PACIFIC LINGUISTICS

Series A - No. 62

PAPERS IN SOUTH-EAST ASIAN LINGUISTICS, No. 8

TONATION

edited by

David Bradley

Department of Linguistics

Research School of Pacific studies

THE AUSTRALIAN NATIONAL UNIVERSITY

All correspondence concerning PACIFIC LINGUISTICS, including orders and subscriptions, should be addressed to:

The Secretary
PACIFIC LINGUISTICS
Department of Linguistics
Research School of Pacific Studies
The Australian National University
P.O. Box 4, Canberra, A.C.T. 2600
Australia.

First Published 1982

The editors are indebted to the Australian National University for assistance in the production of this series.

This publication was made possible by an initial grant from the Hunter Douglas Fund.

National Library of Australia Card Number and ISBN 0 85883 256 9

TABLE OF CONTENTS

INTRODUCTION

This volume contains six studies by researchers in Australia on the suprasegmental systems of various mainland Asian languages - one on a non-Mandarin variety of Chinese; one on Vietnamese dialects; two on dialects of Thai; and two on Burmese.

Three of the contributions contain extensive new acoustic data: on Burmese, on the three major dialects of Vietnamese, and on a Wu dialect of Chinese. The authors have not made the usual, incorrect assumption that 'tone' systems in this area have fundamental frequency, perceived as pitch and contour, as the only parameter in their realisation. Rather, they have also considered such other parameters as duration, intensity, vowel quality, and voice quality. The 'tones' are regarded as a system, each member of which has a complex of parameters involved in its production. While fundamental frequency is a very prominent characteristic, other characteristics may be just as much a part of the realisation of the 'tone'.

It is very often the case that one or more 'tones' in languages of this area have marked voice-quality features: some degree of creaky phonation or breathy phonation. One of the Burmese 'tones' is creaky; in dialects of Vietnamese, up to two 'tones' have creaky phonation, and one may have breathy phonation. Many further instances could be cited; in Bangkok Thai there is some creakiness in the high tone, for example.

The importance of phonation in the suprasegmental systems of this area is recognised in the title of this book; *Tonation*. Of course, this is not to imply that other characteristics can or should be ignored; rather, they should be measured. Even such features as nasalisation may be consistent 'tonation' parameters, as in one tone of Southern Thai. Moreover, in many Austro-Asiatic languages, phonation is regarded as the main feature in a register system.

Two of the articles here are looking at the diachronic development of 'tonation' systems: one within Southern Thai, and the other in Burmese. Diller's paper looks at the process of tonal split and subsequent incipient tonal merger, in a case which also shows the effects of a national standard language on one of its dialects. This is a particularly interesting case, as the change was triggered by changes in segmental phonology: in this case, syllable-final consonants - and is thus parallel to many similar changes, postulated on the basis of reconstruction, in Sino-Tibetan, Austro-Thai, and elsewhere. My paper considers acoustic data as a basis for the analysis of synchronic 'tonation' in Burmese; and suggests that its development is not unconnected with the history of contact and assimilation with speakers of Mon, an Austro-Asiatic language with a register system.

The paper by U Thein Tun is unusually thorough in its investigation of duration and vowel quality as 'tonation' parameters in Burmese, and also provides extensive fundamental frequency data. In his study, Vũ Thanh Phương presents a survey of Vietnamese dialect suprasegmentals, clarifying and highlighting their differences in a way which separate studies of each dialect would not. Rose's contribution is painstakingly excellent in its fundamental frequency, intensity, and duration measurements of a Wu dialect. In this dialect, some speakers have a highly unusual type of voicing in some tones, but those data have been left for a future study.

All three acoustic investigations recognise the importance of phonation; however it is difficult to measure in purely acoustic terms. To get a clearer picture, it would be necessary to observe the articulation involved - in particular, the larynx. Current and future studies by the same authors will doubtless clarify the articulatory questions; in addition, some psychoacoustic studies to elucidate auditory data are now in progress.

The final paper by the Gandours quantifies the relative lexical frequency of occurrence of the tones in standard Thai; very little statistical work on this kind has been done previously, and the present findings may have diachronic as well as synchronic implications.

Some of the references in the articles in this volume appear in small periodicals which are important in the Southeast Asian field, but not yet well-known otherwise. For example, *Linguistics of the Tibeto-Burman Area*, now edited by Graham Thurgood at Fresno, California, has been appearing since 1974, and is now up to volume 6; this is abbreviated *LTBA*. Another is Mantaro Hashimoto's *Computational Analyses of Asian and African Languages* or *CAAAL*, a monograph series appearing since 1975 under the aegis of the Japan National

Inter-University Research Institute on Asian and African Languages and Cultures, which also publishes *Monumenta Serindica*, a series of longer monographs. *CAAAL* now has 18 numbers, and *MS* is up to eight.

David Bradley

ACOUSTIC CHARACTERISTICS OF THE SHANGHAI-ZHENHAI SYLLABLE TYPES

Philip John Rose

0. INTRODUCTION

This paper presents a quantified description of the main acoustic characteristics of the six contrasting syllable-types in the Shanghai-Zhenhai variety of Chinese (Sh-Zh).[1] Zhenhai is a rural county north-east of the municipality of Ningpo in Zhejiang province, and Sh-Zh can be described as a common type of Zhenhai dialect which exhibits to a greater or lesser extent lexical and phonological influence from the neighbouring prestigious dialect of Shanghai.[2]

Sh-Zh belongs to the Wu dialect group of Chinese, and has typical Wu phonology. This includes complex tone sandhi, restriction of syllable-final consonants to [ŋ] and [ʔ], and respective tripartite and bipartite division, according to manner, of syllable-initial occlusives and fricatives (Yuan 1960:59; Chao 1967).

The acoustic parameters investigated in this study were fundamental frequency (F_o), duration, and oral amplitude (Ao). Fundamental frequency is the acoustic correlate of the rate of vibration of the vocal cords, and is usually assumed to correspond to the perceptual dimension of pitch. Duration, the time dimension of the acoustic signal, corresponds to perceived length, and is determined physiologically by the relative timing of articulatory events. Oral amplitude refers to the integrated time-varying sound pressure transduced at a distance from the speaker's lips. It is important to realise that the oral amplitude thus measured is a function of three factors which vary in an essentially independent manner, and which are consequently difficult to isolate. The first of these, the time-varying amplitude of the glottal source, occurs extrinsically as the result of articulatory gestures which affect the sub-glottal pressure (P_s), such as differing

Rose, P.J. "Acoustic Characteristics of the Shanghai-Zhenhai Syllable Types". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation.* A-62:1-54. Pacific Linguistics, The Australian National University, 1982. DOI:10.15144/PL-A62.1

respiratory effort or adjustments in the open quotient of the glottal cycle (Zemlin 1968:198-204). Thus, other things being equal, a change in P_s will be reflected by a change in Ao.

The second factor, the transfer function of the supralaryngeal vocal tract, modifies the amplitude of the glottal source by attenuating the transfer of energy at certain frequencies and passing maximum energy at others (Lieberman 1977:31-36). The transfer function is determined by the shape of the supralaryngeal vocal tract, changes in which - as for example when different vowels are articulated - will then differentially modify the amplitude of the glottal source. Thus, other things being equal, a change in the supralaryngeal vocal tract shape will be reflected by a change in Ao. This is the reason for intrinsic vowel amplitude (Lehiste and Peterson 1959:429).

In addition to the above two factors, changes in Ao will occur as the result of the interaction of harmonic and formant frequencies (House 1959). Thus, if P_s and transfer function are held constant, changes in F_o will be reflected in changes in Ao.

From the articulatory point of view, then, the oral amplitude is of prosodic interest only in so far as it reflects those amplitude features which the speaker is extrinsically controlling, that is, the time-varying amplitude of the glottal source. It is therefore necessary to ensure that fluctuations in Ao due to intrinsic vowel amplitude and the interaction of harmonic and formant frequencies are kept to a minimum. I have done this by 1) selecting examples spoken with as near steady state supralaryngeal configurations as possible, monophthongality being adjudged by reference to formant trajectories on wide band spectrograms, and 2) analysing about equal numbers of open and close vowels in each particular sample. Although it would have been possible to apply a correction factor to the oral amplitude to eliminate the effects of the interaction between formant and harmonic frequencies, I have not done so, because the actual F_o range used by the informant in this study was too narrow to have caused appreciable changes in Ao. The oral amplitude data presented below can therefore be taken to give a reasonable approximation of the time-varying amplitude of the glottal source.

Amplitude is usually assumed to correlate with perceived loudness. In the perception of speech, however, there is evidence that listeners base loudness judgements on features more directly related to the P_s and glottal source amplitude than the oral amplitude (Ladefoged 1967: 35-41).

1. PREVIOUS STUDIES

There have been very few acoustical studies on Wu tones, and there are, to my knowledge, none which pay attention to all three parameters of F_o, Ao, and duration.

Liu (1925) analysed the F_o and duration of two Jiangyin tones kymographically, and there are also some kymographic records of various Ningpo utterances in Tchen (1938). Sokolov (1965) presented data on the F_o and duration of Shanghai citation monosyllables, and the F_o and duration of some Shanghai monosyllabic and polysyllabic utterances have been investigated by Zee and Maddieson (1979). A comparison of their results with those presented below can give an idea of the degree of similarity in F_o shapes between Shanghai and Sh-Zh/Zhenhai dialect.

The nearest site to Zhenhai for which descriptions of citation tone pitch values are available is Ningpo town. Ningpo town was one of the sites visited by Y.R. Chao in 1927 when collecting material for his pioneering monograph on the Wu dialects (Chao 1928). Chao's pitch descriptions compare visually very well with the F_o data obtained in this study.

The pitch of the citation tones of Ningpo town has also been recorded in a recent description (Shi 1979) with the five-point system devised by Chao (1930). The pitch values given, however, do not agree well with those in Chao (1928) or the present study, and the transcriptions must, therefore, be treated with caution.

2. SYLLABLE-TYPES

In Shanghai-Zhenhai, any citation monosyllable or monosyllabic word belongs unambiguously to one of six contrasting types. Below are listed examples of these six syllable-types, together with their main auditory characteristics.

1. The pitch of type 1 syllables falls from high in the speaker's pitch-range to low, with a short initial level component. (In terms of Chao's (1930) five-point pitch scale, its value would be 42 or 442.) The fall often starts higher, and sounds more abrupt, in syllables with nasal codas. The first half of the Final[3] seems the loudest, and the length of the Final is shorter than average for the syllable types. Voice quality is normal, and there is a gradual offset to voicing. Examples[4] are:

[sɿ][5] : *'poem'*; [tĩ] : *'shop'*; [ʔlœʏ] : *'to pick'*, e.g. *'the nose'*; [tɕɪ̃ŋ] : *'needle'*.

2. Syllables of type 2 have a concave pitch contour, starting in

mid pitch-range, falling, and then rising just into the upper third of the pitch-range, where voicing is terminated by a glottal-stop. The pitch shape would be best transcribed by 324 in the Chao notation: this adequately expresses the relative onset and offset heights, and concave shape, but implies too low a dip in pitch. Other possibilities could therefore be 334 or 434.

Loudness is concentrated at the beginning and end of the Final, with the end usually the loudest. Length is above average, and the voice often has a somewhat tense quality. Some examples are:

[tɕiʔ]	: *'chicken';*	[sɿʔ]	: *'water';*
[kɑ̆ʔ]	: *'river';*	[tə̆ŋʔ]	: *'to wait';*
[tɕɪ̆ŋʔ]	: *'well (of water)'.*		

3. Type 3 syllables have striking auditory characteristics, resulting primarily from the combination of a convex pitch contour which starts in the low part of the pitch-range - 232 or 242 - with loudness energetically concentrated on the first half of the syllable. In addition, voice quality over approximately the first third of the Final is whispery [̳]. There is gradual offset to voicing. Syllables of type 3 have average length. Examples are:

[nɛ̳ɩ]	: *'south';*	[mɑ̳̃]	: *'busy';*
[mə̳̆ŋ]	: *'door';*	[d̥ʑ̊ɪ̳̆ŋ]	: *'town';*
[ʑ̊ɿ̳]	: *'to be';*	[d̥ɪ̳̃]	: *'sweet'.*

4. Syllables of type 4 have a concave pitch contour, starting in the low pitch-range, falling to the bottom of the range and then rising again to mid pitch-range, where voicing is optionally terminated in a glottal-stop. The pitch shape would be best transcribed by 213 or 214. Loudness is concentrated at both beginning and end of the Final, but the end sounds considerably louder, with a typical burst-like quality. Type 4 syllables are the longest, and have whispery voice over the first half to two-thirds of their length.

[ʑ̊ɿ̳]	: *'word';*	[d̥ɪ̳ʔ]	: *'ground';*
[d̥ʑ̊ɪ̳̆ŋ]	: *'near';*	[mɑ̳̃ʔ]	: *'dream'*

5. Syllables of type 5 give an auditory impression of very short, high level pitch, which ends in a glottal-stop - 5̲. There is a single burst of loudness, and voice quality sounds somewhat tense. Type 5 syllables sound the shortest of the six types. Examples are:

[fə̆ʔ]	: *'law';*	[tɪ̆ʔ]	: *'target';*
[tɕjə̆ʔ]	: *'foot';*	[tɕhɥɜ̆ʔ]	: *'to eat'.*

6. Syllables of type 6 have a pitch contour which rises abruptly from low in the speaker's pitch-range to upper mid, where it ends in a glottal-stop - 2̲4̲ or 2̲3̲. There is a single burst of loudness. They

seem slightly longer than type 4 syllables, but group together perceptually with them against types 1 to 4 as short vs. long syllables. Voice quality is whispery over the first part of the Final. Examples are:

[ɣ̊ɐ̤̆ʔ] : *'buddha'*; [d̥ʑ̊ɥɔ̤̆ʔ] : *'office'*;
[d̥ʑ̊ɛ̤̆ʔ] : *'straight'*; [ɲɥɔ̤̆ʔ] : *'jade'*.

3. SYLLABLE STRUCTURE

The only relevant details of segmental structure to be mentioned concern the syllable-initial consonants - the 'Initials' - and certain co-occurrence restrictions.

Obstruent initials are divided into two groups on the basis of their co-occurrence with the syllable-types: aspirated and fortis unaspirated voiceless obstruents, i.e. [pʰ p f] etc. - occur only with syllable-types 1, 2, and 5. The sounds transcribed above with IPA voiced symbols plus the voiceless diacritic, i.e. [d̥, ɣ̊, g̊, d̥ʑ̊] etc., occur only with syllable-types 3, 4, and 6. This second group corresponds to what is conventionally termed the 'breathy voiced' Initials in the Wu dialects. For Sh-Zh, and the Zhenhai dialect, at least, this is a misnomer, since the sounds are voiceless, and the occlusives have a voiceless whispered release which is auditorily clearly distinguishable from normal aspiration. I retain the label 'lenes' to describe this group, together with the appropriate IPA transcription, on the basis of their auditory quality vis a vis the decidedly fortis articulation of the voiceless unaspirated series with syllable-types 1, 2 and 5.

There are very few syllables of type 1, 2 or 5 with sonorant Initials in the lexicon, and syllables of type 5 and 6 cannot occur with a syllable-final nasal.

4. PROCEDURE (GENERAL)

The syllables to be analysed were selected from a recording of a list of some 300 Chinese characters made by a young male native speaker. The list was recorded at one sitting in the phonetics laboratory of the University of Manchester's Department of Linguistics. I used a Ferrograph reel-to-reel tape recorder and Ampex 611 professional 1.5 ml acetate tape, and recorded at a speed of 7½ ips. The recording level was set manually. There was no ambient noise, and the recordings are of high quality.

The informant was 25 years old at the time of recording (1973). He was born in 1948 in Shanghai, the youngest son of Zhenhai-speaking parents who came from Qingshuipu, a small town in Zhenhai county.

His family left Shanghai one year after his birth for Hong Kong, where he was brought up and educated. He went to England in 1968 for further education. He speaks fluent English, almost fluent Cantonese, and Standard Chinese with a Sh-Zh accent.

The informant's speech - especially the phonetic values of syllable-type contours, tone sandhi rules and general morphophonemic shapes which have been described in Rose (1974) - is representative of a rather conservative Zhenhai speech form, from which the younger generation of Zhenhai speakers in China already constitutes (in, for example, loss of many nasalised phonemes) a considerable departure.

As far as the composition of the list was concerned, all six types of syllable were represented, and occurred randomly. I selected examples of syllables to be analysed from the list in the following way. The original corpus was first divided into six groups corresponding to the six syllable-types. From each group all syllables were excluded with either aspirated initial consonant, e.g. [ts^hɑ, p^hɪ], or complex vowel finals, e.g. [weɪ, ɪju, jə, jã]. This was to eliminate from the sample any statistical noise coused by the possible influence of aspiration (Hombert 1978) and changing vowel quality (p.2) on the acoustic parameters investigated. Next, each of the six syllable-type groups was further classified, where possible, according to the absence or presence of a sonorant Initial, and the absence or presence of a syllable-final velar nasal. This was because previous studies (Howie 1976; Sauvain 1977; Fok 1974; Kratochvil 1971, 1977) have either demonstrated or mentioned that such factors can substantially effect the F_o, Ao, and duration of syllable-types in Chinese.

This process yielded 15 sub-groups from the original six groups. Each sub-group was then further sifted, if numbers permitted, to ensure a reasonable balance of open and close vowels. This was to minimise the effect of vowel 'height' on Ao (p.2) and F_o (Lehiste 1970:68-71). Table 1, on p.7 below, gives the number of examples analysed in each sub-group. In Table 1, note that examples of C_sV(ŋ) sequences in syllable-types 1, 2 and 5 are lexically rare and did not occur in the corpus. The single example of C_sVŋ in syllable-type 4 was not measured. In the text below, I have used the following slightly redundant, but phonologically relevant transcriptions to designate subgroups within a particular syllable-type:

1. V is re-written as U in syllables of type 1, 2 and 5
2. V is re-written as L in syllables of type 3, 4 and 6
3. a final q indicates syllable-types 5 and 6
4. the diacritics ` ^ and ˇ are used to represent the pitch shapes on syllable-types 1, 3 and 2/4 respectively.

Thus CĽŋ stands for any type 4 syllable with syllable-final velar nasal and obstruent Initial, e.g. [dʑɜ̌ŋ] : *'heavy'*, and C_sLq stands for any type 6 syllable with sonorant Initial, e.g. [mɜ̌ʔ] : *'ink'*.

5. PROCEDURE (MENSURAL) AND INSTRUMENTATION

Each example was segmented, and the duration of the Final (DF), sonorant Initial (where present) and point of onset of the velar nasal (where present) ascertained by reference to wide-band spectrograms, and the wave form in high speed (200 cms/sec) oscillograms.

The offset of the Final was adjudged to occur at the point where the glottal pulse train showed an obvious discontinuity in the regularity of increase of period. In adjudging this point of phonation offset, at most a discrepancy of one glottal pulse, or about 1.0 - 0.8 csec was involved.

Segment boundaries between sonorant Initial consonants and vowels in C_sV(q) syllables - Painter (1979:19) calls them "fault transitions" because they have the appearance of geological faults - are characteristically sharp and unambiguous with respect to location. There was, therefore, no difficulty in determining the point of onset of the Final in these cases.

TABLE 1

NUMBER OF EXAMPLES ANALYSED IN EACH SUBGROUP

(C = unaspirated obstruent; C_s = sonorant; V = monophthongal vowel; ŋ = syllable-final velar nasal; - = phonotactically excluded.)

SYLLABLE TYPE	SEGMENTAL SPECIFICATION CV	CVŋ	C_sV	C_sVŋ
1	14	14		
2	12	5		
3	14	11	9	6
4	11	8	18	1
5	13	-	-	-
6	13	-	7	-

In syllables without sonorant Initial consonants, the point of onset of Final duration was equivalent to phonation onset. In such cases, it was often necessary to ignore the first obvious glottal pulse, because it would have had insufficient amplitude to be audible. This was the approach adopted by Lisker and Abramson (1969:416) in their well-known study of voice onset time.

The point of onset of the velar nasal consonant was defined as the first glottal pulse after the discontinuity in overall intensity distribution expected with nasals had occurred on the wide-band spectrogram. In the vast majority of cases, this point could be ascertained with the precision of one glottal pulse, (less than 1 csec).

In addition to the measurement of duration of Final, in syllables of type 2, 4, and 6 - that is, in syllables with a final rise in pitch contour - the duration from Final onset to F_o peak (DP) was also measured, using synchronised wide and narrow-band spectrograms.

Ao and F_o values were measured at various intervals of DP and DF, depending on the particular syllable-type: the exact points at which measurements were made were chosen to provide an optimal sampling of the curves involved, and can be read from the tables of results below. Generally, in syllables of type 1 and 3, Ao and F_o were samples at 10% intervals of Final duration, i.e. 0%, 10%, 20%...100%DF. In type 5 syllables, sampling points were at 20%DF intervals, i.e. 0%, 20%, 40%, ...100%DF. The Ao of syllables of type 2 and 4 was sampled at 10% DF points, as well as 5%, 15%, 85%, and 95%DF, and the F_o of type 2 and 4 syllables was sampled at 10% intervals of duration to F_o peak. In type 6 syllables, the amplitude was sampled at 20%DF intervals, and at 90%DF, and the F_o at 20%DP intervals. In addition, in syllables of type 2, 4, and 6, the F_o at 100%DF was measured.

In syllables with sonorant Initials, F_o was measured at onset and mid-point of sonorant duration, and Ao was measured at onset, mid-point, and three-quarters of sonorant duration.

In the tables of results below, F_o values are given in Hertz, to the nearest Hz, Ao values are given in decibels, to the nearest 0.1 dB, and duration values are given in centiseconds, to the nearest millisecond.

Onset and offset values of F_o are normally not resolved adequately by pitch extraction devices or narrow band spectrograms (Fant 1968:188, 189; Hombert 1976), and so I measured F_o at these points direct from the wave form in high speed oscillograms. Otherwise, measurements of F_o were calculated from the arithmetical mean of the F_o measurements as derived from the first 3 harmonics in narrow-band spectrograms with expanded frequency scale (1" : 200 Hz). All spectrograms used in this study were made on the model 7049 A Kay Sonagraph 80/85 - 8000 Hz Spectrum Analyser of the Cambridge University Department of Linguistics. This model has a narrow-band-pass filter of 45 Hz, and a wide-band-pass filter of 300 Hz. The high shaping pre-emphasis circuits were not used, all spectrography being done at the Flat 'Fl-1' setting.

In order to determine the accuracy of this spectrographic measurement technique, I compared 96 F_o measurements taken at various %DF points in

12 of the syllables examined with the F_o values for the same %DF points extracted by computer[6]. The 96 spectrographically derived measurements had an arithmetical mean value of 2.1 Hz less than those of the computer, with a standard deviation of 2.5 Hz. This means that - assuming the computerised 'pitch' detection to be accurate - the actual F_o value of a particular spectrographically derived F_o measurement is equal to the spectrographically derived measurement plus 2.1 Hz, plus or minus 4.2 Hz at the 90% confidence level. Thus the spectrographic method used is tolerably accurate - more accurate, in fact, than has been previously surmised possible in measuring F_o from narrow-band spectrograms[7]. I considered the value of 2.1 Hz too small to require adjustment to the actual value. I did not estimate the accuracy of F_o measurements made direct from the wave form: assuming constant and accurate running and playback speeds for the tape recorder and mingograph, these measurements should be more accurate than those made from the spectrograms.

The Ao measurements quoted in this study were made using the F-J Electronics Intensity Meter and Elema Schoenander Mingograph 34 of the Department of Linguistics, School of General Studies, at the Australian National University.

The F-J Intensity Meter employs full-wave, linear rectification. It has two channels: one provides a 500 Hz high-pass filtered output; the other registers amplitude over the full frequency range (50 - 15,000 Hz). The response of its full frequency range channel is given as ± 0.2 dB. In addition, integration time can be set independently for each channel at 2.5, 5, 10, 20, or 40 msec, and an optional linear or quasi logarithmic display is independently available on both channels. I chose the linear display and 20 msec smoothing time : the linear scale permits the most accurate registration at high amplitude levels, i.e. for vowels, and 20 msec is the integration time suggested by Lehiste (1970:124) on the basis of the relatively long time constant reported by Miller (1948) and Small, Brandt and Cox (1962) for auditory perception of loudness judgements.

It is important to note that, for this investigation, the intensity meter channel was chosen which registered amplitude over the full frequency range. Use of the high-pass filter would have complicated the relationship between the glottal and measured oral amplitude by introducing an additional variable, which is difficult to control for, of differential resolution of between-sample differences in spectral characteristics. The flat Ao values obtained in this study may be profitably compared with the corresponding high-pass filtered values given in Rose (1979) for an almost identical corpus.

Using a specially prepared calibration tape, which consisted of a 1000 Hz sine wave signal, the level of which was increased in 2 dB steps through 30 dB, I found that the intensity meter and the mingograph were very accurate indeed. For 14 measurements of the level of the calibration tone at each of the 2 dB steps from 30 db to 4 dB inclusive, the mean discrepancy was -0.33 dB, with a standard deviation of 0.25 dB. This means the actual Ao value for a given Ao measurement is equivalent to the meter measurement minus 0.33 dB, ± 0.4 dB at the 90% confidence level. The amplitude measurements were consequently left unadjusted.

