongkorn University for granting me a leave of absence and to Peter Ladefoged who made my visit to UCLA Phonetics Laboratory possible. My thanks also go to Gérard Diffloth who kindly translated Blagden's article into English and provided me references on Mon language and linguistics.

TABLE 5 : F_1 and F_2 in Hz of tense and lax vowels in CVÇ and CV̤Ç syllable types of 8 speakers.

($H_o : \mu_1 = \mu_2$, $p < .05$, df = 7 (n-1), t = 1.895, * = significant difference.)

	/i/ F₁	/i̤/ F₁		/i/ F₂	/i̤/ F₂
$\bar{x}$	453.75	418.75	$\bar{x}$	1891.88	1866.88
SD	87.50	82.71	SD	114.33	172.15
t	*1.900		t	1.019	
	/e/ F₁	/e̤/ F₁		/e/ F₂	/e̤/ F₂
$\bar{x}$	573.75	561.88	$\bar{x}$	1903.75	1875.00
SD	87.00	96.58	SD	226.78	248.60
t	0.737		t	1.481	
	/ɛ/ F₁	/ɛ̤/ F₁		/ɛ/ F₂	/ɛ̤/ F₂
$\bar{x}$	666.25	657.50	$\bar{x}$	1792.50	1838.75
SD	54.76	66.06	SD	143.70	136.35
t	0.590		t	*2.290	
	/ə/ F₁	/ə̤/ F₁		/ə/ F₂	/ə̤/ F₂
$\bar{x}$	631.88	633.75	$\bar{x}$	1483.75	1516.88
SD	51.54	82.80	SD	152.22	165.81
t	0.087		t	1.475	
	/a/ F₁	/a̤/ F₁		/a/ F₂	/a̤/ F₂
$\bar{x}$	842.50	817.50	$\bar{x}$	1486.88	1513.13
SD	166.00	197.83	SD	132.99	162.68
t	0.863		t	0.882	
	/u/ F₁	/ṳ/ F₁		/u/ F₂	/ṳ/ F₂
$\bar{x}$	414.38	426.88	$\bar{x}$	865.63	840.00
SD	61.96	86.97	SD	111.59	107.30
t	0.352		t	0.518	
	/o/ F₁	/o̤/ F₁		/o/ F₂	/o̤/ F₂
$\bar{x}$	523.13	490.63	$\bar{x}$	910.00	863.75
SD	86.06	85.46	SD	102.12	84.46
t	*5.646		t	1.666	
	/ɔ/ F₁	/ɔ̤/ F₁		/ɔ/ F₂	/ɔ̤/ F₂
$\bar{x}$	646.88	676.25	$\bar{x}$	1106.25	1095.00
SD	76.16	100.42	SD	114.76	101.42
t	1.337		t	0.878	

TABLE 6 : F_1 and F_2 in Hz of tense and lax vowels in CV(C) and CV̤(C) syllable types of 8 speakers.
($H_o : \mu_1 = \mu_2$, $p < .05$, df = 7 (n-1), t = 1.895, * = significant difference.)

	/i/ F_1	/i̤/ F_1		/i/ F_2	/i̤/ F_2
$\bar{x}$	348.75	336.88	$\bar{x}$	2303.75	2267.50
SD	56.17	44.96	SD	245.93	247.60
t	1.221		t	1.163	

	/e/ F_1	/e̤/ F_1		/e/ F_2	/e̤/ F_2
$\bar{x}$	487.50	481.25	$\bar{x}$	2140.00	2090.00
SD	80.13	70.80	SD	247.50	263.76
t	0.413		t	*2.017	

	/ɛ/ F_1	/ɛ̤/ F_1		/ɛ/ F_2	/ɛ̤/ F_2
$\bar{x}$	703.13	667.50	$\bar{x}$	1898.13	1910.63
SD	77.41	72.90	SD	187.12	205.64
t	*3.037		t	0.560	

	/ə/ F_1	/ə̤/ F_1		/ə/ F_2	/ə̤/ F_2
$\bar{x}$	564.38	579.38	$\bar{x}$	1461.25	1484.38
SD	81.74	94.36	SD	115.32	110.95
t	1.381		t	0.868	

	/a/ F_1	/a̤/ F_1		/a/ F_2	/a̤/ F_2
$\bar{x}$	845.00	814.38	$\bar{x}$	1661.88	1653.13
SD	170.29	164.39	SD	63.80	91.45
t	1.060		t	0.228	

	/u/ F_1	/ṳ/ F_1		/u/ F_2	/ṳ/ F_2
$\bar{x}$	350.63	361.88	$\bar{x}$	766.25	821.25
SD	61.67	46.75	SD	89.75	69.32
t	0.551		t	*3.333	

	/o/ F_1	/o̤/ F_1		/o/ F_2	/o̤/ F_2
$\bar{x}$	508.75	488.75	$\bar{x}$	883.75	895.00
SD	69.17	58.42	SD	97.09	86.19
t	*1.965		t	0.492	

	/ɔ/ F_1	/ɔ̤/ F_1		/ɔ/ F_2	/ɔ̤/ F_2
$\bar{x}$	656.25	633.75	$\bar{x}$	1065.00	1059.38
SD	68.44	105.25	SD	116.62	132.57
t	0.655		t	0.262	

References

Blagden, C.O. (1910). Quelques notions sur la phonétique du Talian et son évolution historique. *Journal Asiatique*, 477-505.

Diffloth, Gérard. (1984). *The Dvāravatī-Old-Mon Language and Nyah-Kur*. Monic Studies, Vol. I. Bangkok: Chulalongkorn University Press.

______________. (1985). The registers of Mon vs. the spectrographist's tones. *UCLA Working Papers in Phonetics 60*, 55-58.

Fischer-Jørgensen, E. (1977). Phonetic analysis of breathy (murmured) vowels in Gujarati. *Annual Report of the Institute of Phonetics of the University of Copenhagen 2* : 105-156. (Reprinted from 1967).

Halliday, R. (1922). *A Mon-English Dictionary*. Bangkok: Siam Society.

Haswell, J.M. (1874). *Grammatical Notes and Vocabulary of the Peguan Language*, 2nd Ed. 1918. Rangoon: American Baptist Press.

Henderson, Eugénie J.A. (1952). The main features of Cambodian pronunciation. *Bulletin of the School of Oriental and African Studies 14.1* : 149-174.

Huffman, Franklin E. (1976). The register problem in fifteen Mon-Khmer languages, in *Austroasiatic Studies, Part I*, 25-37. Honolulu: University Press of Hawaii.

Kirk, Paul L. et al. (1984). Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics 59* : 102-113.

Ladefoged, Peter et al. (1987). Investigating phonation types in different languages. To appear in the *Proceedings* of the 5th Conference on Vocal Fold Physiology, Tokyo, January 19-23.

Lee, Thomas. (1983). An acoustical study of the register distinction in Mon. *UCLA Working Papers in Phonetics 57*, 79-96.

Maddieson, Ian and Ladefoged, Peter. (1985). Tense and lax in four minority languages of China. *Journal of Phonetics 13* : 433-454.

Pike, Kenneth L. (1948). *Tone Languages*. Ann Arbor: University of Michigan Press.

Sakamoto, Y. (1974). Notes on the phonology of Modern Mon (Pakkret dialect). *Journal of Asian and African Studies*. Tokyo. (In Japanese).

Shorto, Harry L. (1962). *A Dictionary of Modern Spoken Mon*. London: Oxford University Press.

Shorto, Harry L. (1966). Mon vowels systems : a problem in phonological statement, in *In Memory of J.R. Firth*, ed. by C.E. Bazell et al. London: Longmans.

______________. (1967). The register distinction in Mon-Khmer languages. *Wissenschaftliche Zeitschrift de Karl Marx Universität Leipzig 16*, 245-248.