6. RESULTS

Arithmetical mean values for duration, and fundamental frequency and amplitude as functions of the duration of the Final and/or duration to fundamental frequency peak in the 14 sub-groups examined are given in Tables 2 to 15. Standard deviations are given in brackets. Thus from Table 4, it can be seen that the mean value for the duration of the Final (DF) in type 2 CŬ syllables is 32 csec, with a standard deviation of 3 csec, whereas the mean value for duration to the fundamental frequency peak (DP) is 29.3 csec, with a standard deviation of 2.9 csec. It can also be seen that at 60% of the duration to the fundamental frequency peak (60%DP), or (29.3 csec x 0.6=) 17.6 csec from the onset of the Final, the mean F_0 for CŬ syllables is 135 Hz, with a standard deviation of 6 Hz. Also, at 60% of the duration of the Final (60%DF), or (32 csec x 0.6 = 19.2 csec from the onset of the Final, the mean Ao value for CŬ syllables is 22.7 dB, with a standard deviation of 2.8 dB.

The results are also plotted graphically in Figures 1 to 14. In these figures, the Ao and F_0 shapes are shown synchronised, in order to facilitate accurate comparison between the two parameters at any point in time. F_0 shapes for all syllable-types are compared in Figure 15, and Ao shapes for all syllable-types are compared in Figure 16. In none of these figures has the duration parameter been normalised. Note that the mean Ao values in dB have been replotted against a linear scale. If almost equal differences in loudness are derived by the perceptual mechanism from logarithmically equal intervals of amplitude (Ladefoged 1972:83), plotting the dB values logarithmically can easily give an erroneous visual impression of the way the amplitude may be perceived. Note also that the onset of the Final is shown to occur in all cases at csec 0, irrespective of whether the syllable has a sonorant Initial preceding the Final or not.

7. DISCUSSION

In the discussion of the results below, I shall first make some general observations on the contributions of the three acoustic parameters of F_o, Ao, and duration to contrasts between syllable-types. Next, I shall mention, for each of the acoustic parameters in turn, some salient points of the relationship between their shapes and values in each syllable-type, indicating some of the phonological implications which arise therefrom. I shall then briefly comment on some of the more important implications of the observed correlations between F_o and Ao, and between F_o and duration. Finally, I shall describe the effects on F_o, Ao and duration of the syllable-final velar nasal, and syllable-initial sonorant.

It is clear from the results of this investigation that each syllable-type is characterised acoustically in all three parameters of F_o, Ao, and duration. Indeed, each syllable-type can be defined equally well in any of the three parameters, with the exception of duration in CUq and CLq syllables.

In order to demonstrate the relative contributions of the individual acoustic parameters to a particular contrast between syllable-types, I shall examine the contrast between types 2 and 4 as realised on CŬ and CĽ syllables.

It is necessary first to break down the Ao and F_o shapes into the traditional dimensions of Register and Contour, defined, however, in the following way. Register can be understood as a kind of overall level (in the sense implied by Zee (1978)), and defined as the value for a particular Ao or F_o shape corresponding to the arithmetical mean of the Ao or F_o values at the %DF or %DP points. This absolute mean value can then be relativised by expressing it as a percentage of total relevant F_o or Ao range. For example, the mean F_o value of CŬ syllables is

$$\frac{143+138+136+134+134+134+135+137+141+146+153+140}{12} = 139 \text{ Hz},$$

and the mean F_o value of CĽ syllables is 116 Hz. The total mean F_o range in CV(q) syllables is 169 Hz (CUq, at 0%DF) minus 105 Hz (CĽ, at 30%DP) = 64 Hz. The F_o register value for CŬ syllables is then

$$\frac{139 - 105}{64} \text{ Hz} \times 100 = 53\%,$$

and for CĽ syllables

$$\frac{116 - 105}{64} \text{ Hz} \times 100 = 17\%.$$

Therefore, CĽ and CŬ syllables are separated by a 36% difference in register, which is very highly significant ($p < .001$; $t=8.16$, $df=21$).

TABLE 2

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 14 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 1, WITH MONOPHTHONGAL VOWEL FINAL, AND UNASPIRATED OBSTRUENT INITIAL (CÙ).

Percentage of duration of Final (DF)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	160	165	162	159	156	151	146	139	132	121	110
	(10	6	7	7	8	7	7	7	7	8	9)
Amplitude	17.9	22.1	22.4	22.6	22.5	22.4	22.3	21.5	20.5	18.6	14.9
	(2.3	2.7	2.8	2.5	2.3	2.4	2.3	2.0	2.2	1.9	1.5

Duration of Final = 25.7 (2.8)

TABLE 3

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 14 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 1, WITH UNASPIRATED OBSTRUENT INITIAL, AND SYLLABLE-FINAL VELAR NASAL (CÙŋ).

Percentage of duration of Final (DF)	0	10	20	30	40	45	50	60	70	80	90	100
Fundamental frequency	177	181	176	172	168		161	153	145	136	126	116
	(10	9	8	8	8		9	10	9	9	8	7)
Amplitude	19.8	25.3	24.6	23.5	23.0	22.6	22.7	22.3	21.1	19.8	17.2	13.8
	(3.4	2.5	1.7	1.4	1.3	1.5	1.3	1.1	1.2	1.3	1.2	1.0)

Duration of Final = 22.2 (3.2); onset of nasal consonant occurs at csec 9.2 (1.5), i.e. 41% DF.

TABLE 4

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 12 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 2, WITH MONOPHTHONGAL VOWEL-AND-GLOTTAL-STOP FINAL, AND UNASPIRATED OBSTRUENT INITIAL (CŬ).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	143 (9	138 6	136 6	134 6	134 6	134 6	135 6	137 7	141 8	146 8	153 9)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	18.6 (2.5	22.7 2.8	23.4 2.9	23.4 3.0	23.1 3.0	22.8 2.8	22.3 2.7	22.2 2.8	22.7 2.8	23.4 2.6	24.4 2.8	24.5 3	23.8 3.2	22.9 3.7	18.9 2.9)

Duration of Final = 32 (3); duration to fundamental frequency peak = 29.3 (2.9); fundamental frequency at 100% DF = 140 (7).

TABLE 5

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 5 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 2, WITH UNASPIRATED OBSTRUENT INITIAL, AND SYLLABLE-FINAL VELAR NASAL AND GLOTTAL-STOP (CŬŋ).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	142 (4	138 4	135 5	134 4	135 3	136 4	139 5	142 6	147 7	154 6	159 6)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	20.6 (2.7	24.4 3.1	23.7 3.4	23.3 3.5	23.1 3.3	21.7 3.2	21.1 2.6	20.9 2.6	21.4 3.0	21.9 2.9	22.4 2.8	22.6 2.4	22.1 2.4	21.3 1.9	17.8 2.5)

Duration of Final = 33.8 (2.7), duration to fundamental frequency peak = 31.8 (3.4); fundamental frequency at 100% DF = 145 (7) onset of nasal consonant occurs at csec 12.3 (2.5), i.e. 36% DF, or 39% DP.

TABLE 6

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB) AND DURATION (csec) IN 14 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 3, WITH MONOPHTHONGAL VOWEL FINAL, AND OBSTRUENT INITIAL (CL̂).

Percentage of duration of Final (DF)	0	5	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	122 (8	116 8	117 7	125 8	133 7	140 6	143 6	143 6	140 6	134 6	125 8	110 9)
Amplitude	15.9 (2.8		19.3 3.3	21.1 2.8	22.7 2.4	23.6 2.7	23.7 2.4	23.6 2.5	23.3 2.3	22.0 2.2	19.7 2.1	14.9 1.9)

Duration of Final = 29.6 (3.6)

TABLE 7

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 11 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 3, WITH SYLLABLE-FINAL VELAR NASAL, AND OBSTRUENT INITIAL (CL̂ŋ).

Percentage of duration of Final (DF)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	125 (9	119 8	123 7	135 7	144 6	152 7	153 8	151 8	144 9	133 7	118 6)
Amplitude	16.8 (2.6	20.0 4.0	22.1 3.3	23.7 2.6	24.8 2.3	24.6 2.3	24.4 2.4	23.3 2.3	22.2 2.3	20.1 2.2	15.0 1.7)

Duration of Final = 26.3 (2.5); onset of nasal consonant occurs at csec 10.8 (1.8), i.e. 41% DF.

TABLE 8

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 9 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 3, WITH MONOPHTHONGAL VOWEL FINAL, AND SONORANT INITIAL ($C_s\hat{L}$).

Percentage of duration of Final (DF)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	115 (4	118 4	124 4	131 5	138 5	144 6	145 6	143 6	137 6	128 6	119 7)
Amplitude	16.5 (2.7	19.5 2.4	21.3 1.9	23.0 1.3	24.1 1.5	24.8 1.5	24.9 1.3	25.0 1.7	23.5 2.1	21.1 2.2	15.3 2.8)

Duration of Final = 30.1 (3); duration of sonorant Initial = 8.8 (2.3); fundamental frequency at sonorant onset = 120 (9), fundamental frequency at sonorant mid-point = 108 (4); amplitude at sonorant onset = 10.3 (0.8); amplitude at sonorant mid-point = 14.4 (1.3); amplitude at 3/4 of sonorant duration = 15.1 (1.5).

TABLE 9

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 6 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 3, WITH SYLLABLE-FINAL VELAR NASAL, AND SONORANT INITIAL ($C_s\hat{L}$ŋ).

Percentage of duration of Final (DF)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	117 (4	121 5	128 5	136 5	144 5	151 7	156 11	154 12	147 12	135 11	115 9)
Amplitude	18.5 (3.1	21.9 3.2	23.1 3.4	24.1 2.9	24.7 2.9	24.0 2.9	23.6 2.9	23.0 2.8	22.0 3.2	20.3 3.4	14.7 1.6)

Duration of Final = 27.5 (2); duration of sonorant Initial = 8.6 (2.4); onset of nasal consonant occurs at csec 10.9 (1.7), i.e. 40% DF, fundamental frequency at sonorant onset = 121 (6); fundamental frequency at sonorant mid-point = 115 (4); amplitude at sonorant onset = 9.3 (3); amplitude at sonorant mid-point = 15.8 (3.8); amplitude at 3/4 of sonorant duration = 17 (3.6).

TABLE 10

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 11 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 4, WITH OBSTRUENT INITIAL, MONOPHTHONGAL VOWEL FINAL, AND OPTIONAL SYLLABLE-FINAL GLOTTAL-STOP (CV̆).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	119 (7	108 6	106 7	105 7	106 8	108 8	112 8	116 8	124 7	132 7	141 8)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	14.4 (2.1	16.5 2.6	15.9 2.7	15.4 2.7	15.2 2.6	15.0 2.6	15.3 2.6	16.2 2.9	17.3 3.0	18.6 3.0	20.0 3.1	20.5 3.0	20.7 2.6	20.2 2.6	16.4 2.1)

Duration of Final = 36.3 (4.2); duration to fundamental frequency peak = 33.8 (4) fundamental frequency at 100%DF = 121 (6).

TABLE 11

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 8 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 4, WITH OBSTRUENT INITIAL, AND SYLLABLE-FINAL VELAR NASAL AND OPTIONAL GLOTTAL-STOP (CV̆ŋ).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	126 (5	109 6	102 5	100 4	101 4	103 4	107 5	113 5	123 6	132 5	146 5)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	16.7 (2.8	19.1 4.2	17.4 3.6	16.9 3.2	16.6 2.9	16.2 2.4	16.0 2.6	16.7 2.4	17.7 2.2	18.7 2.3	19.9 2.5	20.1 2.6	20.1 2.4	19.4 2.7	15.9 2.3)

Duration of Final = 36.3 (4.1), duration to fundamental frequency peak = 34.5 (3.5); fundamental frequency at 100%DF = 128 (9); onset of nasal consonant occurs at csec 13.4 (3.1), i.e. 37%DF, or 39%DP.

TABLE 12

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 18 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 4, WITH MONOPHTHONGAL VOWEL-AND-GLOTTAL-STOP FINAL, AND SONORANT INITIAL ($C_s\check{L}$).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	110 (6	107 6	106 6	106 6	107 6	109 6	112 5	117 6	124 7	131 7	140 8)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	16.2 (2.0	17.0 2.2	17.2 2.2	17.2 2.1	17.2 2.4	16.8 2.7	16.9 2.7	17.5 2.9	18.2 2.9	19.5 3.1	20.8 3.0	21.4 3.0	21.6 2.9	21.4 3.1	18.0 2.6)

Duration of Final = 36.9 (3.2); duration to fundamental frequency peak = 34.6 (3.1); duration of sonorant Initial = 9.2 (2.8); fundamental frequency at 100%DF = 124 (9); fundamental frequency at sonorant onset = 114 (5); fundamental frequency at sonorant mid-point = 110 (5), amplitude at sonorant onset 11.3 (2.4); amplitude at sonorant mid-point = 14.5 (2.7); amplitude at 3/4 of sonorant duration = 15.0 (2.3).

TABLE 13

MEAN AND STANDARD DEVIATION VALUES FOR FUNAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 13 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 5, WITH UNASPIRATED OBSTRUENT INITIAL, AND SYLLABLE-FINAL GLOTTAL-STOP (CUq).

Percentage of duration of Final (DF)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	169 (12		167 10		168 11		166 9		160 8		127 11)
Amplitude	20.4 (2.3		25.8 2.1		25.8 2.5		25.8 2.6		25.1 2.7		20 2)

Duration of Final = 9.1 (2.4)

TABLE 14

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 13 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 6, WITH OBSTRUENT INITIAL, AND SYLLABLE-FINAL GLOTTAL-STOP (CLq).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	121 (9		116 7		119 8		131 10		145 9		148 8)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	16.2 (1.9				18.2 2.4		19.7 2.4		22.2 2.5		23.7 1.5		22.8 1.4		19.2 2.1)

Duration of Final = 10.5 (1.5); duration to fundamental frequency peak = 9.2 (1.4), fundamental frequency at 100% DF = 123 (16)

TABLE 15

MEAN AND STANDARD DEVIATION VALUES FOR FUNDAMENTAL FREQUENCY (Hz), AMPLITUDE (dB), AND DURATION (csec) IN 7 SHANGHAI-ZHENHAI CITATION MONOSYLLABLES OF TYPE 6, WITH SONORANT INITIAL, AND SYLLABLE-FINAL GLOTTAL-STOP (C_sLq).

Percentage of duration to fundamental frequency peak (DP)	0	10	20	30	40	50	60	70	80	90	100
Fundamental frequency	110 (2		113 3		118 5		126 6		135 2		147 3)

Percentage of duration of Final (DF)	0	5	10	15	20	30	40	50	60	70	80	85	90	95	100
Amplitude	18.0 (2.0				20.0 1.9		21.8 1.4		23.7 1.5		24.4 1.9		22.9 1.9		19.3 2.2)

Duration of Final = 13.2 (3.6); duration to fundamental frequency peak = 10.8 (3.2); duration of sonorant Initial = 9.4 (1.8); fundamental frequency at 100%DF = 117 (7); fundamental frequency at sonorant onset = 114 (5); fundamental frequency at sonorant mid-point = 108 (3); amplitude at sonorant onset = 11.6 (1.5); amplitude at sonorant mid-point = 15.6 (1.7) amplitude at 3/4 of sonorant duration = 16.5 (1.9).

FIGURE 1

Synchronised Fundamental Frequency and Relative Amplitude Values for 14 Shanghai-Zhenhai Citation Monosyllables of Type 1, with Monophthongal Vowel Final, and Unaspirated Obstruent Initial (CŬ).

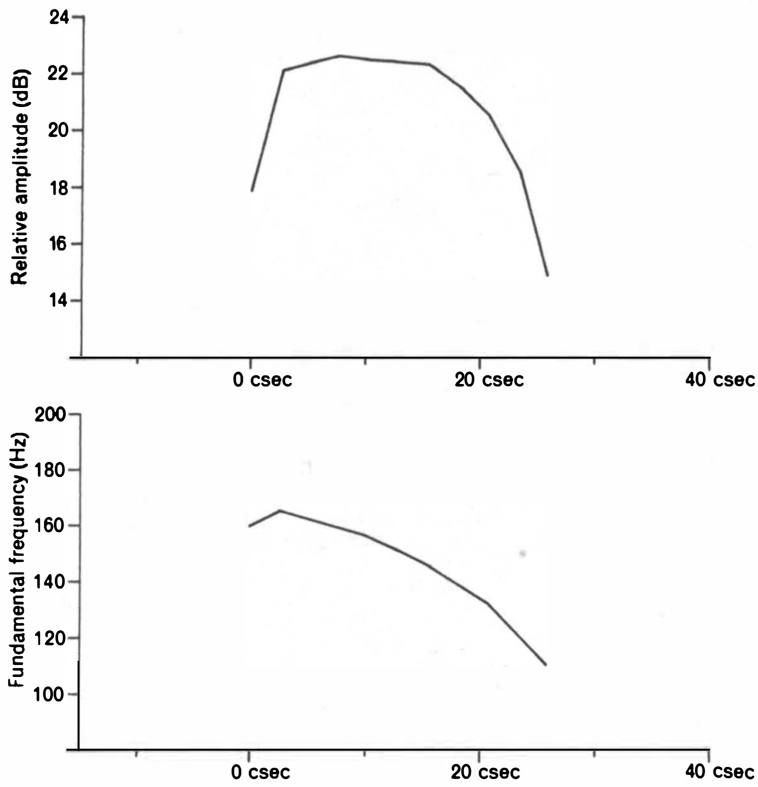


FIGURE 2

Synchronised Fundamental Frequency and Relative Amplitude Values for 14 Shanghai-Zhenhai Citation Monosyllables of Type 1, with Unaspirated Obstruent Initial, and Syllable-Final Velar Nasal (CṺŋ). 'N' Indicates Point of Onset of Velar Nasal

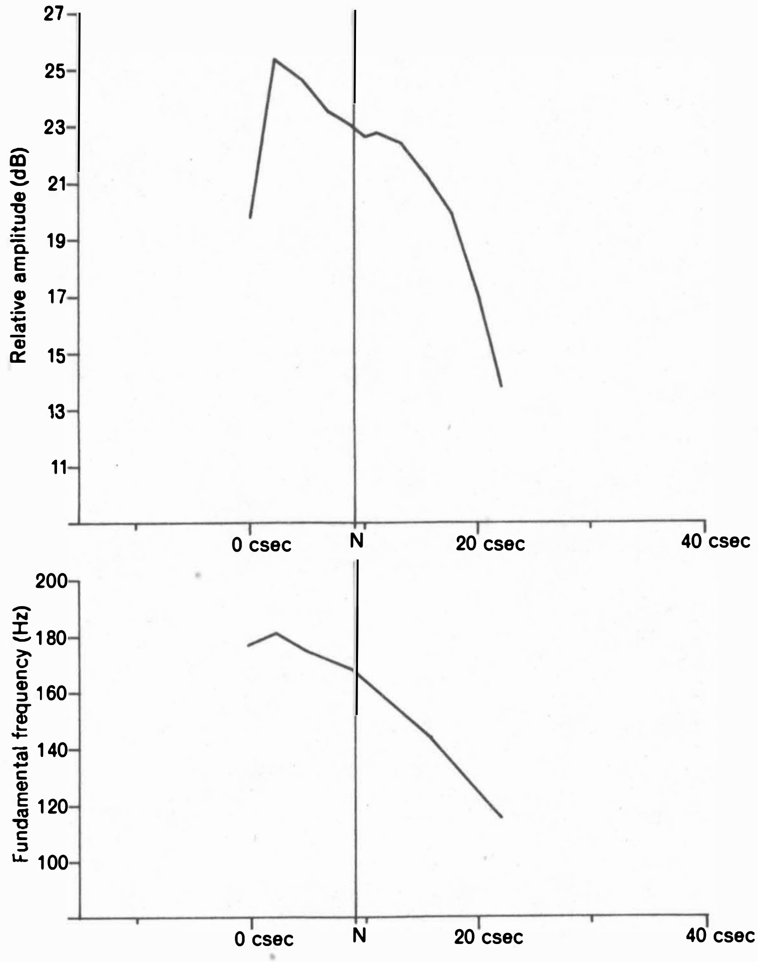


FIGURE 3

Synchronised Fundamental Frequency and Relative Amplitude Values for 12 Shanghai-Zhenhai Citation Monosyllables of Type 2, with Monophthongal Vowel-and-Glottal-Stop Final, and Unaspirated Obstruent Initial (CV̆).

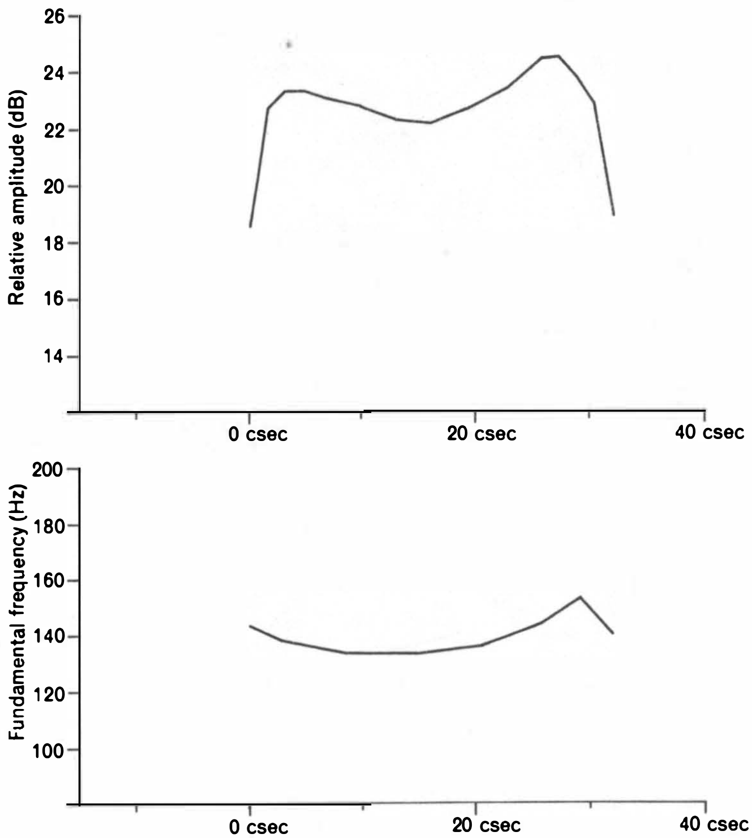


FIGURE 4

Synchronised Fundamental Frequency and Relative Amplitude Values for 5 Shanghai-Zhenhai Citation Monosyllables of Type 2, with Unaspirated Obstruent Initial, and Syllable-Final Velar Nasal and Glottal-Stop (CǓŋ). 'N' Indicates Point of Onset of Velar Nasal.

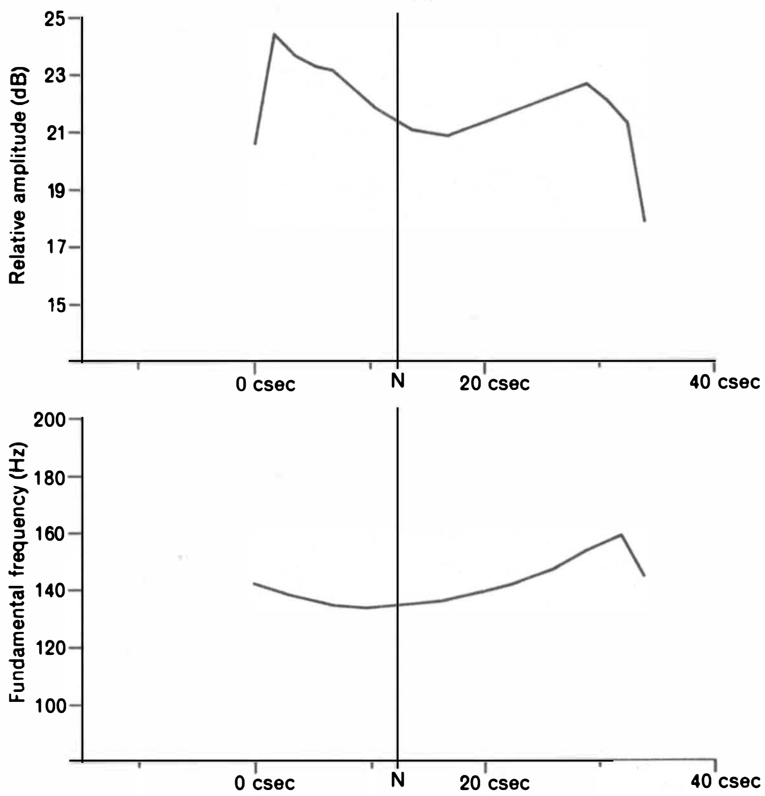


FIGURE 5

Synchonised Fundamental Frequency and Relative Amplitude Values for 14 Shanghai-Zhenhai Citation Monosyllables of Type 3, with Monophthongal Vowel Final, and Obstruent Initial (CL̂).

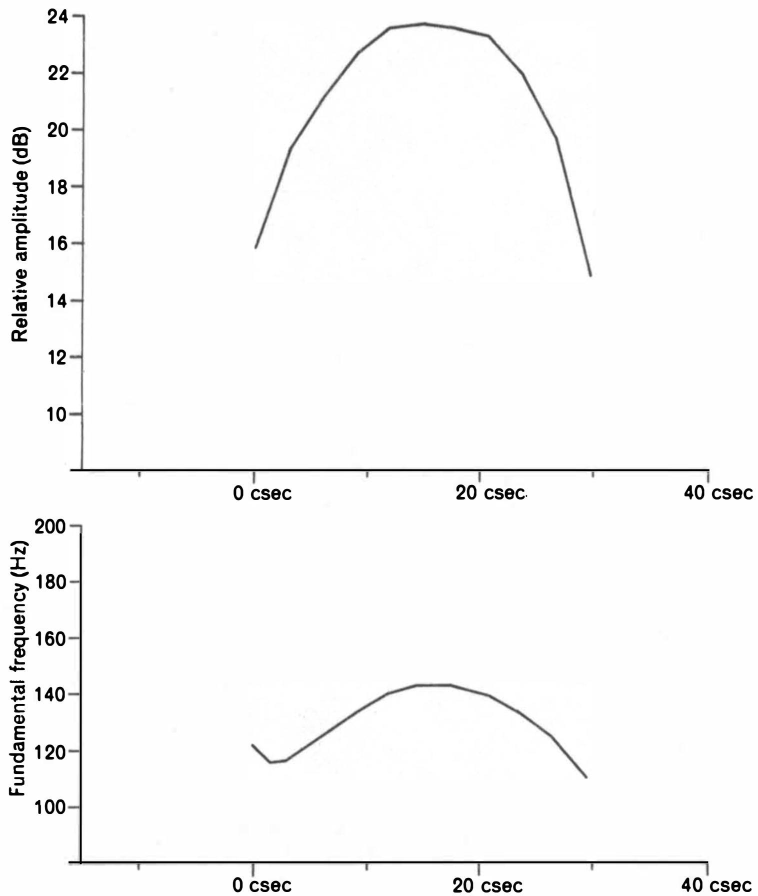


FIGURE 6

Synchronised Fundamental Frequency and Relative Amplitude Values for 11 Shanghai-Zhenhai Citation Monosyllables of Type 3, with Obstruent Initial, and Syllable-Final Velar Nasal (CL̂ŋ). 'N' Indicates Point of Onset of Velar Nasal.

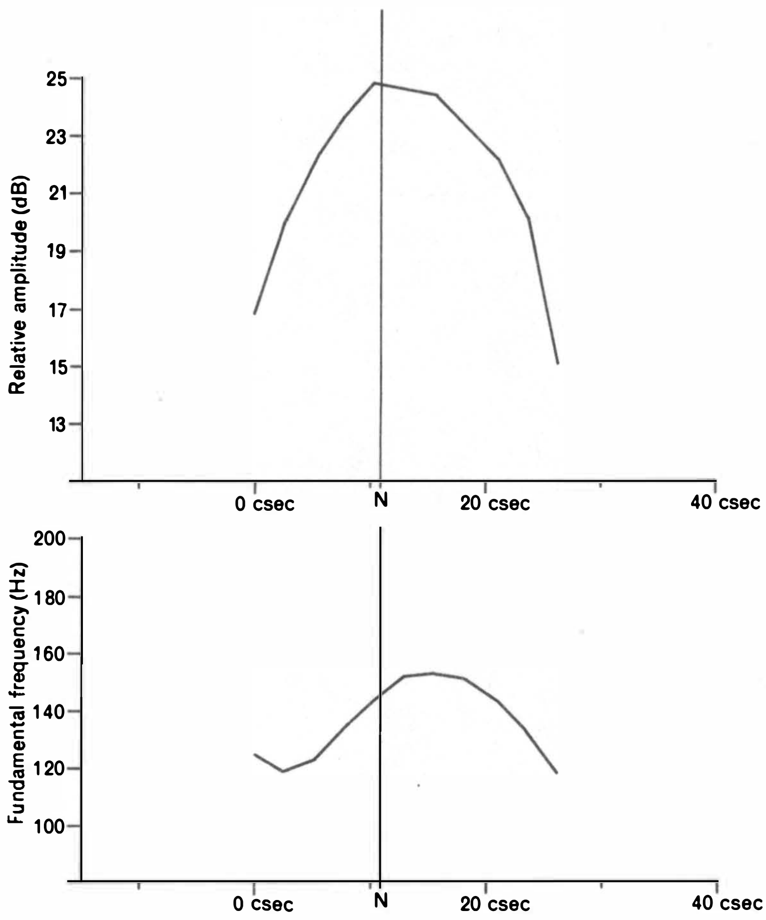


FIGURE 7

Synchronised Fundamental Frequency and Relative Amplitude Values for 9 Shanghai-Zhenhai Citation Monosyllables of Type 3, with Monophthongal Vowel Final, and Sonorant Initial (C_sL). The Point of Onset of the Vowel Final is Located at csec 0.

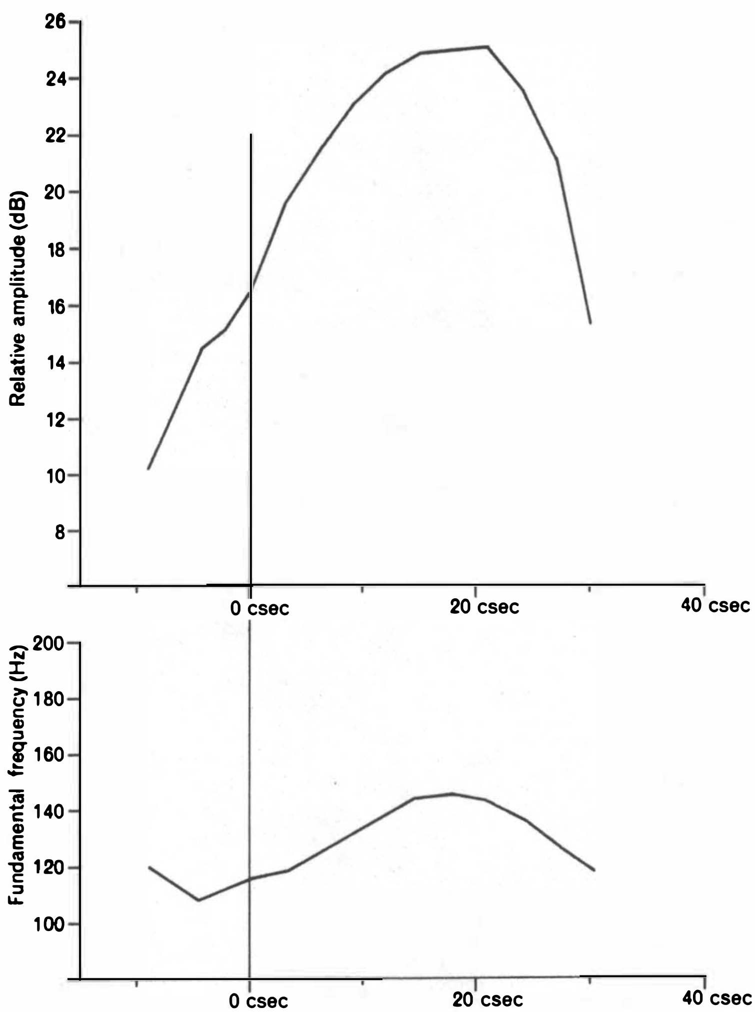


FIGURE 8

Synchronised Fundamental Frequency and Relative Amplitude Values for 6 Shanghai-Zhenhai Citation Monosyllables of Type 3, with Syllable-Final Velar Nasal, and Sonorant Initial (C_sL̂ŋ). 'N' Indicates Point of Onset of Velar Nasal. The Point of Onset of the Vowel Final is Located at csec 0.

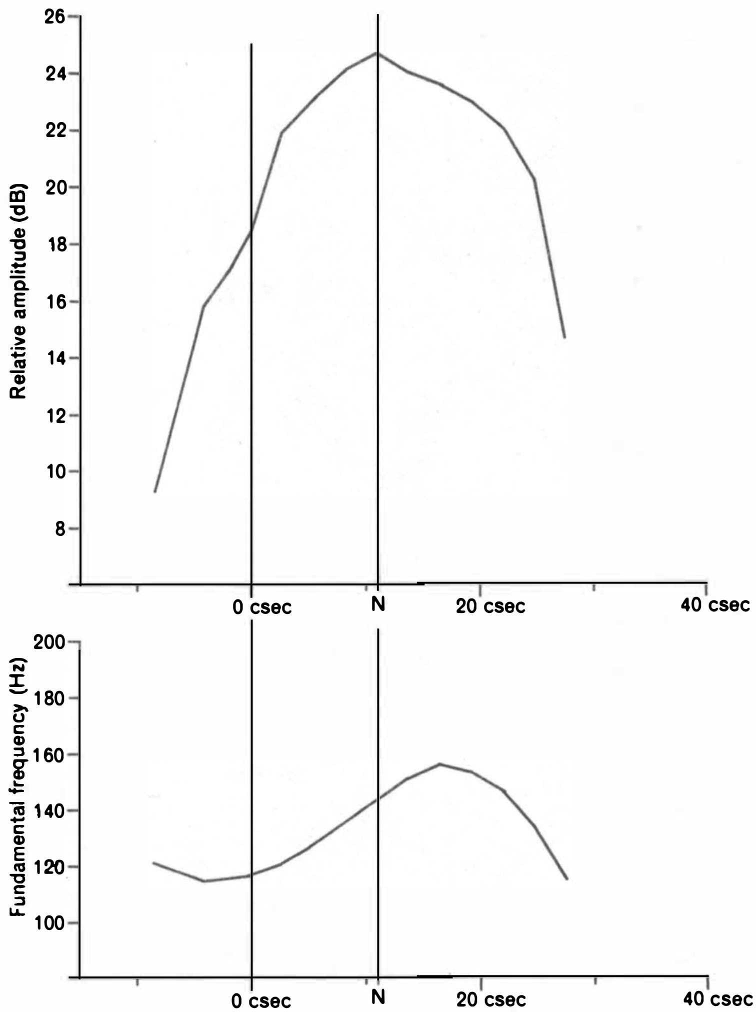


FIGURE 9

Synchronised Fundamental Frequency and Relative Amplitude Values for 11 Shanghai-Zhenhai Citation Monosyllables of Type 4, with Obstruent Initial, and Monophthongal Vowel Final with Optional Glottal-Stop (CĽ).

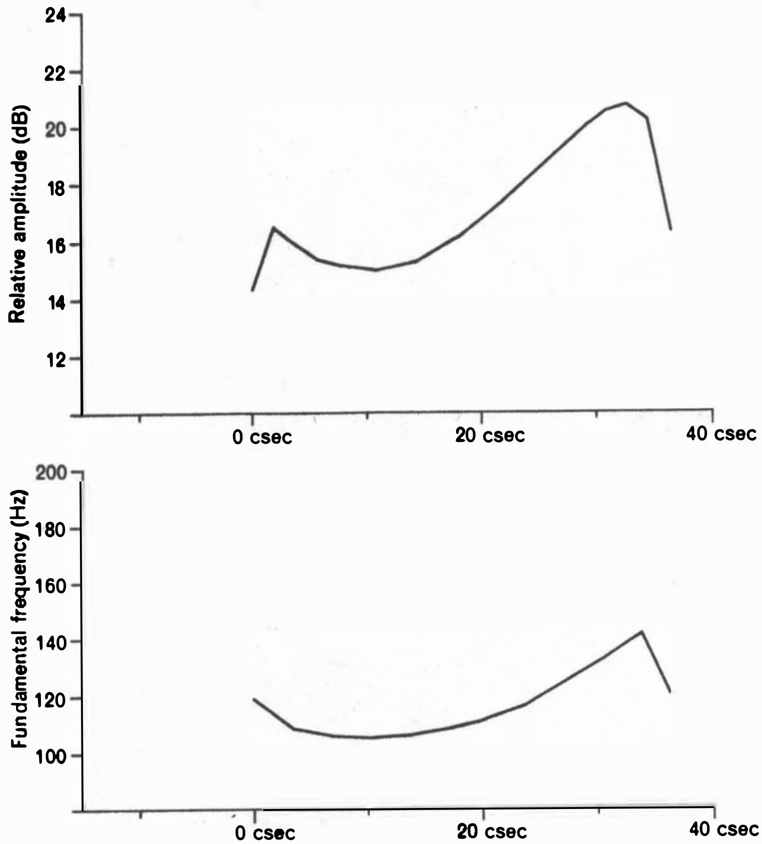


FIGURE 10

Synchronised Fundamental Frequency and Relative Amplitude Values for 8 Shanghai-Zhenhai Citation Monosyllables of Type 4, with Obstruent Initial, and Syllable-Final Velar Nasal and Optional Glottal-Stop (CĽŋ). 'N' Indicates Point of Onset of Velar Nasal.

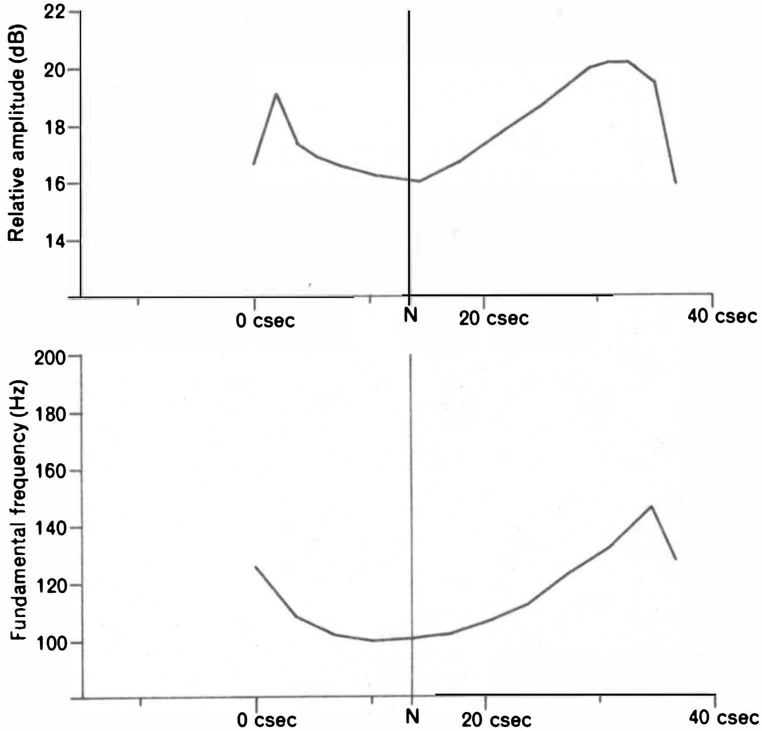


FIGURE 11

Synchronised Fundamental Frequency and Relative Amplitude Values for 18 Shanghai-Zhenhai Citation Monosyllables of Type 4, with Sonorant Initial, and Monophthongal Vowel Final with Optional Glottal-Stop ($C_s\breve{L}$). The Point of Onset of Vowel Final is Located at csec 0.

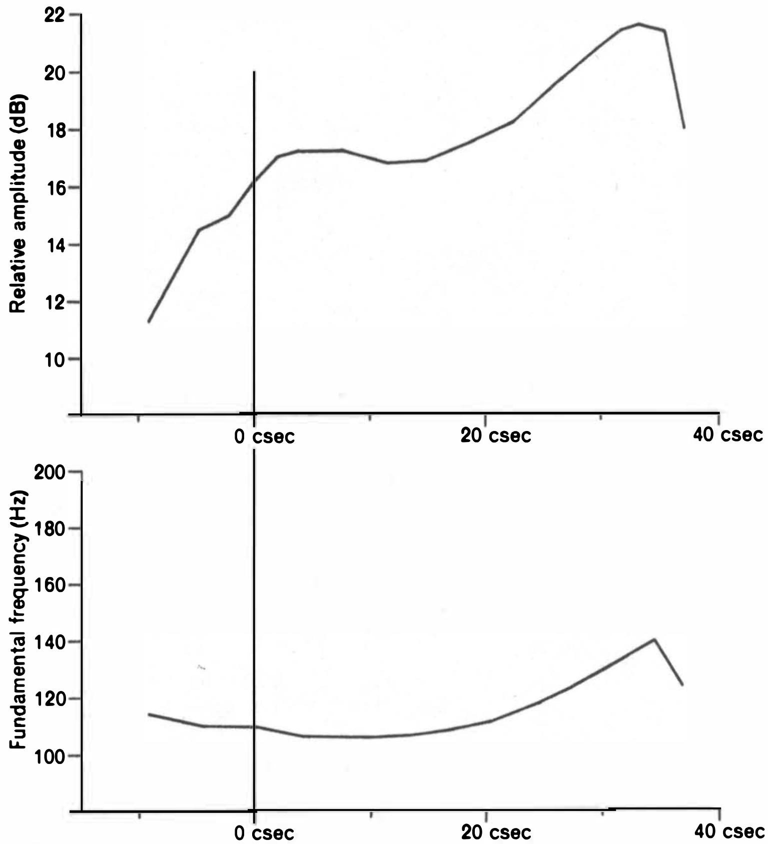


FIGURE 12

Synchronised Fundamental Frequency and Relative Amplitude Values for 13 Shanghai-Zhenhai Citation Monosyllables of Type 5, with Unaspirated Obstruent Initial, and Syllable-Final Glottal-Stop (CUq).

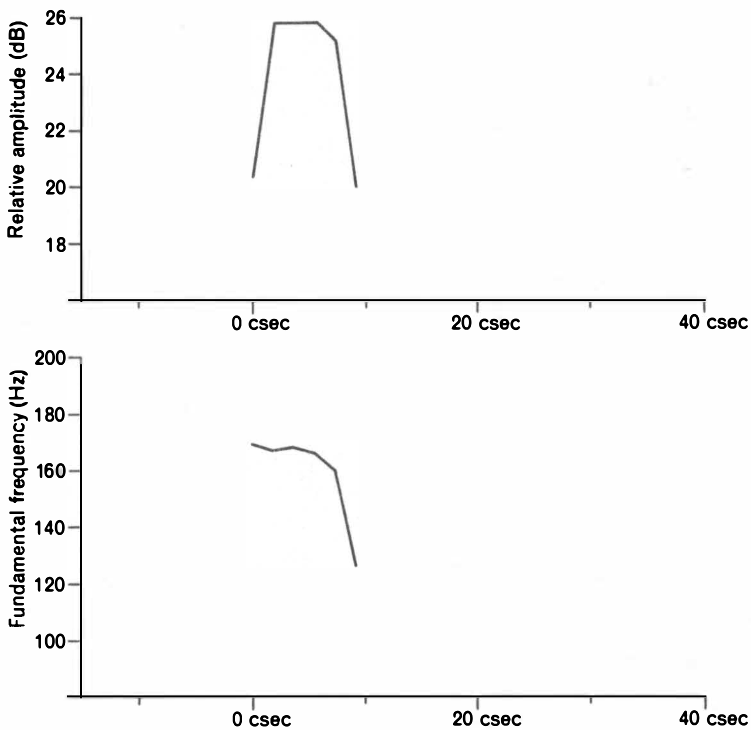


FIGURE 13

Synchronised Fundamental Frequency and Relative Amplitude Values for 13 Shanghai-Zhenhai Citation Monosyllables of Type 6, with Unaspirated Obstruent Initial, and Syllable-Final Glottal-Stop (CLq).

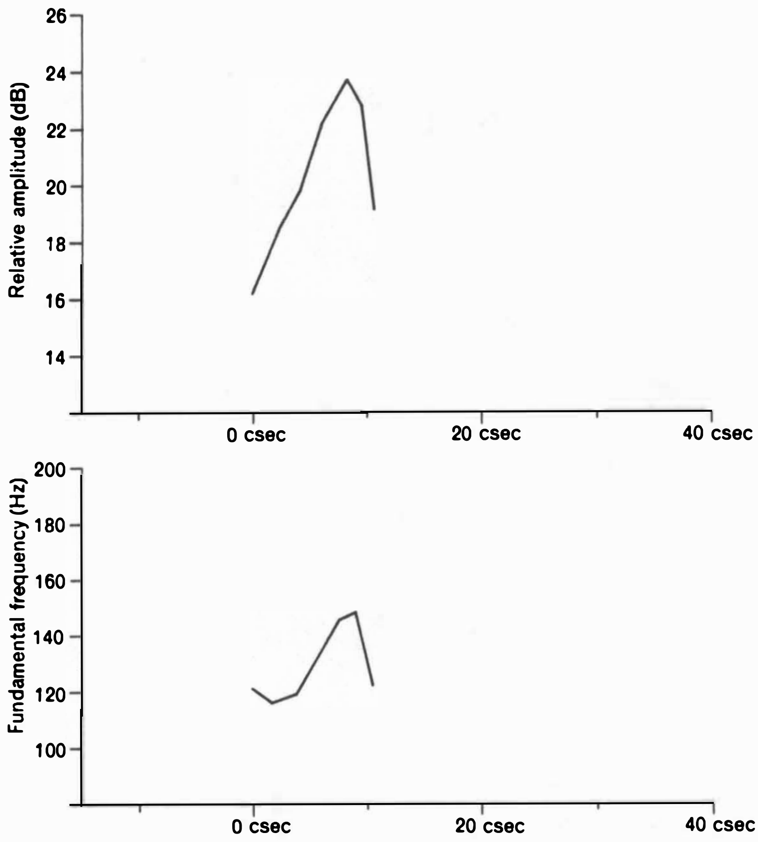


FIGURE 14

Synchronised Fundamental Frequency and Relative Amplitude Values for 7 Shanghai-Zhenhai Citation Monosyllables of Type 6, with Sonorant Initial, and Syllable-Final Glottal-Stop (C_sLq). The Point of Onset of Vowel Final is Located at csec 0.

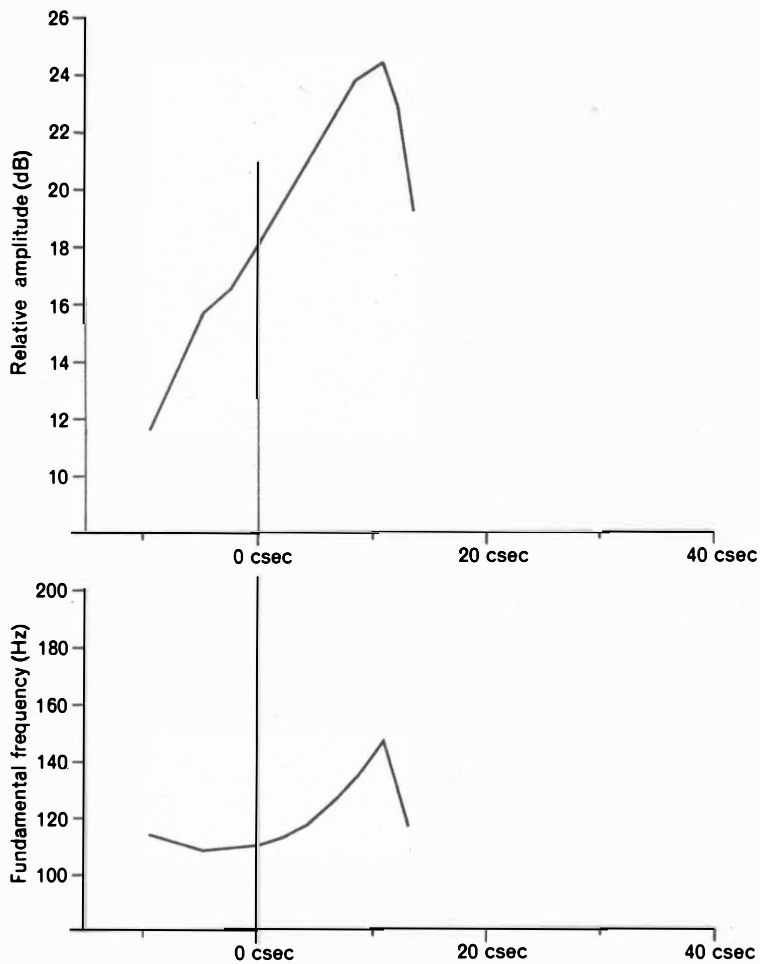


FIGURE 15

Fundamental Frequency Values Compared for

(A) CV(q),
(B) CVŋ,
(C) C_sL(q) / C_sL̂ŋ Syllables.

...... = $C_{(s)}$Vq Syllables;

—- — - — = C_sL̂ŋ Syllables.

Onset of Final is shown at csec 0.