Stevens, Kenneth N. (1981). Vibration modes in relation to model parameters, in *Vocal Fold Physiology*, ed. by K. Stevens and M. Hirano. Tokyo: University of Tokyo Press.

Index to *UCLA Working Papers in Phonetics*

Volumes 58-67

(For volumes 43-57, see *UCLA WPP* 57, April 1983)

Anderton, Alice. See Keating and Anderton 1987.

Antoñanzas-Barroso, Norma. See Javkin et al 1985, M. Jackson et al 1985a,b, Ladefoged and Antoñanzas-Barroso 1985.

Chan, Marjorie K.M. (1986). On the status of "basic" tones. *UCLA Working Papers in Phonetics* 63: 71-94.

Chan, Marjorie K. M., and Hongmo Ren (1986). Wuxi tone sandhi: from last to first syllable dominance. *UCLA Working Papers in Phonetics* 63: 48-70.

Dart, Sarah N. (1984). Testing an aerodynamic model with measured data from Korean. *UCLA Working Papers in Phonetics* 59: 1-17.

Dart, Sarah N. (1987). A bibliography of x-ray studies of speech. *UCLA Working Papers in Phonetics* 66: 1-97.

Diffloth, Gérard (1985). The registers of Mon vs. the spectrographist's tones. *UCLA Working Papers in Phonetics* 60: 55-8.

Disner, Sandra Ferrari (1983). Vowel Quality: the Relation between Universal and Language Specific Factors. *UCLA Working Papers in Phonetics* 58.

Dixit, R. Prakash (1987a). Mechanisms for voicing and aspiration: Hindi and other languages compared. *UCLA Working Papers in Phonetics* 67: 49-102.

Dixit, R. Prakash (1987b). In defense of the phonetic adequacy of the traditional term 'voiced aspirated'. *UCLA Working Papers in Phonetics* 67: 103-11.

Dolan, William B., and Yoko Mimori (1986). Rate-dependent variability in English and Japanese complex vowel F2 transitions. *UCLA Working Papers in Phonetics* 63: 125-53.

Emmorey, Karen (1985). The transition from early to late babbling. *UCLA Working Papers in Phonetics* 62: 61-9.

Emmorey, Karen (1986). Morphemic structure iin the mental lexicon. *UCLA Working Papers in Phonetics* 63: 110-24.

Emmorey, Karen, Diana Van Lancker and Jody Kreiman (1984). Recognition of famous voices given vowels, words, and two-second texts. *UCLA Working Papers in Phonetics* 59: 120-24.

Emmorey, Karen. See also Maddieson and Emmorey 1984, 1985, Van Lancker et al 1984.

Engstrand, Olle (1986). Durational correlates of quantity and sentence stress: a cross-linguistic study of Swedish, Finnish and Czech. *UCLA Working Papers in Phonetics* 63: 1-25.

Hess, Susan A. See Maddieson and Hess 1986, 1987.

Huffman, Marie (1985a). Measures of phonation type in Hmong. *UCLA Working Papers in Phonetics* 61: 1-25.

Huffman, Marie K. (1985b). Place coordination in child speech. *UCLA Working Papers in Phonetics* 62: 70-76.

Huffman, Marie K. (1986). Patterns of coarticulation in English. *UCLA Working Papers in Phonetics* 63: 26-47.

Huffman, Marie. See also Keating et al 1984, M. Jackson et al 1985a,b, Ladefoged, Maddieson, Jackson and Huffman 1987.

Jackson, Ellen. See Keating et al 1984.

Jackson, Michel, Peter Ladefoged, Marie Huffman and Norma Antoñanzas-Barroso (1985a). Automated measures of spectral tilt. *UCLA Working Papers in Phonetics* 62: 77-88.

Jackson, Michel, Peter Ladefoged, Marie Huffman and Norma Antoñanzas-Barroso (1985b). Measures of spectral tilt. *UCLA Working Papers in Phonetics* 61: 72-8.

Jackson, Michel. See also Traill and Jackson 1987, Ladefoged, Maddieson, Jackson and Huffman 1987, Ladefoged, Maddieson and Jackson 1987.