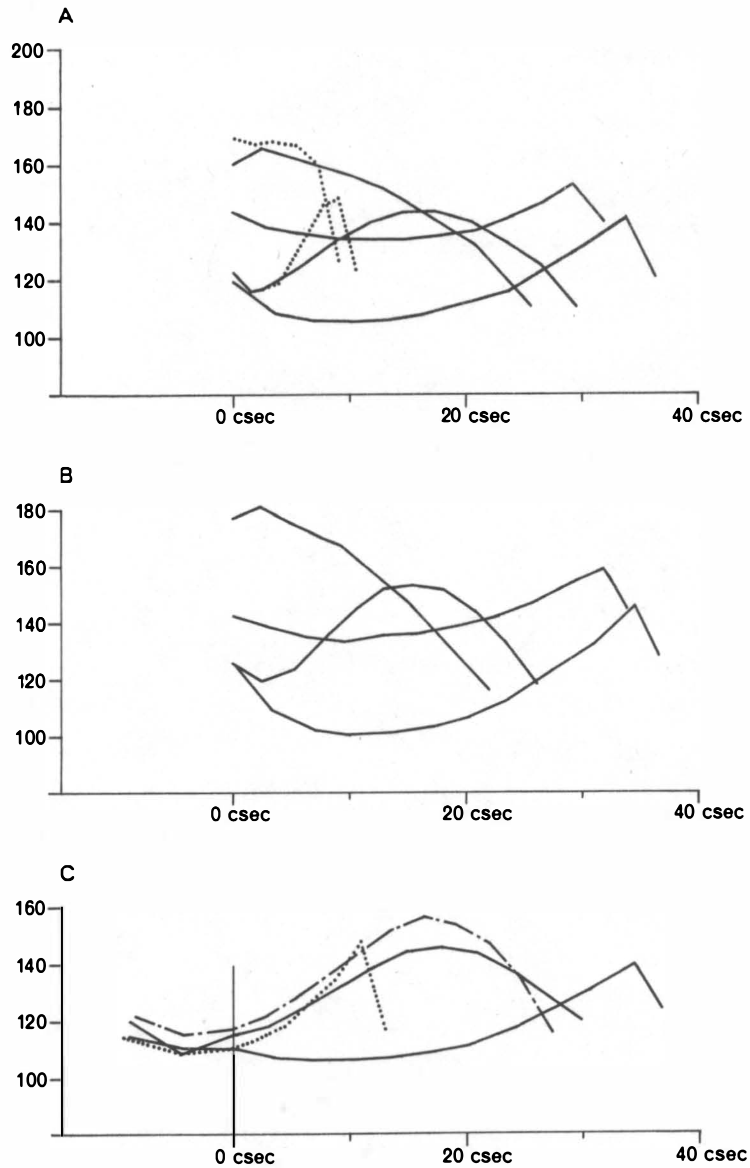
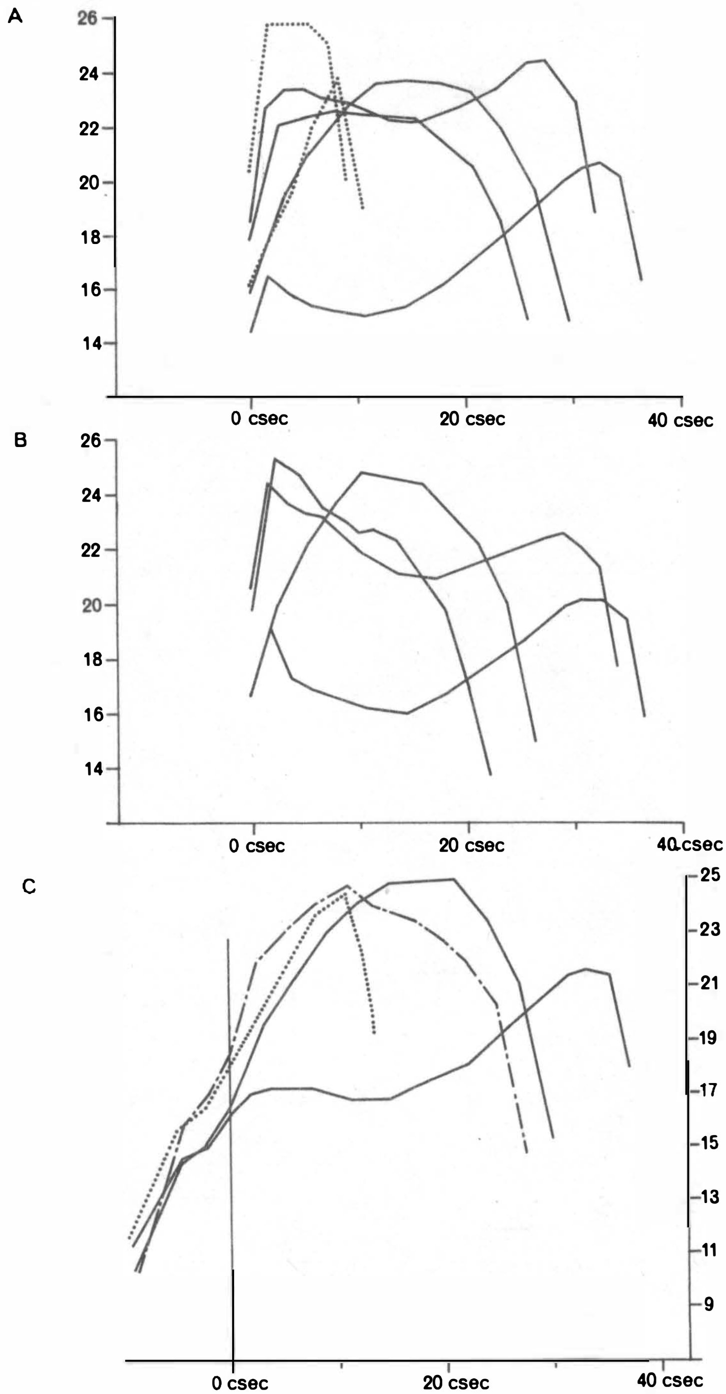


FIGURE 16

Relative Amplitude Values Compared for
(A) CV(q), (B) CVŋ, and (C) C_sL(q) / C_sL̂ŋ Syllables;
...... = $C_{(s)}$Vq Syllables; —-—-— = C_sL̂ŋ Syllables.
Onset of Final is shown at csec 0.



Contour may be defined as the curve obtained by expressing each %DF or %DP point measurement of Ao or F_o as a percentage of the mean Ao or F_o range of the syllable-type in question. Plotting contour values in this way factors out the parameter of duration and the dimension of register. For example, the mean F_o range in $C\check{U}$ syllables is 153 Hz (100%DP) minus 134Hz (30, 40 50%DP) = 19Hz. The mean F_o onset value (143Hz) will have, therefore, a contour value of

$$\frac{143 - 134}{19} \text{ Hz} \times 100 = 47\%.$$

The F_o contours of $C\check{U}$ and $C\check{L}$ syllables are compared in Figure 17. It can be seen that both types share similar onset and offset contour values (at approximately mid-range). Their peaks occur, and lowest points are reached, at very nearly the same %DF points (about 93%DP and 27%DP respectively). However, the $C\check{L}$ contour falls quicker and starts to rise earlier than the $C\check{U}$ contour, so that differences of 13% and 15% of F_o range separate them at 9%DF and 65%DF respectively. Analysis of the contour values for these two types excluding values at 100%DF (Table 16 below) shows that there is no reason at the 5% level to reject the hypothesis that they come from the same populations, whereas the difference in register between the two types is highly significant (p.11).

TABLE 16

REGRESSION ANALYSIS OF $C\check{U}$ AND $C\check{L}$ F_o CONTOURS

Source	Sums of Squares of Residuals (3 degree polynomial)	df	Mean Squares	Variance Ratio
Combined regression, $C\check{U}$, $C\check{L}$	731.18	18	40.62	
Sum of regressions, $C\check{U}$, $C\check{L}$	148.58	14	10.61	
Difference, combined regression - sum of regressions		4	30.01	2.83

(F = 3.11 for 4/14 df at 5% level)

It can be seen, therefore, that within the parameter of F_o, the contrast between $C\check{U}$ and $C\check{L}$ syllables demonstrably resides in the dimension of F_o register.

As far as the Ao parameter is concerned, the difference in Ao register between syllable-types 2 and 4 is greater than for F_o. $C\check{U}$ syllables have an Ao register value of 72%, compared with 24% for $C\check{L}$ syllables. There is therefore an Ao register difference between $C\check{L}$ and $C\check{U}$ syllables of 48% (cf. 36% for F_o register).

The difference in Ao contour between $C\check{U}$ and $C\check{L}$ syllables is shown in Figure 18. It can be seen once again that there is greater contrastiveness in the Ao parameter than in the F_o parameter: the Ao contours are obviously more contrastive than their corresponding F_o contours.

The main difference between the Ao contours of the two types lies in the relatively greater concentration of Ao in the first half of CŬ syllables: note that, in CŬ syllables, the first peak is only 19% of dB range down on the second peak, compared to a difference of 47% of range in CĽ syllables. In CĽ syllables, too, Ao falls, and starts to rise, earlier than in CŬ syllables: a characteristic which, it can be noted, parallels the F_o contours of these syllables (Figure 17).

The contrast in Ao shapes between CĽ and CŬ syllables is therefore a function of both register and contour differences.

CŬ syllables also differ from CĽ syllables in the parameter of duration. The difference, whether between DF values (4.3 csec), or between DP values (4.5 csec), is shown by t-test to be highly significant (DF: $.02 > p > .01$; $t = 2.7$, $df = 21$; DP: $.01 > p$; $t = 2.97$, $df = 21$).

To summarise, the above analysis has established that CŬ and CĽ syllables contrast acoustically in all three parameters of F_o, Ao, and duration. With Ao and F_o, it is the register differences which appear most important, although Ao contours are also contrastive.

FIGURE 17

Comparison of F_o Contours in CŬ (——) and CĽ (—·—·—) Syllables

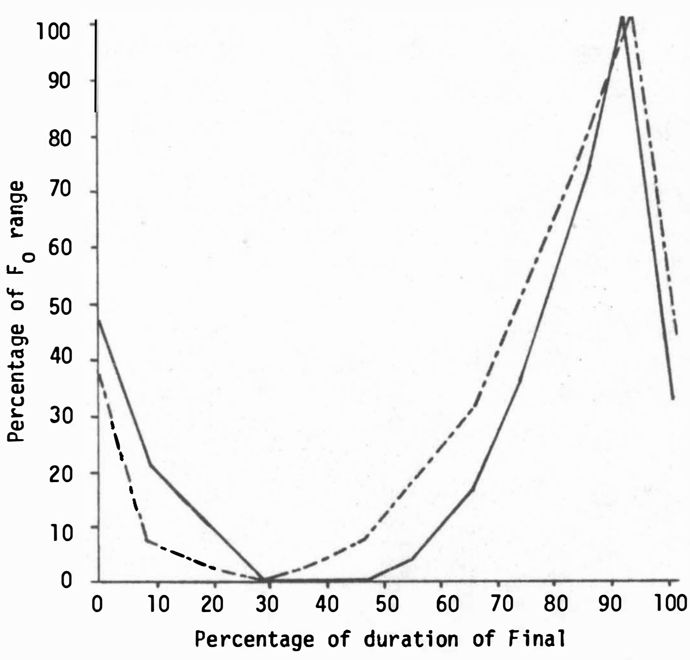
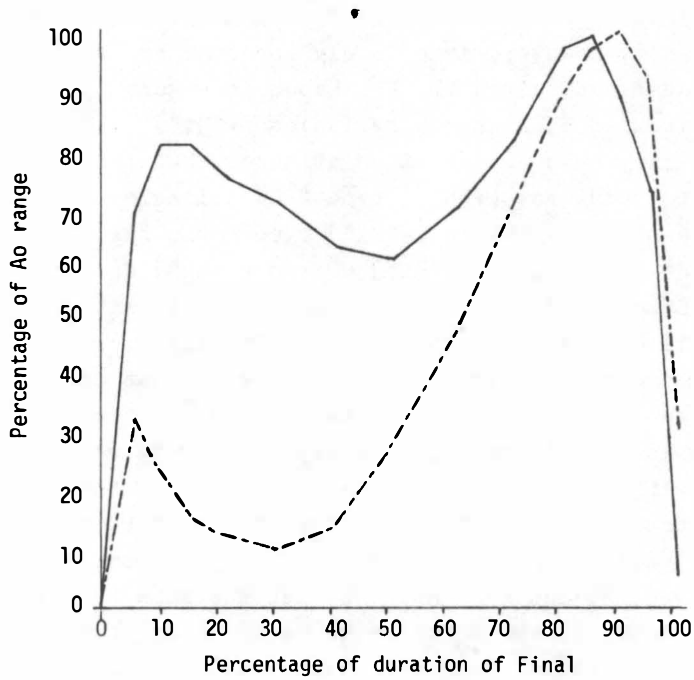


FIGURE 18

Comparison of Ao Contours in CǓ (——) and CĽ (—-—-—) Syllables



8. FUNDAMENTAL FREQUENCY

The general F_o shapes for the individual syllable-types (Figure 15) require little comment. Apart from the CUq syllables, which sounded to me higher, with respect to CÙ syllables, than their mean F_o shapes appear to warrant, the F_o data confirm the auditory impressions. The abrupt fall in F_o at the end of types 2, 4, 5, and 6 is inaudible as such: it is one of the acoustic correlates of the syllable-final glottal stop in these syllables (cf. the abrupt fall in F_o reported for [ʔ] also in Shanghai dialect (Zee and Maddieson 1979).

It can be noted that, although Chao's (1930) notation is of course intended as a method of transcribing pitch, the F_o shapes can be expressed in terms of the five-point scale, without perhaps too much distortion: CÙ: 51 (52?); CǓ: 334 (434? 324?); CĹ: 232; CĽ: 213, CUq: 5; CLq: 23 (24?).

There are several phonological rules in Sh-Zh, and Zhenhai dialect, which divide the syllable-types into two natural classes of 1, 2, and

5 on the one hand and 3, 4, and 6 on the other. (These two natural classes also correspond to the categories of Yin and Yang in traditional Chinese phonology). This kind of situation has been previously handled with undefined pitch features like [+Hi] (Wang 1967) or [Upper] (Yip 1980). It is not, however, immediately apparent how the necessary feature can be defined, given the F_o shapes in Figure 15. In what sense is the pitch of 1, 2, and 5 syllables [+"Hi"]? [+"Hi"] type 1 syllables certainly have a high pitch at onset, but fall to a low pitch; the mean peak F_o value for [+"Hi"] type 2 CǓ syllables is only 5 Hz higher than the mean F_o peak in [-"Hi"] type 6 CLq syllables; only 12 Hz separates the F_o peaks of [+"Hi"] CǓ and [-"Hi"] CĽ types.

If we wished to define this feature acoustically with reference to F_o, two possibilities suggest themselves. The first uses F_o onset values; the second invokes the concept of register as defined on p.11.

Syllable types 3, 4 and 6 have almost identical F_o onset values - in CL(q) syllables, for example, 122, 119 and 122 Hz respectively - and any types with F_o onset value above, say, 40% of total F_o range can be defined as [+Hi], or Yin. (Note that 40% of total F_o range is 131 Hz in CV(q) syllables, and 132 Hz in CVŋ syllables.) However, syllable types, 1, 2, and 5 do not onset at the same F_o value, and in CVŋ syllables, in fact, the F_o onset value for [+Hi] type 2 (CǓŋ) is nearer to the [-Hi] values than the other [+Hi] syllable CÙŋ.

Using the register approach, however, gives an interesting and intuitively satisfying result. The F_o register values of the syllable-types in the present data are as follows:

86%:CUq; 69%:CǓŋ, 64%:CÙ; 53%:CǓ; 52%:CǓŋ; 46%:C_sL̂ŋ, 44%:CL̂ŋ, 41%:C_sL̂, 38%:CL̂; 39%:CLq, 31%:C_sLq; 19%:CĽŋ, 17%:CĽ, 16%:C_sĽ.

These values indicate a demarcation - in acoustic terms at least - at 50% of total F_o range: [+Hi]/Yin syllables in the Ninpo dialects have F_o register values of 50% and above; [-Hi]/Yang syllables have F_o register values of below 50%.

9. DURATION

A constant duration hierarchy obtains for the Finals (i.e. DF measurements) of the syllable-types, viz 4 > 2 > 3 > 1 > 5 > 6. T-tests show that types 5 and 6 are different at the 10% level ($t = 1.76$, $df = 24$), as are also CV types 2 and 3 ($t = 1.71$, $df = 24$). Other types are different at least the 5% level, with the exception of CǓŋ and CĽŋ, which do not differ significantly.

Note that, if DP values are taken to constitute the relevant duration measure in syllable-types 3, 4, and 6, a durational contrast between types 2 and 3 no longer obtains for CǓ vs CL̂, whereas the DP value of CǓŋ syllables still remains significantly different from the DF value of CL̂ŋ syllables ($p < .005$; $t = 3.39$, $df = 14$).

If desired, the durational relationships can be quantified by expressing the duration of individual syllable-types as percentages of the largest duration value, i.e. of type 4. For $C_{(s)}V_{(q)}$ syllables, this gives the following results (the largest duration (DF) value, 36.9 csec, occurs on C_sĽ syllables):

25%:CUq; 28%:CLq, 36%:C_sLq; 70%:CÙ; 80%:CL̂; 82%:C_sL̂;
87%:CǓ; 98%:CĽ, 100%:C_sĽ.

For $C_{(s)}$Vŋ syllables, the values are as follows (the largest duration (DF) value is 36.3 csec):

61%:CÙŋ; 72%:CL̂ŋ; 76%:C_sL̂ŋ; 93%:CǓŋ, 100%:CĽ.

These relativised percentage values fall into two groups. Values less than 40% of reference duration define one specification of another important phonological feature in Sh-Zh, i.e. [-Long], or the traditional Ru category; values of above 40% of reference duration define the opposite specification of the feature.

10. AMPLITUDE

The amplitude shapes for all the examples are shown in Figure 16.

As with the F_o shapes, each syllable-type occupies a definite position in a hierarchy of Ao register values, viz. 5 > 2 > 3 > 1 > 6 > 4. For $C_{(s)}$V(q) syllables, the Ao register values are: 83%:CUq; 72%:CǓ; 64%:C_sL̂, 57%:CL̂; 55%:CÙ; 61%:C_sLq, 52%:CLq; 36%:C_sĽ, 25%:CĽ. For $C_{(s)}$Vŋ syllables, the register values are 70%:CǓŋ; 70%:C_sL̂ŋ; 67%:CL̂ŋ; 65%:CÙŋ; 35%:CĽŋ. Note, however, that the feature [+Hi] cannot be defined in Ao register terms, as was possible with the F_o register. This is because type 3 syllables, which are [-Hi], have higher Ao register values than type 1 syllables which are [+Hi].

In Rose (1979:23), where high-passed amplitude was measured, it was claimed that Ao register was sufficient to provide an additional definition for the [+Long] feature, since syllable-types 5 and 6 had very high Ao register values, separated from the other types by about 40%. With flat amplitude registration, however, it is clear that [+Long] cannot be defined in terms of Ao register, since types 5 and 6 no longer group together. Moreover, the register value of the [-Long] type 5 syllables is only about 10% higher than the [+Long] type 2 syllables.

Within the dimension of Ao contour, there is a contrast between peripherally (CǓ, CĽ, $C_{(s)}$Lq) versus centrally (CÙ, $C_{(s)}$L̂, CUq) located amplitude. Ao contours of CÙ and CǓ could in addition be described as contrasting with $C_{(s)}\hat{L}$ and $C_{(s)}\check{L}$ in the presence vs. absence of an initial Ao peak.

11. INTERPARAMETRIC CORRELATIONS - F_o AND Ao

The type of positive correlation between F_o and Ao registers reported in Zee (1978) for Taiwanese non-Rusheng tones is not as strong in Sh-Zh. For example, although the Sh-Zh syllable-types with the highest and lowest F_o register values (types 5 and 4 respectively) also have the highest and lowest Ao register values, type 1 syllables have F_o registers higher, but Ao registers lower, than type 3 syllables.

Of much greater significance than register correlation is, however, the generally strong similarity in contour between Ao and F_o shapes. This degree of similarity seems to be lacking in other available data on F_o and Ao in tone languages, at least as far as tones with rising F_o are concerned (Sauvain 1977; Phuong 1981; Zee 1978; Li 1971). For example, of the 9 clear examples of tones with rising F_o contours in North, Central, and Southern Vietnamese, (Phuong 1981), there is only one - North Vietnamese tone 6 - which shows a clear increase in Ao corresponding to an increase in F_o. Even in this case, however, the increase is only about 2 dB, and the Ao contour starts to decay well before peak F_o is reached. These facts indicate that tone languages differ considerably with respect to the relative contribution of sub-glottal pressure to production of F_o, and that it is clearly premature to assume (Catford 1977:110; Ohala 1978) that F_o is universally primarily controlled by vocal cord tension (cf. Hombert 1977:18; Zee 1977:117). (It can in fact be shown by applying a version of the technique outlined in Monson's (1978) 'Indirect assessment of the contribution of sub-glottal air pressure and vocal cord tension to changes of fundamental frequency in English' that vocal cord tension plays only a secondary role in the control of F_o in some syllable types in Sh-Zh and Zhenhai dialect (Rose 1982).)

Not only are there clear differences between languages in the control of F_o, however. The present data show that the mode of control of F_o can differ, even within a single language. This can be demonstrated by examining linear changes in absolute mean F_o and Ao over time in different syllable-types. For example, in $C\check{L}$ syllables, in the 8.8 csec between csec 23.7 and csec 32.5, the F_o rises from 116 Hz to 138 Hz. Corresponding to this rise of 22 Hz, the Ao rises 2.6 dB, from 18.1 dB at csec 23.7 to 20.7 dB at csec 32.5.

In CLq syllables, the 22 Hz rise in F_o from 116 Hz to 138 Hz takes 5 csec (from csec 1.5 to csec 6.5), and corresponds to an increase in Ao, from 17.6 to 22.5 dB, of 4.9 dB.

In $C\hat{L}$ syllables, the 22 Hz rise in F_o from 116 Hz to 138 Hz takes 9.5 csec (from csec 1.5 to csec 11), and corresponds to an increase in Ao, from 17.6 to 23.4 dB, of 5.8 dB.

These results can be compared with the corresponding data from types 3, 4, and 6 with sonorant Initial consonants: in $C_s\check{L}$ syllables, a 22 Hz rise in F_o from 115 Hz to 137 Hz takes 10.3 csec (from csec 23 to csec 33.3), and corresponds to an increase in Ao, from 18.4 to 21.6 dB, of 3.2 dB. In C_sLq syllables, a 22 Hz rise in F_o from 115 Hz to 137 Hz takes 9.3 csec (from csec 0 to csec 9.3), and corresponds to an increase in Ao, from 18 to 24 dB, of 6 dB.

In $C_s\hat{L}$ syllables, a 22 Hz rise in F_o from 115 Hz to 137 Hz takes 11.5 csec (from csec 0 to csec 11.5), and corresponds to an increase in Ao, from 16.5 to 23.9 dB, of 7.4 dB.

The linear rates of change of F_o and Ao with respect to time derived from the above data are given in Table 17 below.

TABLE 17

COMPARISON BETWEEN LINEAR RATES OF CHANGE OF F_o AND Ao WITH RESPECT TO TIME IN SYLLABLES OF TYPE 3, 4, AND 6.

Type		$\Delta F_o/\Delta t$ (Hz/csec)	$\Delta dB/\Delta t$ (dB/csec)
4	$C_s\check{L}$	2.1	0.3
	$C\check{L}$	2.5	0.3
3	$C_s\hat{L}$	1.9	0.6
	$C\hat{L}$	2.3	0.6
6	C_sLq	2.4	0.6
	CLq	4.4	1.0

It can be seen from Table 17 that, with the exception of CLq syllables, the rate of change of F_o with respect to time is about the same ($\bar{x}$ = 2.3 Hz/csec) for types 3, 4, and 6. However, there is a clear difference between type 4 syllables on the one hand, and types 3 and 6 on the other, in the rate of change of Ao with respect to time: in types 3 and 6 it is twice the value of that in type 4, i.e. 0.6 dB/csec vs 0.3 dB/csec. (Note that the $\Delta F_o/\Delta t$ value of 4.4 Hz/csec for CLq syllables is (2.4/4.4 =) 0.5 times greater than for C_sLq syllables. If the $\Delta dB/\Delta t$ value of 1 for CLq syllables is adjusted by this amount, the value of 0.5 dB/csec is obtained which is almost the same as the value of 0.6 for CLq and type 3 syllables).

It seems, therefore, that since the rate of change of F_o with respect to time is nearly constant for all the examples, the differences in rates of change of Ao with respect to time between type 4 syllables on the one hand and types 6 and 3 on the other must reflect a difference in the amount of sub-glottal pressure being used by the speaker in these types. Why should this be?

The observed differences probably reflect the differential sensitivity of phonation registers to changes in sub-glottal pressure (Lieberman 1977:89, 90), and the physiological incompatability between the phonatory setting for whispery voice and the use of vocal cord tension to increase F_o.

In type 4 syllables, the 22 Hz rise in F_o examined occurs on the latter part of the syllable (from about 60 to 90%DF), where phonation is normal. Over this particular part of the syllable, vocal cord tension contributes slightly more to the increase of F_o than sub-glottal pressure (Rose 1982).

In syllables of type 3 and 6, however, the 22 Hz rise in F_o occurs near the onset of the syllable, where the vocal cords are adjusted to produce whispery voice (p.4-5). It is clear that increasing the F_o by increasing vocal cord tension is incompatible with the phonatory setting for whispery voice, which involves "...somewhat relaxed vocal cords..." (Catford 1977:101). Hence the obligatory initial rise in F_o on syllable-types 3 and 6 can only be achieved by means of sub-glottal pressure increases. The degree of sub-glottal pressure increase needs to be greater than in type 4 syllables because the contribution of vocal cord tension is excluded, and because, presumably, the vocal cords when set-up for whisper are less sensitive to sub-glottal pressure changes than when in normal phonatory mode.

It is also likely that, in accordance with the Motor Theory of Speech Perception, the auditory impression of 'energetically concentrated loudness' noted for type 3 syllables above (p.4) is referable - via the acoustic cue of Ao - to the abruptly increased sub-glottal pressure over the first half of these syllables.

From the above discussion, it can be seen that the production of syllable-types in Sh-Zh involves deliberate and considerably differentiated control over sub-glottal pressure. This has interesting implications for diachronic Chinese tonology, since it provides a tonal feature other than F_o/pitch, i.e. sub-glottal pressure, which characterises syllable-types 3 and 6 against type 4. From Middle Chinese, voiced occlusives developed differently into the modern dialects, depending on the syllable-type. Thus in Cantonese and Mandarin, aspirated reflexes of Middle Chinese voiced occlusives are found in syllable-types cognate with Sh-Zh syllable-type 3, (and in some cases syllable-type 6), but unaspirated reflexes are found in syllable-types cognate with Sh-Zh syllable-type 4. It is therefore possible that the differential development of the voiced occlusives reflects a syllable-type contrast in Middle Chinese similar to that in modern Sh-Zh, especially since aspiration has been found to correlate in some languages with increased sub-glottal pressure.

12. INTERPARAMETRIC CORRELATIONS - F_0 AND DURATION

In recent studies, an intrinsic correlation between syllable duration and F_0 contour has been posited. Ohala (1978:31) notes that speakers appear to be able to produce falling F_0 contours faster than rising ones, which explains why tones with rising F_0 have a longer duration than tones with falling F_0. Also, correlations have been reported between duration and F_0 register, and duration and F_0 range: in syllables with falling F_0, duration decreases with F_0 register, whereas in syllables with rising F_0, duration depends on the range of F_0 covered - the smaller the range, the shorter the duration (Hombert 1977:15).

To what extent do these correlations hold for the Sh-Zh data? The duration (DF) hierarchy in $C_{(s)}V(q)$ syllables, repeated here from p.38, is 4 < 2 < 3 < 1 < 6 < 5. The duration contrast between the [+Long] (4,2,3,1) and [-Long] (5,6) syllables is, of course, extrinsically controlled by the speaker.

The durational relationship between types 4 and 2 appears to be an example of a positive correlation between duration and F_0 range, and is therefore a plausible extension of the previously noted correlation for syllables with rising F_0 : syllables of type 4 and type 2 both have the same concave F_0 contour, but the mean F_0 range in type 4 syllables is almost twice that of type 2 syllables, viz. 36 vs. 19 Hz respectively.

There is nothing exceptional about the difference in duration between type 1 and types 4 and 2 : if syllables with falling F_0 contours are intrinsically shorter than those with rising contours, then they are bound to be shorter than those with falling-rising contours. The same reasoning applies to the duration contrast between type 1 and type 3 syllables.

There remain, however, two cases which do not exhibit the expected durational relations. Type 3 syllables, with rising-falling F_0 contours, are shorter than types 2 and 4, which have falling-rising contours, and type 6 syllables, with rising F_0 contours, are shorter than type 1 syllables, which have falling contours.

Now a convex F_0 contour, like that of Sh-Zh type 3 syllables, is very rare in Chinese dialects: of the 3433 F_0 shapes investigated in a survey of Chinese dialect tones, only 80 were found to have such a contour (Cheng 1973:103). In current terminology, then, the F_o contour of Sh-Zh type 3 syllables would be considered, as for example in Wang (1967), highly marked and (in some unspecified way) 'complex'.

However, if the type 3 syllable F_o contour is more complex than the concave contours in types 2 and 4, we should surely expect it to have a longer duration, applying the inference from relative duration to relative complexity of articulation implied in Ohala (1973:15). (Note that in Cheng's (1973) survey, 352 of the 3433 F_o shapes were concave).

Clearly, in this case, the idea of relative duration of an F_o contour as an indicator of complexity is incompatible with the idea of complexity of an F_o contour being reflected in its frequency of occurrence/markedness. I think this incompatibility can be resolved by recalling (p.42) that the mode of control of F_o in type 3 syllables (with convex contour) differs from that in type 4 (with concave contour). In type 3 syllables, the participation of vocal cord tension in the control of F_o is ruled out by the phonatory setting for whispery voice, so that sub-glottal pressure has to be used. If it is possible that a convex F_o contour can be produced quicker by a burst of sub-glottal pressure (akin to the effect of the push-in-the-chest technique described in Ladefoged (1967)) than by changes in vocal cord tension, the shorter duration of type 3 syllables relative to types 4 and 2 would be explained.

The question of mode of F_o control is also of relevance to the problem of the durational relationship between types 6 and 1. Normally, we would expect type 1 syllables, with falling F_o contours, to be shorter than type 6 syllables, with rising F_o contours. Therefore, the fact that type 1 syllables have a much greater duration than type 6 syllables requires comment.

It will be recalled that type 6 syllables are [-Long], whereas type 1 syllables are [+Long] (p.39). Since this phonological feature is primarily manifested by extrinsically controlled relative duration, type 6 syllables must be kept shorter than type 1 syllables. This, however, only explains half of the problem, because putative intrinsic relationships, like the correlation between F_o contour and duration, should not by definition be phonologically suspendible. There must, therefore, be an additional factor present which enables the rising F_o required for type 6 syllables to be produced quicker than the falling F_o required for type 1 syllables. In the light of the results reported above, this factor seems once again to be the mode of F_o control in type 6 syllables. As in type 3 syllables, a more abrupt increase in F_o is achieved by a burst of sub-glottal pressure, rather than increased vocal cord tension.

To summarise, in order to explain differences in duration between the Sh-Zh [+Long] syllables as intrinsically correlated with their F_o characteristics, it is necessary to consider, in addition to the factors of F_o range and contour, their mode of F_o control. The mode of F_o control must also be considered in accounting for the absence of an expected correlation between F_o contour and duration in types 1 and 6.

13. EFFECT OF SYLLABLE-FINAL VELAR NASAL

All syllable-types capable phonotactically of showing the final velar nasal, i.e. types 1, 2, 3, and 4, were analysed. The results are shown graphically in Figures 2 (CÙŋ), 4 (CǓŋ), 6 (CL̂ŋ), 8 (C_sL̂ŋ) and 10 (CĽŋ), where 'N' marks the onset of the velar nasal. Many interesting and perhaps unexpected correlations exist between the final nasal and all prosodic parameters. Note, for example, that the presence of a syllable-final nasal appears to correlate with a significantly higher F_o at syllable onset in type 1 syllables. I shall restrict my comments here, however, to correlations between syllable-final nasal and Ao, and between duration and point of onset of nasal consonant.

14. FINAL NASAL AND F_o

The effect of a final nasal on Ao can be seen from a comparison between the Ao shapes on $C_{(s)}$V and $C_{(s)}$Vŋ syllables in Figure 16.

Probably the most obvious effect is the slightly lower Ao values, relative to CV syllables, which are found on that part of the Ao shape corresponding to the nasal consonant. In CǓŋ syllables, for example, the Ao rise over the nasal is 1.7 dB compared to 2.3 dB in CǓ syllables, and in CĽŋ syllables the Ao rise is 4.1 dB, compared to values of 5.7 dB and 4.8 dB in CĽ and C_sĽ syllables respectively. Because of their relatively large, acoustically absorbent surface area, the nasal passages have a high damping factor, and therefore an intrinsic difference in amplitude between nasal and non-nasal sounds is to be expected (Ohala 1975:292).

A drop in Ao caused by a syllable-final nasal has been noted for very different Chinese dialects, namely Yangzhou (Sauvain 1977:203), Cantonese (Fok 1974:23, see also Figures 2, 3, pp.139, 140), and Peking dialect (Kratochvil 1977:29). Unfortunately, the type of amplitude involved - i.e. whether flat or pre-emphasised - is usually not made clear, and the differences not quantified. This is a regrettable omission, because a drop in Ao can have perceptual significance : in casual Peking dialect speech, where nasal codas are often dropped, the morphophonemic presence of the nasal is still signalled by a mid-syllable drop in Ao, and also, interestingly, in F_o (Kratochvil 1971). Moreover, a syllable-final nasal can be synthesised merely by abruptly decreasing Ao in mid-syllable (Kratochvil 1977). Now, it is clear that the flat Ao differences between some CV and CVŋ syllables in the second half of their duration are unlikely to be above the difference limen for the perception of amplitude[8], whereas the pre-emphasised amplitudes of Sh-Zh CV and CVŋ syllables show much greater contrastivity (Rose 1979). This invites the speculation that the perceptual mechanism involves some kind of high-pass filtering in order to be able to detect amplitude drops *qua* cues for final nasals.

A second characteristic of syllables with final nasals is the higher Ao values (relative to CV syllables), registered just after the onset of the Final. For example, the Ao value of CÙŋ syllables at 10%DF is 25.3 dB, 3.2 dB up on the corresponding Ao value of 22.1 dB for CÙ syllables. In type 2 syllables the difference is 1.7 dB, and in type 4 syllables, 2.6 dB. For CÙŋ syllables, at least, this difference in Ao seems to reflect extrinsic control, possibly in order to maximise contrastivity in Ao between the first and second halves of the syllable, and thus enhance the function of Ao as a perceptual cue for nasals (Rose 1982). Note also in this connection that, in the absence of compensatory adjustments in vocal cord tension, we should expect to find an increase in F_o concomitant with such a deliberate increase in Ao. This is, of course, exactly what is observed in the higher F_o on CÙŋ (and CL̂ŋ) syllables (p.33).

Finally, note that most of the drop in Ao in syllables with final nasals occurs on the vowel prior to the onset of the nasal consonant. This is probably an indication both of early anticipatory nasalisation of the short vowel preceding the [ŋ], and (later) of narrowing of the vocal tract caused by movement of the back of the tongue towards velaric occusion.

15. NASAL ONSET POINT AND DURATION

Examination of the measurements for the onset of the nasal consonant shows that the value ^{N}/DF in CÙŋ and CL̂ŋ syllables, and ^{N}/DP in CǓŋ and CĽŋ syllables is a constant 40%. The speaker appears, therefore, to be timing the onset of the nasal consonant with reference to the duration characteristics of the syllable in which it occurs, so that velaric occlusion is effected at 1/3rd of the relevant duration (DF in CL̂ŋ, and CÙŋ syllables; DP in CǓŋ and CĽŋ syllables). This means, of course, that the proportions of short vowel (1/3rd) and velar nasal 2/3rd) are maintained constant within the syllable-type while their absolute duration values will differ. For example in CÙŋ syllables the short vowel and [ŋ] have mean durations of 9.2 csec and 13 csec respectively, compared to 13.4 csec and 22.9 csec in CĽŋ syllables. It is also interesting to note that it is the duration to F_o peak in CǓŋ and CĽŋ syllables, rather than the duration of the Final - or duration to Ao peak - which appears to be the relevant measure for the timing of this articulatory pattern.

16. EFFECT OF INITIAL SONORANT

A comparison of the F_o shapes of syllables having a sonorant Initial in Figure 16(C) with those in Figure 16(A) and (B) having obstruent Initials, gives a clear illustration of the different effect on F_o onset value and contour caused by the sonority feature. In syllables of type 3, 4, and 6 with (voiceless 'lenes') obstruents, there is an initial drop in F_o after Final onset. The size of the drop in F_o, and the duration of this consonantally induced perturbation depends upon the syllable-type. In types 3 and 6, the drop is 6 Hz and the effect lasts from 1 to 3 csecs; in type 4 syllables, the drop is much greater: 14 Hz (CĽ) and 26 Hz (CĽŋ), and the perturbation lasts longer too - approximately 12 csec.

The initial drop in F_o after voiceless obstruents in syllable-types 3, 4, and 6 contrasts markedly with the flat F_o onset contour on syllables with a sonorant Initial. In these syllables, the F_o at Final onset is already at a low value, similar to the value of the F_o inflexion point in syllables with non-sonorant Initials.

Comparisons like the one above allow us to separate those acoustic features which are characteristic of a syllable-type from those which are attributable to other independently variable features, like differences in the nature of the Initial consonant. Thus the initial fall in F_o observable in CL($^{q}_{ŋ}$) syllables appears to be due to articulatory and/or aerodynamic features connected with the Initial consonant, and is not a direct characteristic of the F_o contour of type 4 syllables.

Although the duration of the voiced part of the syllable is considerably increased by the addition of a sonorant Initial, the duration of the Final remains unaffected, i.e. it retains the same absolute and relative duration values as Finals in syllables without sonorant Initials. For example, the DF value of 29.6 csec in type 3 syllables without sonorant Initials (CL̂) does not differ significantly from that of 30.1 csec in type 3 syllables with sonorant Initials (C_sL̂); c.f. the DF values in CL̂ŋ and CL̂ŋ syllables of 26.3 and 27.5 csec respectively and in $C\check{L}$ and $C_s\check{L}$ syllables of 36.3 and 36.9 csec respectively. Type 6 syllables, for some as yet obscure reason, are an exception: the 2.4. csec difference in Final duration between CLq and C_sLq syllables is significant at the 5% level ($.05 > p > .02$; $t = 2.23$, $df = 18$).

The results given above agree with those reported for Yangzhou (Sauvain 1977) and MSC (Howie 1974), and hence provide additional evidence that duration, F_0, (and also Ao) characteristics of Initial sonorants can in most cases be excluded from tonetic domain in Chinese.

17. SUMMARY

In the above investigation, I have attempted to give some idea of the complexity in the physical reality underlying citation syllable-type contrasts in a Chinese dialect. I have also tried to illustrate the potential of, and necessity for, a polydimensional approach to tonal investigation *contra* the prevalent monodimensional stance which ignores from the outset all parameters except F_0/pitch variation (Kratochvil 1977:22). I have used acoustic data to demonstrate that Sh-Zh syllable-types are characterised in at least the three acoustic parameters of F_0, Ao, and duration, and that the physiological mechanisms which produce the acoustic effects are differentially controlled, depending on the syllable-type. I have illustrated some proposals for quantifying the relative contributions of each parameter, including definitions of register and contour dimensions within the F_0 and Ao parameters, and have also shown how the data might be used in defining the phonological features [Hi] and [Long] in Sh-Zh.

NOTES

1. An earlier version of this paper - Rose (1979) - was circulated at the 12th International Conference on Sino-Tibetan Languages and Linguistics in Paris. The present paper differs from the earlier version mainly in the larger number of examples analysed per sample and in the measurement of flat, as opposed to high-passed amplitude. I am grateful to Paul Kratochvil for criticisms and suggestions.

The data quoted are from my own field work on the dialects of the Ningpo area (Ningpo town, Cixi, Zhenhai, Dinghai, Fenghua, and Xiangshan) carried-out over the past seven years with native speakers in Hong Kong, Taiwan, Shanghai, and Manchester. More extensive data on Zhenhai dialect are contained in Rose (1982).

2. For some details of the past and present sociolinguistic situation in the Shanghai area, see Sherard (1972:3,5) and Hu (1978). Phonological interaction between the Shanghai and Zhenhai dialects is complex. For the purposes of this paper it is sufficient to note that Shanghai dialect influence is restricted to segmentals. Sh-Zh speakers normally retain characteristic Zhenhai suprasegmental sounds and rules, which are very different from those of Shanghai dialect.

3. 'Initial' and 'Final', when written with capitals, are the conventional translations of the traditional Chinese phonological terms for the immediate segmental constituents of the syllable. Initial refers to the syllable-initial consonant, and Final to the rest of the syllable. Thus in the syllables [ti] and [dʑɥɔ̆ŋ], t and dʑ are Initials, and i and ɥɔ̆ŋ are Finals.

4. Loudness, pitch and length have not been transcribed.

5. For typographical convenience, the symbol [ɿ] has been chosen to represent the Sh-Zh unrounded non-retroflex apical vowel.

6. The 'pitch' detection algorithm I used is a software equivalent of the hardware method presented in Dubrowski et al. (1975), which is a modified autocorrelation method using clipping. My thanks to Geoff Bristow, Department of Engineering, Cambridge University, and Francis Nolan, Department of Linguistics, Cambridge University, for their help in making the program available and running it for me.

7. Cf. Cheng (1972:289) "...this degree of accuracy [$\pm$ 4 Hz] is more than one can ask for when measuring from a narrow band spectrogram..."

8. It is very doubtful whether the difference limen in overall dB level for this kind of natural speech within this range of fundamental frequencies and levels is smaller than 1 dB. At levels of 60 or 70 dB above threshold, the just noticable difference (JND) for normal listeners is less than 0.5 dB for a typical psychophysical stimulus tone of 1000 Hz, but for a synthetic vowel at an obviously much lower F_0, the JND is apparently $\pm$ 1 dB (Lehiste 1970:115, 116).

BIBLIOGRAPHY

CATFORD, J.C.

1977 *Fundamental problems in phonetics.* Edinburgh University Press.

CHAO, Yuen Ren

1928 *Studies in the modern Wu dialects.* Tsing Hua College Research Institute Monograph No.4. Peking.

1930 ə sɩstəm əv "toun-lɛtəz". *Le maître phonétique* 30:24-27.

1967 Contrastive aspects of the Wu dialects. *Language* 43/1:92-101. Baltimore.

CHENG, Chin-chuan

1973 A quantitative study of Chinese tones. *Journal of Chinese Linguistics* 1/1:93-110.

CHENG, T.M.

1972 Review of F.M. Weingartner's *Tones in Taiwanese: an instrumental investigation. Lingua* 30:285-294.

DUBROWSKI, J.J., R.W. SCHAFER and L.R. RABINER

1976 Real-time digital hardware pitch detector. *Institute of Electrical and Electronic Engineers transactions on acoustics, speech, and signal processing* 24/1:2-8.

FANT, Gunnar

1968 Analysis and synthesis of speech processes. In Malmberg, ed. 1968:173-277.

FERGUSON, Charles A., Larry M. HYMAN and John J. OHALA, eds

1975 *Nasålfest: papers from a symposium on nasals and nasalisation.* Stanford University Press.

FOK, Chan Yuen-yuen

1974 *A perceptual study of tones in Cantonese.* Publications of the Centre of Asian Studies of the University of Hong Kong, No.18. Hong Kong.

FROMKIN, Victoria A., ed.

1978 *Tone: a linguistic survey.* New York: Academic Press.

HOMBERT, J.M.

1976 Perception of tones of bisyllabic nouns in Yoruba. *Studies in African Linguistics.* Supplement 6:109-121.

1977 Difficulty of producing different F_o in speech. *UCLA working papers in phonetics* 36:12-19.

1978 Consonant types, vowel quality, and tone. In Fromkin ed., 1978:77-111.

HOUSE, Arthur S.

1959 A note on optimal vocal frequency. *Journal of Speech and Hearing Research* 2/1:55-60.

HOWIE, J.M.

1976 *Acoustical studies of Mandarin vowels and tones.* Cambridge: Cambridge University Press.

HU Ming-yang

1978 Some changes in Shanghai dialect during the last hundred years. *Zhongguo Yuwen* 3:199-205.

KRATOCHVIL, Paul

1971 An experiment in the perception of Peking dialect tones. In Inga-Lill Hansson, ed. *A symposium on Chinese grammar,* 7-40. Scandinavian Institute of Asian Studies Monograph Series No.6. London: Curzon Press.

1977 Traditions in Chinese linguistics - fact or fiction? *Cahiers de linguistique asie orientale* 1/1:7-40.

LADEFOGED, Peter

1967 *Three areas of experimental phonetics.* London: Oxford University Press.

LEHISTE, Ilse and Gordon E. PETERSON

1959 Vowel amplitude and phonemic stress in American English. *JASA* 31/4:428-435. Reprinted in Lehiste ed., 1967:183-190.

LEHISTE, Ilse, ed.

1967 *Readings in acoustic phonetics.* Cambridge, Mass.: MIT.

LI, Lin-nei

1971 An instrumental study of Mandarin and Cantonese tones. Ph.D. dissertation, Australian National University, Canberra.

LIEBERMAN, Philip

1977 *Speech physiology and acoustic phonetics.* New York: Macmillan.

LISKER, Leigh and A.S. ABRAMSON

1969 A cross-linguistic study of voicing in initial stops: acoustical measurements. *Word* 20/3:384-422.

LIU, Fu

1925 *Etude expérimentale sur les tons du chinois.* Collection de l'Institut de Phonétique et des Archives de la Parole, Fascicule 1. Paris and Peking.

MALMBERG, Bertil, ed.

1968 *Manual of phonetics.* Amsterdam: North-Holland.

MILLER, G.A.

1948 The perception of short bursts of noise. *JASA* 20:160-170.

MONSEN, Randall B., A. Maynard ENGEBRETSON and N. Rao VEMULA

1978 Indirect assessment of the contribution of sub-glottal air pressure and vocal cord tension to changes of fundamental frequency in English. *JASA* 64/1:65-80.

OHALA, John J.

1975 Phonetic explanations for nasal sound patterns. In Ferguson et al., eds, 1975:289-316.

1978 Production of tone. In Fromkin, ed., 1978:5-39.

PAINTER, Colin

1979 *An introduction to instrumental phonetics.* Baltimore: University Park Press.

ROSE, Philip John

1974 The phonology of the Ningpo dialect. M.A. thesis, University of Manchester.

1979 Tones in the Wu dialect of Zhenhai: an acoustic description. Paper circulated at the 12th International Conference on Sino-Tibetan Languages and Linguistics, Paris.

1982 An acoustically based phonetic description of the syllable in the Zhenhai dialect. Dissertation submitted for the Ph.D. degree, University of Cambridge.

SAUVAIN, D.L.

1977 An acoustical analysis of selected Yangchow syllables. Dissertation submitted for the Ph.D. degree, Cambridge University.

SHERARD, Michael

1972 Shanghai phonology. Ph.D. dissertation, Cornell University.

SHI Wentao

1979 Etymological notes on the Ningbo dialect. *Fangyan* 3:161-170.

SMALL, A.M., Jr, BRANDT, J.F. and P.C. COX

1962 Loudness as a function of signal duration. *JASA* 34:513-514.

SOKOLOV, M.V.

1965 Eksperimental'noe issledovanie tonov shankhajskogo dialekta [An experimental investigation of Shanghai dialect tones]. *Phonetica* 12:197-200.

TCHEN, Ting-Ming

1938 *Etude phonétique des particules de la langue Chinoise.* Paris.

VŨ THANH PHƯỜNG

1981 Acoustic and perceptual nature of tone in Vietnamese. Ph.D. thesis, Australian National University.

WANG, Wm. S-Y.,

1967 Phonological features of tone. *International Journal of American Linguistics* 33/2:93-105.

YIP, Moira Jean W.

1980 *The tonal phonology of Chinese.* Indiana University Linguistics Club.

YUAN Jiahua et al.

1960 *Hanyu Fangyan Gaiyao* [Survey of Chinese dialects]. Peking: Wenzi Gaige Chubanshe.

ZEE, Eric

1978 Duration and intensity as correlates of F_o. *Journal of Phonetics* 6:213-220.

ZEE, Eric and Ian MADDIESON

1979 Tones and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *UCLA working papers in phonetics* 45:93-129.

ZEMLIN, Willard R.

1968 *Speech and hearing science anatomy and physiology.* New Jersey: Prentice-Hall.

PHONETIC PROPERTIES OF VIETNAMESE TONES ACROSS DIALECTS

Vũ Thanh Phương

0. INTRODUCTION[1]

Although Vietnamese tones have been extensively studied in quite a few works, the more detailed phonetic descriptions based on instrumental analysis have all concentrated on Northern Vietnamese (henceforward NV), e.g. Lê Văn Lý (1948), Andreev and Gordina (1957), Nguyễn Hàm Dương (1962), Han (1969), Han and Kim (1972) and Earle (1975), while Central Vietnamese and Southern Vietnamese (henceforward CV and SV) have had fewer and mainly impressionistic descriptions, with the exception of Trần Huong Mai (1969) which was only partially based on instrumental records. Mine is an attempt to provide a more comprehensive description of Vietnamese tones by presenting data from all three major dialects in their various aspects.

1. AN OVERVIEW OF VIETNAMESE TONES

The official spelling recognises six tones in Vietnamese, which represent what can be termed the underlying phonological tones of standard literary Vietnamese and also of NV, which is regarded as a prestige dialect. Table 1 summarises the system in three dialects. The English labels, taken from Han (1969) and preferred to others because they are short and suggestive of the basic contours of each tone, and the phonological notations, taken from official spelling diacritics with the addition of the macron for the level tone, will be used throughout this work. The phonetic notations, a modified version of Chao's (1930), was first based on auditory impressions and later readjusted in some cases by taking pitch values calculated from the data through various normalisation and conversion procedures described elsewhere (Vũ Thanh Phương 1981).