Javkin, Hector, Norma Antoñanzas-Barroso and Ian Maddieson (1985). Digital inverse filtering for linguistic research. *UCLA Working Papers in Phonetics* 60: 87-100.

Keating, Patricia A. (1984a). Aerodynamic modeling at UCLA. *UCLA Working Papers in Phonetics* 59: 18-28.

Keating, Patricia A. (1984b). Physiological effects on stop consonant voicing. *UCLA Working Papers in Phonetics* 59: 29-34.

Keating, Patricia A. (1984c). Universal phonetics and the organization of grammars. *UCLA Working Papers in Phonetics* 59: 35-49.

Keating, Patricia A. (1985a). Linguistic and nonlinguistic effects on the perception of vowel duration. *UCLA Working Papers in Phonetics* 60: 20-39.

Keating, Patricia A. (1985b). CV phonology, experimental phonetics, and coarticulation. *UCLA Working Papers in Phonetics* 62: 1-13.

Keating, Patricia A. (1985c). The phonology-phonetics interfact. *UCLA Working Papers in Phonetics* 62: 14-33.

Keating, Patricia (1987). A survey of phonological features. *UCLA Working Papers in Phonetics* 66: 124-50.

Keating, Patricia and Alice Anderton (1987). Survey of phonetics computers in North American laboratories. *UCLA Working Papers in Phonetics* 66: 104-23.

Keating, Patricia A., Marie Huffman and Ellen Jackson (1984). Vowel allophones and the vowel-formant phonetic space. *UCLA Working Papers in Phonetics* 59: 60-61.

Keating, Patricia A. See also Westbury and Keating 1985.

Kempler, Daniel. See Van Lancker and Kempler 1986.

Kirk, Paul L., Peter Ladefoged and Jenny Ladefoged (1984). Using a spectrograph for measures of phonation types in a natural language. *UCLA Working Papers in Phonetics* 59: 102-13.

Kreiman, Jody, and Diana Van Lancker (1986). Hemispheric specialization of voice recognition: evidence from dichotic listening. *UCLA Working Papers in Phonetics* 63: 95-102.

Kreiman, Jody, and George Papçun (1985). Voice discrimination by two listener populations. *UCLA Working Papers in Phonetics* 61: 45-50.

Kreiman, Jody. See also Emmorey et al 1984, Van Lancker and Kreiman 1985, Van Lancker et al 1984.

Ladefoged, Jenny. See Kirk et al 1984.

Ladefoged, Peter (1985a). Macintosh usage for linguists. *UCLA Working Papers in Phonetics* 60: 84-6.

Ladefoged, Peter (1985b). Redefining the scope of phonology. *UCLA Working Papers in Phonetics* 60: 101-8.

Ladefoged, Peter (1985c). Macintosh models and plots for phoneticians. *UCLA Working Papers in Phonetics* 61: 65-71.

Ladefoged, Peter (1987). Macintosh computer programs for phoneticians. *UCLA Working Papers in Phonetics* 66: 98-103.

Ladefoged, Peter, and Norma Antoñanzas-Barroso (1985). Computer measures of breathy phonation. *UCLA Working Papers in Phonetics* 61: 79-86.

Ladefoged, Peter, and Ian Maddieson (1986). Some of the Sounds of the World's Languages. *UCLA Working Papers in Phonetics* 64.

Ladefoged, Peter, Ian Maddieson, Michel Jackson and Marie Huffman (1987). Characteristics of the voice source. *UCLA Working Papers in Phonetics* 67: 119-25.

Ladefoged, Peter, Ian Maddieson and Michel Jackson (1987). Investigating phonation types in different languages. *UCLA Working Papers in Phonetics* 67: 126-31.

Ladefoged, Peter and Zongji Wu (1984). Places of articulation: an investigation of Pekingese fricatives and affricates. *UCLA Working Papers in Phonetics* 59: 62-76.