Phuong, V.T. "Phonetic Properties of Vietnamese Tones Across Dialects". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation*.
A-62:55-76. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.55*

TABLE 1

THE TONES OF VIETNAMESE

Number	1	2	3	3B	4	4B	5	6
Vietnamese Names	ngang	huyền	sắc	sắc (tắc)	nặng	nặng (tắc)	hỏi	ngã
English Labels*	level tone	falling tone	rising tone	stopped rising tone	drop tone	stopped drop tone	curve tone	broken tone
Phonological Notations*	/-/	/ˋ/	/ˊ/	/ˊs/**	/./	/.s/**	/ˀ/	/˜/
Phonetic Notations*								
NV	[33]	[21]	[35]	[45s]**	[21]**	[21s]**	[212]	[325]**
CV	[55]	[42]	[24]	[34s]**	[31]	[31s]**	[312]**	
SV	[33]	[21]	[35]	[35s]**	[212]	[21s]**	[214]	
Examples	/hāj/	/hàj/	/háj/	/hát/	/hạj/	/hạt/	/hảj/	/hãj/
	'two'	*'slipper'*	*'to pick (fruit)'*	*'to sing'*	*'harm'*	*'grain'*	*'sea'*	*'scared'* (in compounds only)

* See comments in the text.

** s represents the syllable-final voiceless stop which conditions the occurrence of the tone.

̳ marks the laryngealisation characteristic of the tone.

2. PROCEDURES

2.1. *INFORMANTS*

This study was based on the recorded voices of thirty-four native speakers of Vietnamese (11 NV, 12 CV and 11 SV), whose home towns are indicated on Map 1 (p.58). They included 14 females and 20 males, respectively represented by F and M and numbered in increasing order in the southward direction within each sex group and each dialect. Being mostly university students or staff, they spoke an educated and standardised variety of their respective dialects.

2.2. *WORD LISTS*

In order to pinpoint dialectal variations of tones in similar phonetic environments, I decided on a restricted number of syllables in two word lists. One consisted of five syllables

/ta/ /t^ha/ /da/ /na/ /sa/

occurring with all the six tones (five in CV and SV), and

/tak/ /t^hak/ /dak/ /nak/ /sak/

occurring with the two stopped tones.

The other consisted of the syllables /ta/ (for non-stopped tones), and /tak/ (for stopped tones), each repeated three times after a frame sentence.

The idea was to minimise possible perturbations caused by consonants and vowels of various types which might differ phonetically in the three dialects. Comparison with data from Han (1969) and Earle (1975), which included a greater variety of syllables, showed that the tone shapes obtained from my material were essentially the same as theirs.

2.3. *INSTRUMENTS*

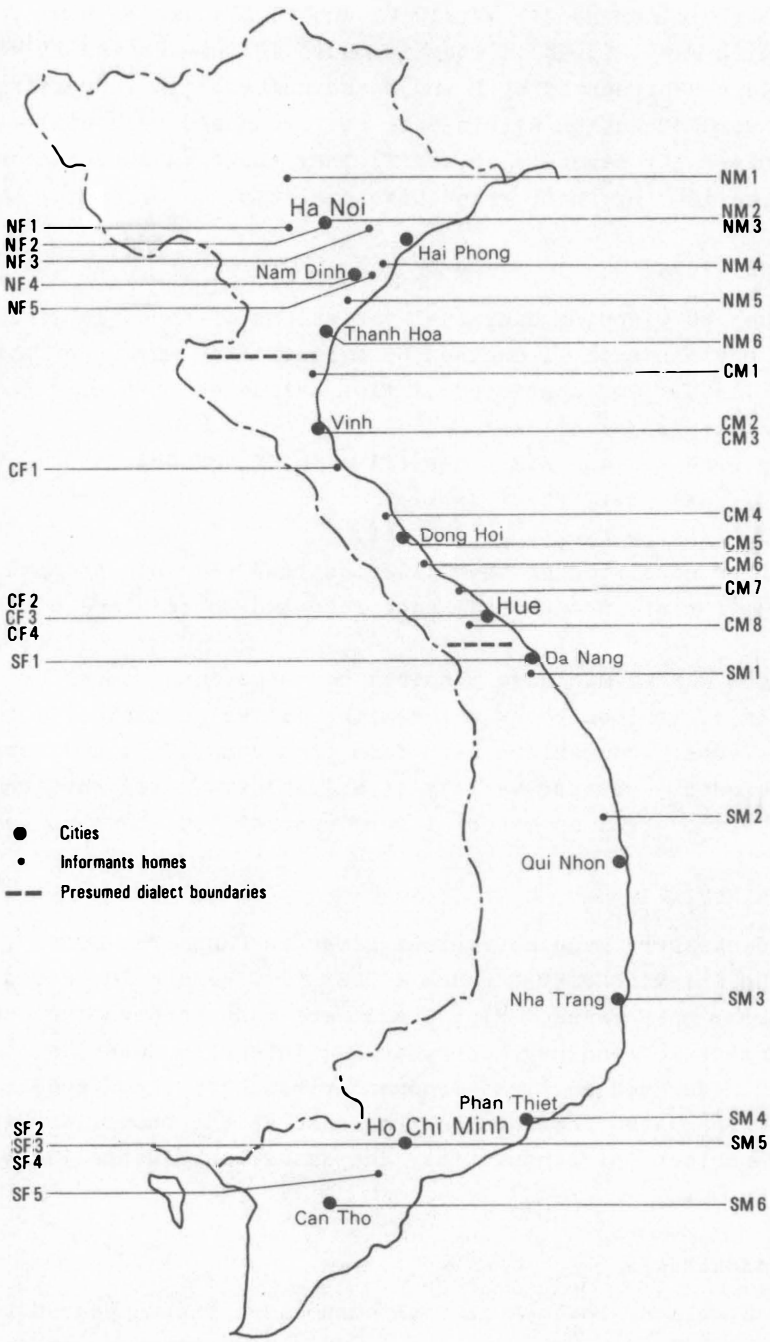
Recordings were made at various times in Canberra, Sydney, Hanoi, Hue and Ho Chi Minh City through a UHER 4000 Report IC Recorder with reel-to-reel BASF tapes. Mingograms were made through the use of an F-J Fundamental Frequency Meter, an F-J Intensity Meter, a Sony 8-Channel Mixer and an Elema-Schonander Mingograf, and spectrograms were made from a Voiceprint Spectrograph, at the Phonetics Laboratory of the Department of Linguistics, The Faculties, Australian National University.

2.4. *MEASUREMENTS*

After I decided to look at four parameters that appeared to characterise Vietnamese tones, namely Fo, intensity, duration and

MAP 1

DISTRIBUTION OF INFORMANTS' NATIVE PLACES IN VIET NAM



laryngealisation, measurements were made manually from mingograms, supplemented by spectrograms only in cases of unclear traces.

Fo values were measured at six timepoints for non-stopped tones designated P1, P2, P3, P4, P5, and P6, corresponding respectively to vowel onset, 10% (after onset), 37% (midpoint 1), 63% (midpoint 2), 90% (before endpoint) and endpoint. For the stopped tones, P1, P2, P3 and P4 were defined as onset, midpoint 1, midpoint 2 and endpoint, because with their much shorter duration, values at 10% after onset and before endpoint would not alter their Fo contours. For syllables beginning with voiced consonants, vowel onsets started later as the Fo values for consonant onsets were ignored, because they could cause deviations from the typical Fo contours of the tones.

Intensity was measured at four timepoints, I1, I2, I3 and I4, corresponding to P2, P3, P4 and P5 of the Fo measurements, and on the same syllables /ta/ and /tak/ for the various tones.

Duration was measured in centiseconds and the determination of onset and endpoint was made in the same way as for Fo and intensity.

As I know of no tested method of measuring laryngealisation, I studied the tones auditorily, noted the occurrence of breathy voice, creaky voice and glottal closure in various tones, and referred back to places in mingograms and spectrograms where they were supposed to occur and measured their rough duration in centiseconds.

3. RESULTS

The Fo, intensity and duration values obtained from measurements described above were treated statistically by calculating the arithmetical means and standard deviations. They are given in tabular forms in the following pages. Tables 2 to 4 present the Fo data, Tables 5 to 7 the intensity data, Table 8 the duration data and Table 9 the laryngealisation data.

TABLE 2

MEAN Fo IN HERTZ AND STANDARD DEVIATIONS OF NV TONES AT SIX TIMEPOINTS (FOUR TIMEPOINTS FOR STOPPED TONES)*

Tone	n	P1	P2	P3	P4	P5	P6
Level /-/	72	212	212	212	210	207	202
		55	56	55	55	54	53
Falling /ˋ/	72	178	175	171	163	159	155
		45	46	45	43	44	43
Rising /ˊ/	72	198	194	197	217	246	257
		52	51	51	51	54	56
St. Ris. /ˊs/	72	221	232	250	268		
		61	59	57	60		
Drop /./	72	189	186	180	165	162	163
		47	47	45	43	44	47
St. Drop /.s/	72	182	174	167	158		
		45	43	41	43		
Curve /ʾ/	72	176	170	157	140	155	166
		45	46	42	37	38	38
Broken /˜/	72	202	196	169	211	244	245
		55	54	47	55	55	58

*Data from nine NV informants: NF1, NF2, NF3, NF4, NM1, NM2, NM3, NM4 and NM6.

For each tone, mean Fo values on first line, SD on second line.

TABLE 3

MEAN Fo IN HERTZ AND STANDARD DEVIATIONS OF CV TONES
AT SIX TIMEPOINTS (FOUR TIMEPOINTS FOR STOPPED TONES)*

Tone	n	P1	P2	P3	P4	P5	P6
Level /-/	96	184	182	185	187	188	186
		53	50	50	50	52	52
Falling /ˋ/	96	174	171	167	162	159	156
		49	48	47	46	45	44
Rising /ˊ/	96	158	153	150	162	177	181
		43	39	37	37	47	52
St. Ris. /ˊs/	96	163	159	166	179		
		42	39	39	46		
Drop /./	96	163	158	153	148	148	149
		42	40	38	39	39	40
St. Drop /.s/	96	169	161	155	150		
		43	39	39	39		
Curve /ʾ/	96	166	163	156	148	153	156
		45	43	41	42	46	48

*Data from twelve CV informants: CF1, CF2, CF3, CF4, CM1, CM2, CM3, CM4, CM5, CM6, CM7 and CM8.
For each tone, mean Fo values on first line, SD on second line.

TABLE 4

MEAN Fo IN HERTZ AND STANDARD DEVIATIONS OF SV TONES AT SIX TIMEPOINTS (FOUR TIMEPOINTS FOR STOPPED TONES)*

Tone	n	P1	P2	P3	P4	P5	P6
Level /-/	72	191 61	191 60	192 61	191 60	189 61	185 61
Falling /\`/	72	166 52	161 52	155 51	151 50	149 51	149 54
Rising /'/	72	194 63	192 60	201 61	224 68	250 80	255 84
St. Ris. /'s/	72	201 67	208 66	228 71	256 86		
Drop /./	72	166 57	157 52	148 49	150 51	162 55	166 56
St. Drop /.s/	72	170 56	159 52	155 49	160 52		
Curve /'/	72	173 61	162 56	149 48	175 54	219 66	224 71

*Data from nine SV informants: SF1, SF2, SF3, SF5, SM1, SM2, SM4, SM5, and SM6.

For each tone, mean Fo values on first line, SD on second line.

TABLE 5

MEAN INTENSITY IN dB AND STANDARD DEVIATIONS OF NV TONES ON SAME SYLLABLES AT FOUR TIMEPOINTS*

Tone	n	I1	SD	I2	SD	I3	SD	I4	SD
Level /-/	32	7.1	2.1	6.3	1.1	5.4	1.6	2.0	1.8
Falling /ˎ/	32	5.1	2.1	4.5	1.5	3.5	1.7	1.1	0.2
Rising /ʹ/	32	5.5	1.5	4.6	1.3	5.4	2.5	1.6	1.1
St. Ris. /ʹs/	32	5.6	2.0	5.0	1.6	4.5	2.1	1.5	1.3
Drop /./	32	4.5	1.3	4.1	1.5	2.3	1.0	0.8	0.2
St. Drop /.s/	32	5.0	2.1	4.4	1.8	3.6	1.6	1.0	0.05
Curve /ʾ/	32	4.4	1.5	3.2	1.2	2.5	1.2	1.4	0.5
Broken /˜/	32	5.2	1.9	2.4	0.9	4.8	2.3	1.4	0.5

*Data from 8 NV informants: NF1, NF2, NF5, NM1, NM3, NM4, NM5 and NM6. The syllables were /ta/ and /tak/ for sonorant-ending and stopped tones respectively.

For each tone, mean intensity values on first line, SD on second line.

TABLE 6

MEAN INTENSITY IN dB AND STANDARD DEVIATIONS OF CV TONES ON SAME SYLLABLES AT FOUR TIMEPOINTS*

Tone	n	I1	SD	I2	SD	I3	SD	I4	SD
Level /-/	32	8.0	2.5	7.2	1.1	5.5	1.8	2.7	1.3
Falling /ˎ/	32	8.4	2.7	6.5	1.8	4.1	1.6	1.9	0.6
Rising /ʼ/	32	5.9	2.8	5.0	1.8	5.1	1.8	2.3	1.2
St. Ris. /ʼs/	32	5.3	2.6	4.9	2.1	4.6	1.8	1.6	0.5
Drop /./	32	6.5	3.1	4.7	1.5	3.1	1.9	1.7	0.9
St. Drop /.s/	32	6.6	3.2	5.3	1.5	3.8	1.3	1.4	0.8
Curve /ʾ/	32	7.3	2.7	5.8	2.2	3.2	1.6	1.4	0.7

*Data from 8 CV informants: CF1, CF2, CF4, CM1, CM4, CM5, CM7 and CM8. The syllables were /ta/ and /tak/ for sonorant-ending and stopped tones respectively.

For each tone, mean intensity values on first line, SD on second line.

TABLE 7

MEAN INTENSITY IN dB AND STANDARD DEVIATIONS OF
SV TONES ON SAME SYLLABLES AT FOUR TIMEPOINTS*

Tone	n	I1	SD	I2	SD	I3	SD	I4	SD
Level /-/	32	4.9	2.0	5.8	1.4	4.1	1.3	2.3	1.7
Falling /`/	32	3.5	1.3	3.8	0.9	2.4	0.8	1.2	0.6
Rising /´/	32	4.4	1.8	4.9	1.3	4.2	2.2	2.5	2.3
St. Ris. /´s/	32	4.2	2.3	5.8	2.0	5.0	1.5	2.4	1.7
Drop /./	32	3.2	1.3	2.9	1.1	2.0	0.8	1.2	0.4
St. Drop /.s/	32	3.9	1.2	3.7	1.2	2.7	0.3	1.2	0.4
Curve /’/	32	3.6	1.5	2.5	0.9	3.0	1.0	2.0	1.7

*Data from 8 SV informants: SF1, SF4, SF5, SM1, SM3, SM4, SM5 and SM6. The syllables were /ta/ and /tak/ for sonorant-ending and stopped tones respectively.

For each tone, mean intensity values on first line, SD on second line.

TABLE 8

MEAN DURATION IN CENTISECONDS AND STANDARD DEVIATIONS
OF NV, CV AND SV TONES ON SAME SYLLABLES*

Dialect & Tone	n	$\overline{D}$	SD	Dmax	Dmin**
NV Level /-/	36	25	6	36	12
Falling /\`/	32	25	6	38	12
Rising /ʹ/	32	25	5	38	14
St. Rising /ʹs/	32	15	3	22	10
Drop /./	32	20	4	30	14
St. Drop /.s/	32	15	3	22	8
Curve /ʾ/	32	26	6	40	16
Broken /~/	32	25	4	32	14
CV Level /-/	32	26	5	40	16
Falling /\`/	32	28	5	40	16
Rising /ʹ/	32	26	5	40	18
St. Rising /ʹs/	32	16	3	26	10
Drop /./	32	28	7	50	18
St. Drop /.s/	32	17	3	26	10
Curve /ʾ/	32	24	5	40	15
SV Level /-/	32	30	8	50	16
Falling /\`/	32	29	6	44	18
Rising /ʹ/	32	28	6	40	15
St. Rising /ʹs/	32	18	4	28	12
Drop /./	32	28	5	40	18
St. Drop /.s/	32	17	4	24	10
Curve /ʾ/	32	30	6	42	20

*Data from same syllables and same informants as for intensity data in Tables 5 to 7.

**Dmax and Dmin are the longest and shortest tokens found in each tone sample.

TABLE 9

LARYNGEALISATION IN NV, CV AND SV TONES*

Dialect and Tone	Degree (a)	Duration (b)	Timing (c)
NV Level /-/. Rising /ʹ/, St. Rising /ʹs/ and St. Drop /.s/	0		
Falling /ˎ/	0 (1)	(i)	(E)
Drop /./	2 (3)	4.5	E
Curve /ʾ/	0 (1)	(i)	(M,E)
Broken /˜/	2 (3)	3.7	M
CV Level /-/, Falling /ˎ/, St. Rising /ʹs/ and St. Drop /.s/	0		
Rising /ʹ/	0 (1)	(i)	(M)
Drop /./	0 (1,2)	(i)	(E)
Curve /ʾ/	2 (1)	4.6	E
SV Level /-/, Rising /ʹ/, St. Rising /ʹs/ and St. Drop /.s/	0		
Falling /ˎ/	0 (1)	(i)	(E)
Drop /./	0 (1,2)	(i)	(E)
Curve /ʾ/	0 (1)	(i)	(M)

*Based on auditory and acoustic studies on same syllables and same informants as for intensity data in Tables 5 to 7.

(a) 0:regular voicing; 1:breathy voice; 2:creaky voice; 3:glottal closure. () indicates alternative occurrences with some speakers and in some contexts only.

(b) Number indicates mean duration of laryngealised part in centiseconds; (i) indicates irregular durations.

(c) Laryngealisation may occur at the middle (M) or end (E) of the syllable.

4. DISCUSSION

4.1. *PHYSICAL PHONETIC PARAMETERS IN NORMALISED VALUES*

To understand and evaluate the common characteristics of Vietnamese tones and their variations across dialects, it is necessary to make the data comparable by using the same sets of parameters in describing them. However, the absolute and mean values given for those parameters as the results of direct measurements do not always make meaningful generalisations possible, because of the wide range of variations in non-linguistic parameters such as different Fo ranges between male and female speakers, differences in speech tempo or the power of their voices, etc. Therefore some normalisation procedures are proposed below to bring the data presented in 3. into directly conparable forms.

For comparison of Fo data, I devised a method of normalisation involving the notion of Fo Differential in function of the mean $\overline{F}$, or FD($\overline{F}$), expressed in percent in the following formula:

$$FD(\overline{F}) = Itg\left(\frac{Fi - \overline{F}}{\overline{F}} \times 100\right)$$

where Fi is any individual Fo value, $\overline{F}$ is the mean Fo of a sample, used as a reference level, and Itg stands for integer, i.e. the FD will be expressed in integer digits, any decimals being automatically dropped off.

The intensity and duration data were normalised according to two similar formulae:

$$I = Itg\left(\frac{Ii \times 10}{\overline{I}max} + 0.9\right)$$

$$D = Itg\ \frac{Di \times 10}{\overline{D}max} + 0.9)$$

where Ii and Di are any individual intensity and duration values to be normalised,

$\overline{I}$max and $\overline{D}$ max are the highest mean values of I and D in the samples in question, and 0.9 is a correcting factor.

These formulae will give normalised values for I and D in decimal scales where only integers are retained.

The application of the foregoing formulae gives results in Table 10 and is illustrated in Figures 1, 2 and 3. Table 10 gives the normalised values for the physical phonetic parameters of NV, CV and SV tones in their standard forms. Figure 1 gives diagrams of these tones in actual mean Fo plotted against mean duration, for comparison with Figure 2 where Fo in FD($\overline{F}$) percent of the same tones were plotted against normalised duration, and Figure 3 presents normalised intensity plotted against normalised duration.

TABLE 10

PHYSICAL PHONETIC PARAMETERS (Fo, L, I & D IN NORMALISED VALUES) OF NV, CV AND SV TONES*

Dialect & Tone	Fo in FD($\bar{F}$) Percent						Laryng.**			Intensity				Duration
	P1	P2	P3	P4	P5	P6	a	b	c	I1	I2	I3	I4	
NV /-/	9	9	9	8	7	4	0			10	9	8	3	10
/\`/	-7	-9	-10	-15	-17	-19	0			8	7	5	2	10
/'/	2	0	2	12	27	33	0			8	7	8	3	10
/'s/	14	20	29	38			0			8	7	7	3	6
/./	-2	-3	-6	-14	-16	-15	2	2	E	7	6	4	2	8
/.s/	-5	-9	-13	-17			0			7	7	5	2	6
/ʾ/	-8	-11	-18	-26	-19	-13	0			7	5	4	2	10
/~/	4	1	-12	9	26	26	2	2	M	8	4	7	2	10
CV /-/	12	10	12	14	14	13	0			10	9	7	4	10
/\`/	6	4	1	-1	-3	-4	0			10	8	5	3	10
/'/	-3	-6	-8	-1	7	10	0			7	6	6	3	10
/'s/	0	-3	1	9			0			7	6	6	2	6
/./	0	-3	-6	-9	-9	-9	0			8	6	4	2	10
/.s/	3	-1	-5	-8			0			8	7	5	2	6
/ʾ/	1	0	-4	-9	-6	-4	2	2	E	9	7	4	2	9
SV /-/	4	4	4	4	3	1	0			9	10	7	4	10
/\`/	-9	-12	-14	-17	-18	-18	0			6	7	5	2	10
/'/	6	4	9	22	36	39	0			8	9	8	5	10
/'s/	9	13	24	39			0			8	10	9	5	6
/./	-9	-14	-19	-18	-11	-9	0			6	5	4	2	10
/.s/	-7	-13	-15	-12			0			7	7	5	2	6
/ʾ/	-5	-12	-18	-4	19	22	0			7	5	6	4	10

*Calculated from data in Tables 2 to 9.

**In the Laryngealisation parameter,

(a) indicates degree,

(b) duration (same scale as for whole tone), and

(c) timing.

Intensity and duration values are in decimal scales.

4.2. COMMON CHARACTERISTICS AND DIFFERENCES

A number of observations can be made about the similarities and differences between NV, CV and SV tones from the data thus processed.

First, one can see that across dialects, all but one of the tones have basically similar Fo contours (level, falling, rising or concave), while varying in relative Fo level (high, mid or low) and in the presence or absence of laryngealisation, and the remaining one, the drop tone, displays difference in Fo contour in one dialect only, namely SV, as laryngealisation marks the contrast between the other two dialects. This suggests the primacy of Fo contour over other parameters as a major feature for differentiating tones in the Vietnamese system, and this fact is borne out in the analysis of tone perception by native speakers, as I reported elsewhere (1979, 1981). Together with the fact that Vietnamese has only one level tone out of the seven or eight phonetic tones in each dialect, Vietnamese can be typologically classified as a "contour tone language with register overlap" as defined by Pike (1948). This is further supported by analysis of subdialectal and individual variations, which showed that Fo contour is mainly characterised by the general direction of the Fo change, while great differences could occur in Fo slopes and Fo ranges. For example, the mean Fo differential between onset and endpoint of the NV rising tone is only 18% with Informant NM6 and as great as 53% with Informant NM1. It is also interesting to note that beside the expected difference in Fo ranges between male and female speakers, the use of Fo ranges differs markedly between CV and the other two dialects (see Figures 1 and 2).

Second, intensity shows no great differences between dialects, and correlation coefficients calculated for Fo and intensity values at the same timepoints indicate a fair degree of correlation. Duration is not significantly different between sonorant-ending tones, except for the creaky-ending NV drop tone and CV curve tone, which are significantly shorter by 20% and 10% respectively, and for the stopped tones where duration is 40% shorter in all three dialects. I take this to mean that duration is not an independent factor in tone production but is conditioned by the presence of laryngealisation or the voiceless final stop at syllable endings, which cause the shortening. Both intensity and duration may thus be characterised as independent parameters at the physical phonetic level only, and would become redundant at higher levels of analysis.