Ladefoged, Peter. See also M. Jackson et al 1985a,b, Kirk et al 1984, Maddieson and Ladefoged 1985.

Lindau-Webb, Mona (1985a). Hausa vowels and diphthongs. *UCLA Working Papers in Phonetics* 60: 40-54.

Lindau, Mona (1985b). Testing a model of intonation in a tone language. *UCLA*

Working Papers in Phonetics 61: 26-39.
Lindau, Mona, Kjell Norlin and Jan-Olaf Svantesson (1985). Cross-linguistic differences in diphthongs. *UCLA Working Papers in Phonetics* 61: 40-44.
Maddieson, Ian (1984). Phonetic cues to syllabification. *UCLA Working Papers in Phonetics* 59: 85-101.
Maddieson, Ian (1985). Borrowed sounds. *UCLA Working Papers in Phonetics* 61: 51-64.
Maddieson, Ian and Karen Emmorey (1984). Is there a valid distinction between voiceless lateral approximants and fricatives? *UCLA Working Papers in Phonetics* 59: 77-84.
Maddieson, Ian, and Karen Emmorey (1985). Cross-linguistic issues in the relationship between semivowels and vowels. *UCLA Working Papers in Phonetics* 61: 87-98.
Maddieson, Ian, and Susan A. Hess (1986). "Tense" and "lax" revisited: more on phonation type and pitch in minority languages in China. *UCLA Working Papers in Phonetics* 63: 103-9.
Maddieson, Ian, and Susan Hess (1987). The effect on F_0 of the linguistic use of phonation type. *UCLA Working Papers in Phonetics* 67: 112-8.
Maddieson, Ian, and Peter Ladefoged (1985). "Tense" and "lax" in four minority languages of China. *UCLA Working Papers in Phonetics* 60: 59-83.
Maddieson, Ian. See also Javkin et al 1985, Ladefoged and Maddieson 1986, Ladefoged, Maddieson, Jackson and Huffman 1987, Ladefoged, Maddieson and Jackson 1987.
Mimori, Yoko. See Dolan and Mimori 1986.
Norlin, Kjell. See Lindau et al 1985.
Papçun, George. See Kreiman and Papçun 1985.
Ren Hongmo (1985). Linguistically conditioned duration rules in a timing model for Chinese. *UCLA Working Papers in Phonetics* 62: 34-49.
Ren, Hongmo (1986). On the Acoustic Structure of Diphtongal Syllables. *UCLA Working Papers in Phonetics* 65.
Ren, Hongmo. See also Chan and Ren 1986.
Svantesson, Jan-Olaf. See Lindau et al 1985.
Thongkum, Therapan L. (1987a). Phonation types in Mon-Khmer languages. *UCLA Working Papers in Phonetics* 67: 20-48.
Thongkum, Therapan L. (1987b). Another look at the register distinction in Mon. *UCLA Working Papers in Phonetics* 67: 132-66.
Traill, Anthony, and Michel Jackson (1987). Speaker variation and phonation types in Tsonga nasals. *UCLA Working Papers in Phonetics* 67: 1-19.

Van Lancker, Diana, and Daniel Kempler (1986). Comprehension of familiar phrases by left- but not right-hemisphere damaged patients. *UCLA Working Papers in Phonetics* 63: 154-65.

Van Lancker, Diana, and Jody Kreiman (1985). Unfamiliar voice discrimination and familiar voice recognition are independent and unordered abilities. *UCLA Working Papers in Phonetics* 62: 50-60.

Van Lancker, Diana, Jody Kreiman and Karen Emmorey (1984). Recognition of famous voices forwards and backwards. *UCLA Working Papers in Phonetics* 59: 114-9.

Van Lancker, Diana. See also Emmorey et al 1984, Kreiman and Van Lancker 1986.

Westbury, John R. and Patricia A. Keating (1985). On the naturalness of stop consonant voicing. *UCLA Working Papers in Phonetics* 60: 1-19.

Wu, Zongji. See Ladefoged and Wu 1984.