FIGURE 1

Mean Fo of NV, CV and SV Tones Plotted Against Mean Duration
(Data from 9 NV, 12 CV and 9 SV Informants)

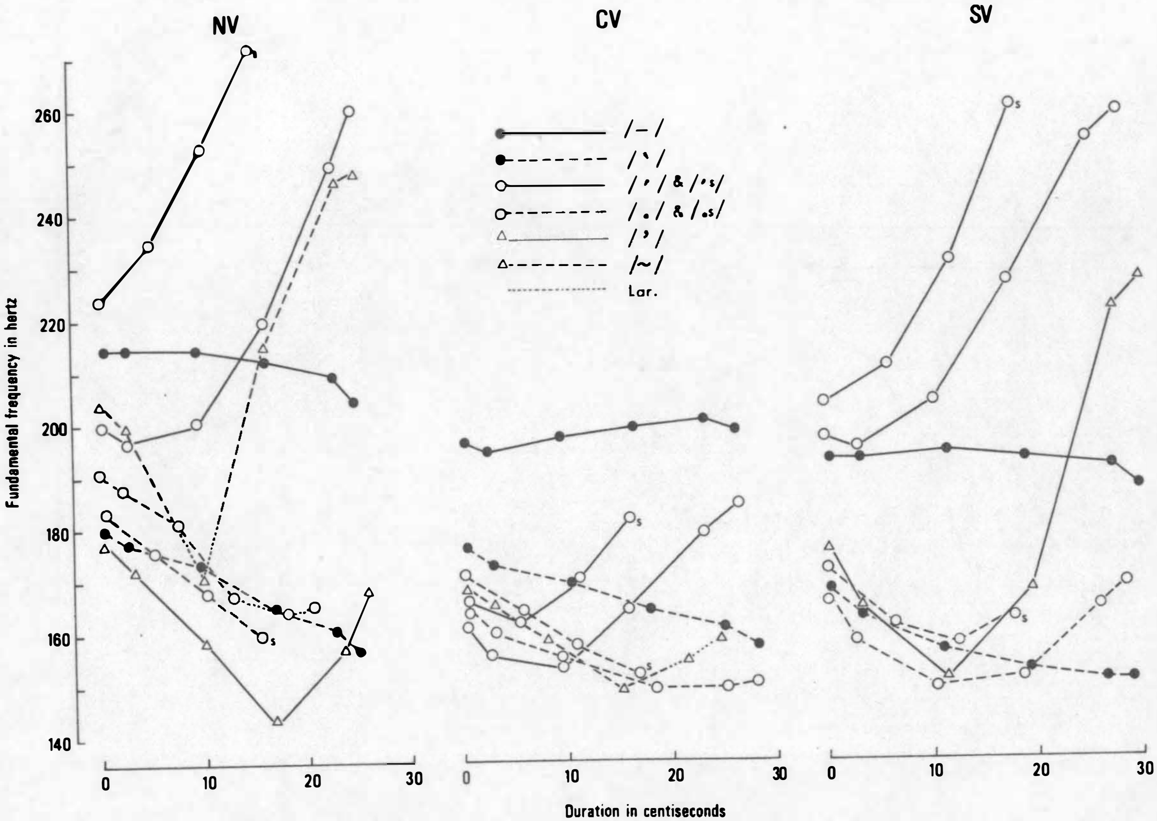


FIGURE 2

Normalised Mean Fo in FD($\overline{F}$) Percent of NV, CV and SV Tones Plotted Against Normalised Duration

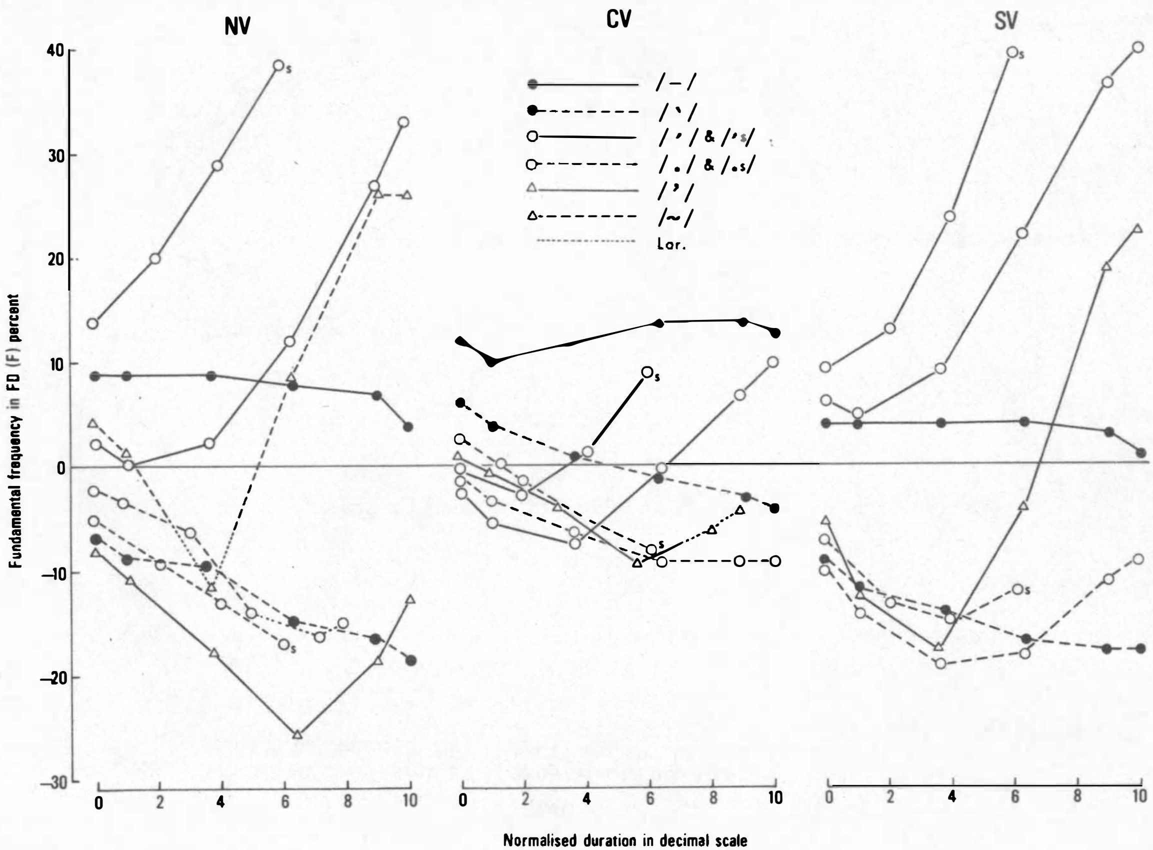
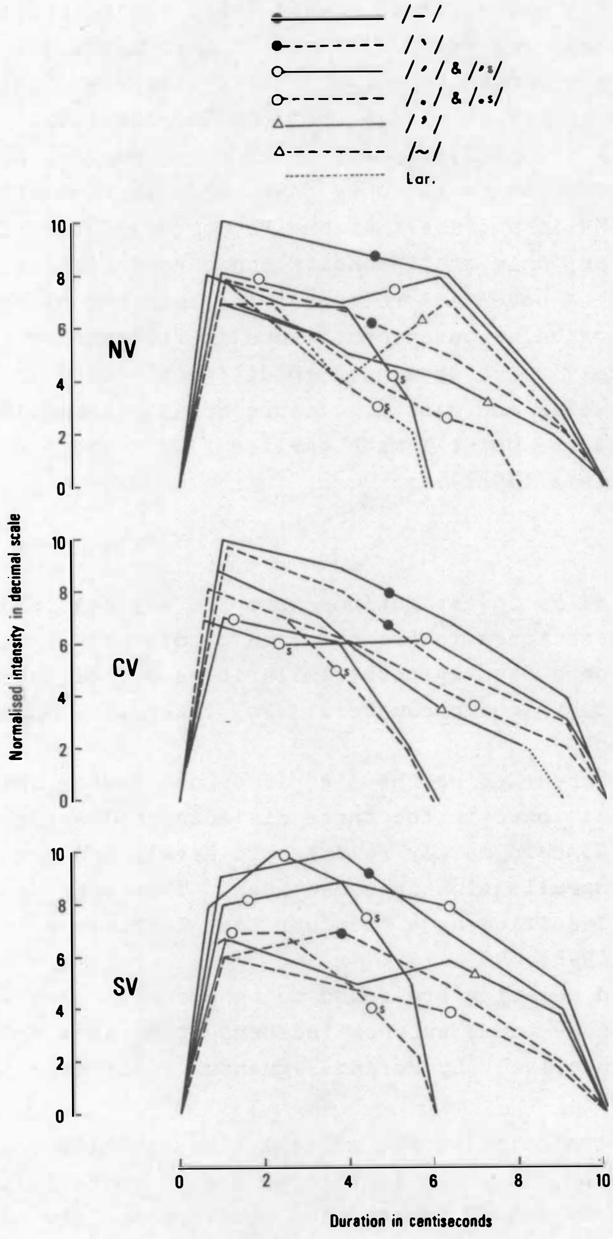


FIGURE 3

Normalised Mean Intensity of NV, CV and SV Tones Plotted Against Normalised Duration (Data from Table 10)



Third, laryngealisation or creaky voice, a characteristic of the NV drop tone and broken tone, and of the CV curve tone, is not a regular feature of any SV tone. While breathy voice and glottal closure may occur with some tones in all three dialects, they are not a regular characteristic of any, and can best be regarded as alternatives in free variation for some standard forms. It is interesting to note that what is auditorily perceived as creaky voice may be realised differently in acoustic terms. For example, three NV informants display marked differences in the broken tone: NF3: heavy laryngealisation, sharp drop in Fo and in intensity at middle; NM3: no laryngealisation, sharp drop in Fo and slight drop in intensity at middle; and NM4: no laryngealisation, no sharp drop in Fo but only sharp drop in intensity at middle. For Informant NM4 it appears that the Fo curves of the rising and broken tones are similar; only the intensity contours differ sharply.

This fact is of potential relevance for both the historical evolution and the physiological production of tone in Vietnamese: it might explain how creaky voice developed in different tones in NV and CV, and why creaky voice and glottal closure occur alternatively in some tones. This was the point I made earlier (1980) and discussed in more detail in my thesis (1981).

5. CONCLUSION

The results of my investigations into the physical phonetic properties of Vietnamese tones in the three major dialects have shown that NV, CV and SV tones display both similarities and differences and can be characterised by four parameters: Fo, intensity, duration and laryngealisation.

Fo contours appear to be the most important factor that unites the same phonological tones in the three dialects, below the surface differences which concern mainly relative Fo level, and presence or absence of laryngealisation in some tones. This suggests that Vietnamese can be classified as a "contour tone language with register overlap" (Pike 1948).

Intensity and duration are found to be phonetic parameters characterising some tones but not independently; they are probably conditioned respectively by Fo and segmental environment or laryngealisation.

Apart from breathy voice and glottal closure which occur irregularly as free variations, laryngealisation or creaky voice is a distinctive feature of some NV and CV tones. Its auditory quality may be the effect of different acoustic realisations and this fact might have implications for historical tone evolution in Vietnamese.

NOTES

1. This is an abridged version of a chapter from my thesis on Vietnamese tones. I am gratefully indebted to my successive supervisors, Dr David Bradley (now at the University of Melbourne), Dr Tim Shopen and Mr Phil Rose, of the Department of Linguistics, The Faculties, Australian National University, for their valuable advice in my research.

BIBLIOGRAPHY

ANDREEV, N.D. and M.V. GORDINA

1957 Sistema tonov vjetnamskogo jazyka. *Vestnik Leningradskogo gosudarstvennogo Universiteta: Serija Istorii, Jazyka i Literatury* 8:132-148. Leningrad.

CHAO, Y.R.

1930 ə sɩstəm əv "toun-lɛtəz". *Le Maître Phonétique* 30:24-27.

EARLE, M.A.

1975 *An acoustic phonetic study of Northern Vietnamese tones.* Santa Barbara, California: Speech Communications Research Laboratory, Inc.

HAN, Mieko S.

1969 *Studies in the phonology of Asian languages VIII: Vietnamese Tones.* Los Angeles: University of Southern California.

HAN, Mieko S. and KIM Kong-On

1972 *Studies in the phonology of Asian languages X: Intertonal influence in two-syllable utterances of Vietnamese.* Los Angeles: University of Southern California.

LÊ VĂN LÝ

1948 *Le parler vietnamien.* Paris: Huong Anh.

NGUYỄN HÀM DƯƠNG

1962 Sistema tonov i spectry glasnyx vjetnamskogo jazyka. Kandidatskaja dissertacija. Moskva.

PIKE, K.L.

1948 *Tone languages.* Ann Arbor: University of Michigan Press.

TRẦN HƯƠNG MAI, Aurélie

1969 Stress, tones and intonation in South Vietnamese. Ph.D. thesis, Australian National University, Canberra.

VŨ THANH PHƯƠNG

1979 Tone perception in Vietnamese. Paper presented at the 11th Annual Conference of the Australian Linguistic Society. Newcastle, New South Wales.

1980 The dynamics of tone change: phonetic and historical evidence from Vietnamese. Paper presented at the 12th Annual Conference of the Australian Linguistic Society. Melbourne, Victoria.

1981 The acoustic and perceptual nature of tone in Vietnamese. Ph.D. thesis, Australian National University, Canberra.

Phuong, V.T. "Phonetic Properties of Vietnamese Tones Across Dialects". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation.*
A-62:55-76. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.55*

SOME ACOUSTIC PROPERTIES OF TONES IN BURMESE

U Thein Tun

1. PHONOLOGICAL TONES

On the basis of Trubetzkoy's definitions of phoneme, phonological unit, phonological contrast, and his rules 2 and 3 for the differentiation of phonemes and variants (Trubetzkoy 1968:7, 1969:48-49), the following phonological tones can be assigned to the phoneme inventory of Burmese phonology.

Four tones:

Tone I	/tɕá/	*'lotus', 'to take a long time'*
Tone II	/tɕâ/	*'to hear'*
Tone III	/tɕà/	*'to fall down'*
Tone IV	/tɕaˀ/	*'to be stringent', 'to be tight'*

The Burmese phonological tones, however, are classified differently by various writers. Taylor (1920) and Firth (1933) classify the Burmese tones into three. Firth (1936) and McDavid (1945) suggest that there are five tones in Burmese. Cornyn (1944), Burling (1967), Stewart (1955) and Becker (1964) agree in classifying the Burmese tones into four.

If the CVC phonological syllable pattern is recognised, as in present-day written Burmese and old epigraphic Burmese, it is true that /tɕaˀ/ (written tɕap), (Tone IV in the above example) could be treated as belonging to the CVC pattern, and thus there would be only three tones in Burmese. In present-day spoken Burmese, however, the final stops of the written CVC pattern are replaced by a final stop, which is a glottal stop in isolation. If the CVC pattern is regarded as one phonological syllable pattern, then this characteristic will be the sole representation of the final C of the CVC pattern. Therefore it is more acceptable to analyse the Burmese phonological syllable pattern as V and CV and to treat the final stop as a tonal feature.

Thein-Tun, U. "Some Acoustic Properties of Tones in Burmese". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation*.
A-62:77-116. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.77*

Firth (1936) and McDavid (1945) classify the Burmese tones into five because they treat the neutral vowel [ə] as a separate tone. The neutral vowel [ə] appears as a non-final **neutral** syllable in words consisting of two or more syllables, but never in monosyllabic words. Burmese is basically a monosyllabic language and any non-monosyllabic word in Burmese with a neutral vowel [ə] is either a borrowing or a lexical compound or the result of a derivational process. Therefore, in this analysis, the Burmese phonological tones are classified into four. [Note that Tones I, II and II in this analysis are generally described as even or level tone, heavy or breathy tone and creaky tone respectively in the literature. The cardinal order and symbols of the tones in this analysis are somewhat different from Okell (1969) but same as Cornyn (1944) and Cornyn and Roop (1968).]

2. EXPERIMENT 1

There are altogether fifty vocalic nuclei in Burmese apart from the neutral vowel [ə].

Basic symbol	non-nasalised				nasalised			
	Tone I	Tone II	Tone III	Tone IV	Tone I	Tone II	Tone III	Tone IV
i	1. í	2. î	3. ì	4. iʔ	5. ĩ́	6. ĩ̂	7. ĩ̀	X
e	8. é	9. ê	10. è	11. eiʔ	12. ẽ́ĩ	13. ẽ̂ĩ	14. ẽ̀ĩ	X
ɛ	15. ɛ́	16. ɛ̂	17. ɛ̀	18. ɛʔ				X
ai				19. aiʔ	20. ã́ĩ	21. ã̂ĩ	22. ã̀ĩ	X
a	23. á	24. â	25. à	26. aʔ	27. ã́	28. ã̂	29. ã̀	X
ɔ	30. ɔ́	31. ɔ̂	32. ɔ̀	33. auʔ	34. ã́ũ	35. ã̂ũ	36. ã̀ũ	X
o	37. ó	38. ô	39. ò	40. ouʔ	41. ṍũ	42. õ̂ũ	43. õ̀ũ	X
u	44. ú	45. û	46. ù	47. uʔ	48. ṹ	49. ũ̂	50. ũ̀	X

([ei], [au] and [ou] can be regarded as the different realisations of /e/, /ɔ/, and /o/ in the environment of nasalisation or a final stop and [ãĩ] could be the allophone of /ai/ in the nasalised environment.)

All these fifty vocalic nuclei of two male native speakers and two female native speakers were recorded in [hVdá] (/hVdá/) frame for the monophthongal nuclei and in [hṽndá] (/hṼdá/) and [hvvʔdá] (/hVʔdá/) frames for the diphthongal nuclei. The initial consonant [h] was chosen in the frames because it is the weakest consonant to bring forth any co-articulation effect on the following vowel; the medial

[d] was chosen to make the segmentation easy; and the final [ǝ́] was chosen to make the utterances natural for the Burmese speakers as this is a frequent vowel.

Some of the utterances in these frames are nonsense words whether the last syllable [dǝ́] is eliminated or not. These nonsense words however have to be tolerated because the purpose of the experiment is to analyse the phonetic qualities of the vocalic nuclei of the utterances and they are not foreign for the Burmese speakers to pronounce. The established writing system enables the subjects to pronounce them easily and naturally whether they are meaningful or not.

Every speaker uttered every sound twice and both utterances were recorded. All the recorded sounds were processed with a Frokjaer-Jensen intensity and fundamental frequency meters, using the linear intensity output and the duplex oscillogram and fundamental frequency output. The outputs were displayed by a mingograph. The necessary precautions were observed both in recording and processing. When the fundamental frequency of every utterance was measured, a slight difference between the first and the second utterances of the same sound sometimes occurred. In such a case the average value of the two was taken.

Of the three outputs observed on the mingograph (i.e. intensity display trace, duplex oscillogram trace and the fundamental frequency trace), the intensity display trace of the tones does not show any consistency or any common factor by which the contrastive features of the tones can be distinguished. It is tempting to say that in the null context that these fifty utterances are in, the intensities of the tones do not behave regularly as they do in sentences. The duplex oscillogram trace also enables us only to distinguish the vowel segments from the adjacent consonants along the fundamental frequency trace. Therefore, of the three parameters, only the fundamental frequency trace of the tones is dealt with in this present analysis.

The fundamental frequency as shown in the mingograms of all the four subjects, rises gradually from Tone I to Tone IV. The fundamental frequency for Tone I starts at a relatively level range and tends to go down slightly; the fundamental frequency for Tone II starts at a relatively level range, goes up, and then falls down relatively low; the fundamental frequency for Tone III starts at a relatively high range, usually higher than or as high as the peak of Tone II, and falls down relatively low; the fundamental frequency for Tone IV starts at a high range, frequently higher or as high as the peak of Tone II and falls low, but not as low as Tone III because it stops very suddenly before it can drop lower. The general contrastive features of the four phonological tones offered by the analysis of their fundamental frequency can be described as:

I level, low
II high, rising, falling
III high, falling
IV high, falling, abrupt end.

For the detailed contrastive characteristics of the tones, in terms of their fundamental frequencies, there is no special point to be observed for Tones I and II because Tone I simply starts at a level range and falls slightly in the end and Tone II starts at a level range, goes up and falls down to the level range. However there are some detailed contrastive features to be observed for Tones III and IV in comparison with Tone II.

The fundamental frequency of all the utterances in Tone III by the two female speakers starts with a high range, higher than the peak of Tone II with one exception. For the one exception also, the starting range is the same as the peak of Tone II. The fundamental frequency of the utterances of the first male subject in Tone III starts with a relatively high range, usually as high as the peak of Tone II. Of the fourteen utterances of the first male subject in Tone III, the fundamental frequency for nine he starts with the high range equal to the peak of Tone II, the fundamental frequency for three starts with the high range higher than the peak of Tone II; and the fundamental frequency for the remaining two starts at a range slightly lower than the peak of Tone II. The fundamental frequency for all the utterances in Tone III by the second male subject starts at the high range, higher than the peak of Tone II.

For the two female subjects, the commencing high range of Tone IV is generally either higher than or as high as the peak of Tone III. However, in one case for one female subject and in three cases for the other, the commencing range of Tone IV is either equal to or slightly lower than the peak of Tone III. Of the eight utterances in Tone IV of one male subject, four are equal to and four are slightly lower than the peak of Tone III; of the eight utterances in Tone IV of the other male subject, three are equal to and five are slightly lower than the peak of Tone III.

The following table summarises the average fundamental frequencies of the four tones for the four subjects, out of their 200 utterances i.e. 50 each (14 in Tone I, II and III and eight in IV) among the four subjects i.e. out of 56 utterances in each of the first three tones and 32 utterances in Tone IV.

TABLE 1

Tone	Subject	Initial (High) HZ	Middle (Peak) HZ	Final (Low) HZ
I	1st female	226	-	174
	2nd female	230	-	176
	1st male	122	-	106
	2nd male	135	-	105
II	1st female	237	255	191
	2nd female	241	254	213
	1st male	144	165	142
	2nd male	146	152	137
III	1st female	302	-	190
	2nd female	300	-	160
	1st male	166	-	117
	2nd male	184	-	92
IV	1st female	297	-	204
	2nd female	294	-	175
	1st male	159	-	122
	2nd male	172	-	110

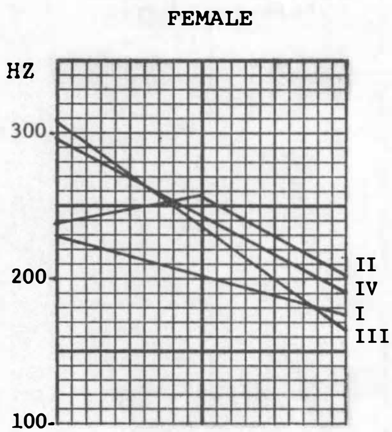
The following table summarises the average fundamental frequencies of the four tones for two females and two males.

TABLE 2

Tone	Initial HZ		Middle HZ		Final HZ	
	F	M	F	M	F	M
I	228	123	-	-	175	105
II	239	145	254	158	202	144
III	301	175	-	-	165	104
IV	295	165	-	-	190	116

This table can be visualised more clearly on the graph.

TABLE 3



In these graphic descriptions of the tones, there is very little distinction between Tone III and Tone IV. This kind of distinction is to be made clear in the next experiment and analysis where the factor of duration of the tones is taken into account.

3. EXPERIMENT 2

The aim of this experiment is to find out the further acoustic qualities of the tones in terms of length. The author's original intention was to put the aforesaid fifty vocalic nuclei in one frame which can occur both in null context i.e. by itself as a meaningful syllable; and in phrase or sentence context. As the most suitable syllable pattern for this purpose, the writer originally chose [tv(v)] (/tV/) frame. The initial consonant [t] was chosen to make the segmentation easy. All the fifty vocalic nuclei can fit meaningfully into this frame. The author made up fifty sentences with these fifty vocalic nuclei in [tv(v)] frame. As a tentative effort, four male native speakers were asked to utter these fifty nuclei in [tv(v)] frame, both in null context and sentence context. Their utterances were recorded and processed with the Kay sound spectrograph. It was found that the correct analysis was almost impossible if these fifty vocalic nuclei are put in CV frame. The following major defects presented problems.

1. /V/ of /tV/ can be followed by any consonant which is the initial consonant of the next CV syllable and the following consonant can affect the F2 off-glide frequency value of the vocalic nucleus.

2. When /V/ of /tV/ is followed by a voiced stop which is the initial consonant of the next CV syllable in a close juncture, the

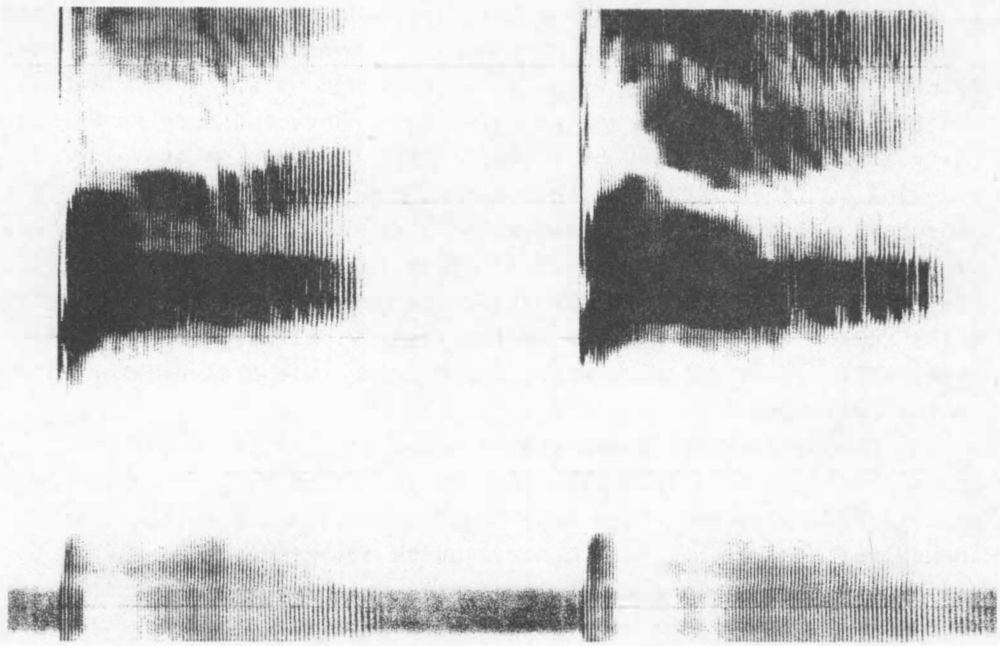
following voiced stop becomes continuant in the form of a fricative. For instance, /tí/ means *'earth-worm'* and /káu/ is a classifier for animals and insects. When /tí/ and /káu/ are combined in close juncture, /k/ becomes voiced according to the voicing rule and it becomes [tígáu]. However the process of the change from /k/ does not stop at [g]. That [g] which is the phonetic realisation of /k/ often becomes voiced and is shown on the sonogram as voiced velar fricative [ɣ]. In such a case it is very difficult to segment the preceding [V] from [ɣ] which shows the characteristics of a resonant and continuant like a vowel. Such a problem occurs especially when [a] is followed by a voiced bilabial and [i] is followed by a voiced velar consonant, and more so when the [tv(v)] frame is followed by a vowel. For the problem of voiced stops changing to fricatives and continuants, one might like to choose the consonant which does not become voiced, for the initial C of the next CV syllable which follows [tv(v)]. This is impossible because to get the right length of the /V/ of /tV/, the next syllable should immediately follow it, and when one CV immediately follows, two syllables are combined in close juncture where the voicing rule always takes place on the following C.

3. Phonologically, a meaningful vowel or a syllable which is grammatically a free form can occur by itself in null context. People do say and understand these individual monosyllables which occur by themselves. However, these monosyllables are generally followed by a particle. If it is a noun it is usually followed by a classifier or a particle and if it is a verb it is usually followed by a particle, either stative, imperative or polite and so on. As people are so accustomed to adding particles to monosyllables, they unconsciously find it awkward to articulate one individual syllable by itself. Hence they either repeat it or put a strong stress on it when they have to articulate one individual syllable. That kind of stress is especially strong when they pronounce it consciously before the microphone in the studio. That is why the vocalic nuclei in null context of these subjects are extremely and unnaturally long, and hence the writer was not happy to use the sonagrams they produced, as reliable data. The following sonagram is an example of it.

FIGURE 1

Sonagram of /ti^{II}/ Repeated Twice.

/t/ followed by /i/ in Tone II.
The average length for /i/ in Tone II is only 21 Cs, whereas it is 30.8 Cs in this sonagram.



For these reasons, the writer had to eventually use the frames from Experiment 1. These frames /hVdá/, /hṼdá/ and /hVˀdá/ ([hvdá], [hṽvndá], [hvvˀdá]) do not create either null context or sentence context. The last CV syllable [dá] in these frames has its grammatical function sometimes as a demonstrative particle (that or this) and sometimes as a nominaliser. For instance, in [hódá] ([hvdá] with [ó]), [hó] means there, [dá] means that, and hence [hó] and [dá] are two separate words forming a phrase. In [háundá] (/hṼdá/ with /áu/), [háun] (/háu/) means *'to bark'*, [dá] functions as a nominaliser, and hence [háundá] means *'barking'*, forming a separate word as a noun. In this way these three frames form either word or phrase context which lies between the two extremes - null context with one syllable where the nucleus length is extremely long, and sentence context where the nucleus length is extremely abrupt. Moreover, in these three frames, the nuclei are consistently followed by the same consonant i.e. [d] and hence there is no disturbance of different co-articulation effect on the F2 offglide value. Therefore in this experiment as well, the writer used, with satisfaction and assurance, the same three frames from Experiment 1.

The same procedure of utterance, recording and averaging between the first and second utterances, was also used in this experiment. However, in this experiment, to get the accurate average measurements of formant frequencies and lengths, ten more male native speakers were added to the original number of two (first and second male subjects, subjects 1 and 2) and thus the number of subjects was raised to twelve, and no female voice was used. This time all these recorded utterances (1,200 utterances altogether; 50 nuclei each uttered twice by every subject) of the twelve male subjects were processed with the Kay sound spectrograph. The analysis was made on 1,200 sonagrams produced from these 1,200 utterances.

4. TABLES SHOWING THE AVERAGE LENGTHS OF THE TONES

The following tables summarise the average component durations of the four phonological tones among the twelve male subjects. For monophthongal vowels, transitions toward and away from the steady state are given as onglide and offglide; for diphthongal vowels, first and second steady states, and transition between the two are shown, as well as onglide and offglide transitions if any.

TABLE 4

Basic Symbol	Tonal Symbol	Onglide Cs	Steady State (SS) Cs	Offglide Cs	Total Cs
/i/	I /í/	4.8	9.0	5.0	18.8
	II /î/	4.4	11.8	5.5	21.7
	III /ì/	3.9	8.9	3.2	16.0
	IV /iˀ/	2.0	4.3	2.8	9.1
/e/	I /é/	4.8	9.6	5.0	19.4
	II /ê/	5.7	12.0	5.2	22.9
	III /è/	3.6	9.1	4.1	16.8
			SS1 Trans SS2		
	IV [eiˀ]/eˀ/	0	3.8 6.0 3.8	0	13.6

TABLE 4 (cont.)

Basic Symbol	Tonal Symbol	Onglide Cs	Steady State (SS) Cs	Offglide Cs	Total Cs
/ɛ/	I /ɛ́/	4.5	11.5	4.5	20.5
	II /ɛ̂/	6.6	11.4	4.5	22.5
	III /ɛ̀/	4.1	8.5	4.8	17.4
	IV /ɛˀ/	1.7	7.6	1.6	10.9
			SS1 Trans SS2		
/ai/	IV /aiˀ/	0	4.6 6.7 3.7	0	15.0
/a/	I /á/	5.1	11.0	5.9	22.0
	II /â/	6.8	12.3	5.4	24.5
	III /à/	5.0	10.0	5.0	20.0
	IV /aˀ/	1.9	6.8	2.0	10.7
/ɔ/	I /ɔ́/	5.6	10.6	5.6	21.8
	II /ɔ̂/	5.6	12.4	5.2	23.4
	III /ɔ̀/	3.8	7.9	5.5	17.2
	IV [auˀ]/ɔˀ/	0.5	6.0 6.0 3.0	0	15.5
/o/	I /ó/	5.7	10.6	5.0	21.3
	II /ô/	6.6	11.6	5.9	24.1
	III /ò/	4.7	8.7	4.0	17.4
	IV [ouˀ]/oˀ/	0	4.2 6.7 3.3	.4	14.6
/u/	I /ú/	5.0	9.8	4.7	19.5
	II /û/	5.1	12.4	4.7	22.2
	III /ù/	3.9	7.4	3.6	14.9
	IV /uˀ/	2.2	4.8	2.7	9.7
/ɪ̃/	I [ɪ̃́]	3.4	8.1	2.8	14.3
	II [ɪ̃̂]	3.8	8.3	4.2	16.3
	III [ɪ̃̀]	2.1	6.3	2.4	10.8
/ã/	I [ã́]	3.6	6.4	3.4	13.4
	II [ã̂]	3.9	8.0	4.1	16.0
	III [ã̀]	2.73	6.53	2.73	12.0

TABLE 4 (cont.)

Basic Symbol	Tonal Symbol	Onglide Cs	Steady State (SS) Cs			Offglide Cs	Total
/ũ/	I [ṹ]	3.2	7.1			3.9	14.2
	II [ũ̂]	4.1	7.8			4.7	16.6
	III [ũ̀]	2.8	6.1			2.4	11.3
/ẽ/			SS1	Trans	SS2		
([ẽĩ])	I [ẽ́ĩ]	0.3	5.3	7.8	4.5	0	17.9
	II [ẽ̂ĩ]	0.7	4.5	9.1	5.5	0.5	20.3
	III [ẽ̀ĩ]	0	4.3	6.5	4.1	0	14.9
/ãĩ/	I [ã́ĩ]	0	5.4	10.1	5.0	0	20.5
	II [ã̂ĩ]	0	6.0	10.6	5.7	0	22.3
	III [ã̀ĩ]	0	4.7	7.2	4.4	0	16.3
/ɔ̃/							
([ãũ])	I [ã́ũ]	0.9	6.3	6.0	5.7	0	18.9
	II [ã̂ũ]	0	7.1	7.2	5.6	0.6	20.5
	III [ã̀ũ]	0.3	5.2	6.5	3.9	0	16.0
/õ/							
([õũ])	I [ṍũ]	0	4.8	5.6	5.2	0.6	16.2
	II [õ̂ũ]	0	6.13	5.13	8.43	1.12	20.8
	III [õ̀ũ]	0.5	4.4	5.5	4.1	0.3	14.8

4.1. *TOTAL LENGTHS OF THE TONES*

The analysis of the overall lengths of the tones shows that the total length is a very consistent contrastive factor of the tones. For all the vocalic nuclei Tone I is of moderate length, Tone II is longer than Tone I, Tone III is shorter than Tone I, and Tone IV is so short and abrupt that it is even shorter than Tone III. Hence the conclusion that can be made for the general contrastive feature of the tones on the basis of their total length is:

Tone I moderate
II long
III short
IV abrupt

Of 600 (50 x 12) average (average of first and second utterances) measurements, there are only 34 exceptions that do not agree with the above-mentioned generalisation. The common feature of these exceptions is that either Tone II is slightly shorter or as short as Tone I, or Tone IV is slightly longer or as long as Tone III. There are some points that can explain the cause of these 34 deviations from the standard features.

1. All the subjects were instructed to utter every sound as they usually say it, but taking care to be accurate. When the subjects were told to articulate the sounds accurately, they were so cautious about glottis articulation and tongue and lip articulation (which shape the lower formant frequencies), that they tended to give less attention to the length factor; and this is more likely to happen especially at word and phrase level where the utterance frames are.

2. Of these 34 deviations more than half (18) are in [hṽvndá] /hṼ̃dá/ frame, nasalised environment where the later part of nasalisation changes to the homorganic nasal of the following [d] i.e. [hṽvndá]. In such a case, in the sonagram, the nasal bar is sometimes present all along underneath both F1 and F2. Sometimes F1 and F2 overlap on the nasal bar underneath them and in such a case, it must be admitted that, the segmentation of [vv] from [n] is not clear cut. Therefore out of 34 deviations, 18 can be due to the inaccuracy of the segmentation.

3. It is a well known fact that the linguistic factors such as frequency, length etc. are significant only if they are compared with each other in sentence context. In an isolated context such as a word or phrase consisting of only two syllables, the linguistic factor may not be as significant as in sentence context.

Even if these three points fail to explain the cause of these deviations, the number of deviations (34) is tolerable compared to the overall number of 600; that is, the deviations from the standard represent only 5.5% of all the utterances. Therefore, the above-mentioned generalisation for the contrastive features of the four phonological tones, on the basis of their length, can be accepted as satisfactory.

As the table shows, these 34 exceptions do not carry any significant weight on the average. The average total length of all the four tones entirely agrees with the above-mentioned length contrastic features of the tones. The average length difference of the three tones from Tone I are summarised in the following table for basic vowel phonemes and allophones.

TABLE 5

Basic Symbol	Tone II Cs	Tone III Cs	Tone IV Cs
/i/	+2.9	-2.8	-9.2
/e/	+3.5	-2.6	-5.8
/ɛ/	+2.0	-3.1	-9.6
/a/	+2.5	-2.0	-11.3
/ɔ/	+1.6	-4.6	-6.3
/o/	+2.8	-3.9	-6.7
/u/	+2.7	-4.6	-9.8
[ĩ]	+2.0	-3.5	
[ã]	+2.6	-1.4	
[ũ]	+2.4	-2.9	
[ẽi]	+2.4	-3.0	
[ãi]	+1.8	-4.2	
[ãu]	+1.6	-2.9	
[õu]	+4.6	-1.4	
The average length difference of the three tones from Tone I, for all 14 basic vocalic nuclei	+2.53	-3.06	-8.40

The average length of Tone I for all the 14 vocalic nuclei is 18.5 Cs.

In the previous analyses the average lengths of both monophthongs and diphthongs in stop-final environment (Tone IV) are compared with those of their oral counterparts in the other three tones. It should be of some interest to compare the average lengths of these stop-final nuclei with those of their nasalised counterparts in the other three tones. The following tables show that the alternative comparison of tones in terms of their lengths also follows the pattern of moderate (Tone I), long (Tone II), short (Tone III), and abrupt (Tone IV).

TABLE 6

Basic Symbol	Tonal Symbol	Total Length (Cs)
/i/	I [í]	14.3
	II [î]	16.3
	III [ì]	10.8
	IV [iʔ]	9.1
/a/	I [á]	13.4
	II [à]	16.0
	III [à]	12.0
	IV [aʔ]	10.7
/u/	I [ú]	14.2
	II [û]	16.6
	III [ù]	11.3
	IV [uʔ]	9.7
average for monophthongs	Tone I 13.97	
	II 16.30	
	III 11.37	
	IV 9.83	

TABLE 7

Basic Symbol	Tonal Symbol	Total Lengths (Cs)
[ei]	I [éi]	17.3
	II [êi]	20.3
	III [èi]	14.9
	IV [eiʔ]	13.6
[ai]	I [ái]	20.5
	II [âi]	22.3
	III [ài]	16.3
	IV [aiʔ]	15.0
[au]	I [áu]	18.9
	II [âu]	20.5
	III [àu]	16.0
	IV [auʔ]	15.5

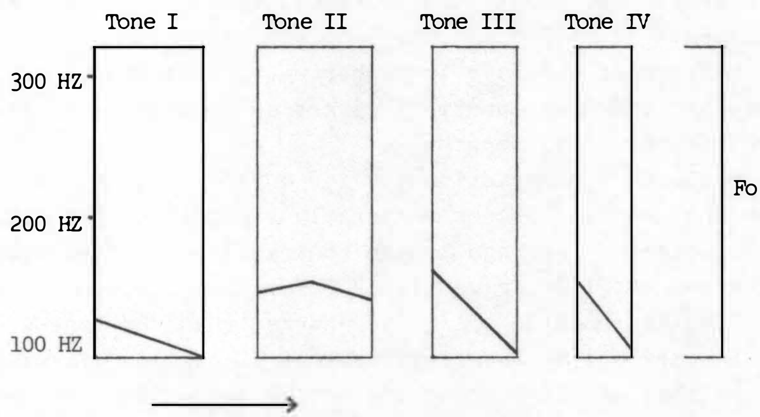
TABLE 7 (cont.)

Basic Symbol	Tonal Symbol	Total Lengths (Cs)
[ou]	I [óu]	16.6
	II [ôu]	20.8
	III [òu]	14.8
	IV [ou$^?$]	14.6
average for diphthongs	Tone I 18.48	
	II 20.98	
	III 15.56	
	IV 14.68	

If we add the contrastive features of the tonal fundamental frequency found in Experiment 1 to the length contrastive features of the tones found in this experiment, we will have the following complete contrastive features of the tones, fundamental frequency and duration.

Tone I level, low; moderate
II high rising, falling; long
III high falling; short
IV high falling; abrupt

The complete contrastive features of the tones can be visualised more clearly in terms of their average length and average fundamental frequency as shown in the following chart for the male speakers:



Time Scale: 1 cm = 1 Cs

5. COMPONENT DURATIONS OF THE MONOPHTHONGAL NUCLEI

As the table of average component durations of the twelve subjects shows, the average durations of onglides, steady states and offglides of the four tones in all the monophthongal nuclei for the twelve subjects, tend to behave in conformity with the contrastive feature of the total length of the four tones i.e. the pattern of 'I moderate, II long, III short and IV abrupt' as mentioned above. Out of the ten monophthongal vocalic nuclei (seven vowel phonemes which occur in all the tones and three nasalised variants which occur in the first three tones) the onglide of /ì/ and /ɔ́/, the steady state of /ɛ̂/ and /ầ/, and the offglide of /ɛ̂/, /â/ and /ɔ̂/, show exception in that they do not behave in conformity with the above-mentioned pattern.

For the onglide, Tone II of /i/ is 0.4 Cs shorter than Tone I of /i/ and Tone II of /ɔ/ is the same as Tone I of /ɔ/.

For the steady state, Tone II of /ɛ/ is 0.1 Cs shorter than Tone I of /ɛ/ and Tone III of /ã/ is 0.4 Cs longer than Tone I of /ã/.

For the offglide, Tone II of /ɛ/ is the same as Tone I of /ɛ/, Tone II of /a/ is 0.5 Cs shorter than Tone I of /a/, and Tone II of /ɔ/ is 0.4 Cs shorter than Tone I of /ɔ/.

Although there are only two exceptions for onglide, two for steady state and three for offglide, the writer would not be inclined to conclude that the average onglide, steady state and offglide lengths of the tones on all the monophthongal nuclei among the twelve subjects generally follow that moderate-long-short-abrupt pattern of the tonal length. There are two points that make the writer reluctant to draw that conclusion:

1. The pattern of moderate-long-short-abrupt length on onglide, steady state and offglide durations is not as consistent at the individual level as the total length.

2. The pattern of contrastiveness for onglide, steady state and offglide on the average, is not remarkably significant. In other words, that pattern of average length contrastiveness for onglide, steady state and offglide is very weak. For instance, on the average, though the Tone III onglide of /a/ is shorter than the Tone I onglide of /a/, it is only 0.1 Cs shorter; the Tone II onglide of /u/ is only 0.1 Cs longer than the Tone I onglide of /u/ and so on. There are quite a number of similar instances for the average tonal lengths of steady state and offglide as well.

Therefore the writer has to be content with a broad generalisation that the average durations of onglide, steady state and offglide of the four tones on all the monophthongal nuclei for the twelve subjects tend to behave in conformity with the total length contrastive pattern of 'moderate-long-short-abrupt'.

6. COMPONENT DURATIONS OF THE DIPHTHONGAL NUCLEI

The four two-target nuclei [ẽĩ], [ãĩ], [ãũ] and [õũ] occur only in first three tones. On the individual level they either have very short onglides and offglides or no onglide and offglide at all. Their overall length is usually the total of SS1, transition and SS2. That is why they show very limited duration of onglide and offglide or neither of the two. It is clear that the following consonant [d] has limited co-articulation effect on the diphthongal nuclei as Burmese has no VC or CVC phonological syllable pattern. As in onglides, steady states, and offglides of one-target nuclei, SS1, transition and SS2 lengths of diphthongal nuclei merely tend to follow the contrastive pattern of the total length with some exceptions and little significance.

7. CONCLUSION FROM THE TWO EXPERIMENTS

In the above discussions, some consistent acoustic features of the tones namely, fundamental frequency, total length, and component durations were explained. Of these three properties, the consistency of component duration is only of small significance. Only fundamental frequency and total length can be regarded as consistent factors. It may also be of interest to examine whether fundamental frequency or total length is the more consistent parameter of the tones. So far as these two experiments with the utterances of word and phrase level are concerned, it is difficult to determine which of the two factors is more consistent. However, a close look at the mingograms and sonagrams of the sentences shows that total length is more consistent than fundamental frequency. In the mingograms of the sentences, the fundamental frequency of Tone II does not rise under certain circumstances.

Though deciding the phonetic cues without any hearing experiment may be open to criticism, it is tempting to say that total length is the more important phonetic cue by which the herearers can distinguish the phonological tones.

8. EXPERIMENT 3

8.1. *SPECTRAL QUALITIES OF THE TONES ON VOCALIC NUCLEI*

In the previous experiments, fundamental frequency, total length and component durations were examined as acoustic properties of the phonological tones. In this analysis, an attempt will be made to examine the different spectral qualities of the vowels with different tones. The 1,200 sonagrams obtained from experiment 2 were used again for this analysis.

8.2. *MONOPHTHONGAL NUCLEI*

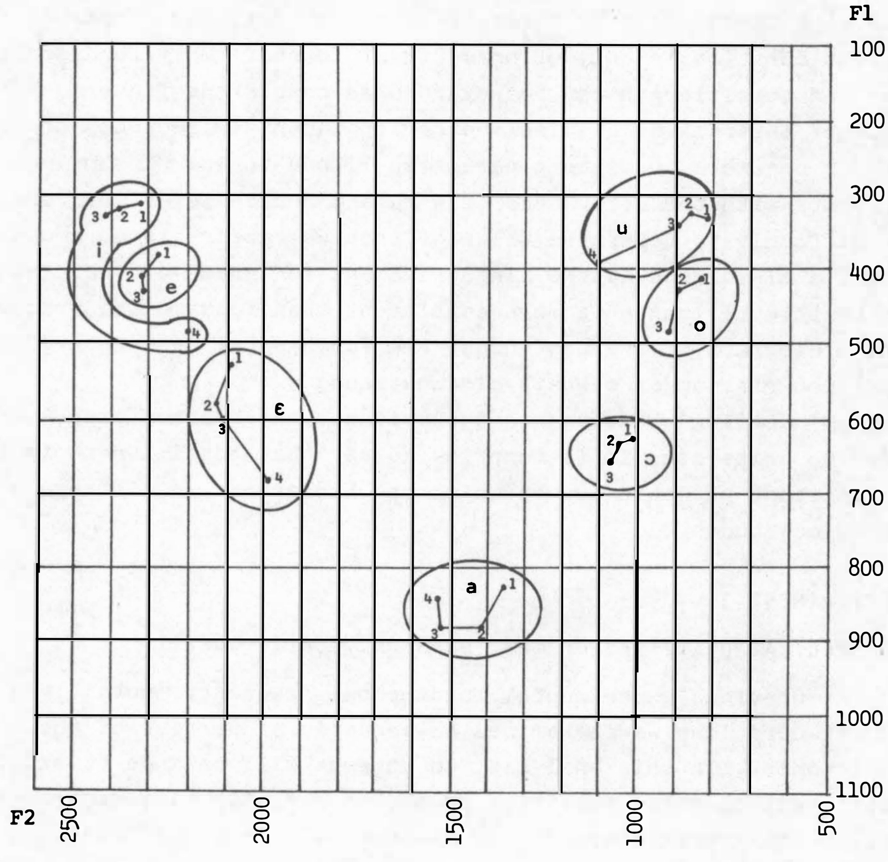
For all the monophthongal nuclei ([i, ī, e, ε, a, ã, ɔ, o, u, ũ]) three lower formants of all the sonagrams in all the different tones were measured at their steady state positions. All monophthongal nuclei can be explained under two categories, that is, non-nasalised monophthongal nuclei and nasalised monophthongal nuclei.

8.3. *NON-NASALISED MONOPHTHONGAL NUCLEI*

The F3 steady state frequency values of all monophthongal nuclei do not show any significant or consistent rise or fall in the four different tones. The F1 and F2 steady state frequency values of all monophthongal nuclei, however, show a fairly consistent rise on the average from Tones I to II and II to III. Graph 1 shows the average positions of the seven monophthongal nuclei in different tones for twelve subjects.

GRAPH 1

Movements of the Seven Basic Vowels in Different Tones Traced on their Average F1 and F2 Steady Values Calculated among the Twelve Subjects.



/i/ moves gradually forward and downward from Tones I to II and II to III. In other words, for /i/ both F1 and F2 frequency values rise gradually from Tones I to II and II to III. /i/ in Tone IV, however, as against the general tendency, is less peripheral than in the first three tones; so much so that it is even lower than /e/ in the first three tones.

/e/ in Tone II moves slightly forward and very noticeably downward from Tones I to II, that is, both F1 and F2 frequency values rise. /e/ in Tone III moves downward very slightly from /e/ in Tone II, that is, F1 frequency keeps rising and F2 frequency is more or less the same.

/ɛ/ moves downward and slightly forward from Tone I to Tone II, that is, both F1 and F2 frequency values rise. /ɛ/ in Tone III moves slightly downward from /ɛ/ in Tone II, that is, F1 frequency slightly increases. /ɛ/ in Tone IV, however, as against the general tendency, is less peripheral (much further back and lower and thus far more central) than /ɛ/ in the first three tones, that is, F1 frequency significantly increases and F2 frequency significantly decreases.

/a/ in Tone II moves downward and forward from /a/ in Tone I, that is, both F1 and F2 values increase. /a/ in Tone III moves horizontally forward from /a/ in Tone II, that is, F2 frequency significantly increases and F1 frequency is more or less the same. /a/ in Tone IV, however, as against the general tendency, moves upward and very slightly forward from /a/ in Tone III, that is, F1 frequency decreases.

/ɔ/ and /o/ become more fronted and lower gradually from Tones I to II and II to III, that is, both F1 and F2 frequency values increase gradually.

/u/ like /ɔ/ and /o/ has a gradual more frontal and lower movement from Tones I to II and II to III except for the fact that /u/ in Tone II tends to move slightly upward from /u/ in Tone I, that is, F1 frequency tends to decrease slightly.

From the points observed so far, it can be concluded that the non-nasalised monophthongal nuclei have a fairly strong tendency to move forward and downward from tones I to II, II to III, and III to IV, that is, both F1 and F2 values increase gradually from Tones I to II, II to III, and III to IV. There are three cases, however, which significantly differ from that general tendency, viz. /i/, /ɛ/ and /a/ in Tone IV.

This kind of deviation from the general tendency may be due to the influence of the final stop.

The following table summarises the average steady state frequency of the lower three formants among the twelve subjects, for all non-nasalised monophthongal nuclei.

TABLE 8

TABLE OF AVERAGE STEADY STATE FORMANT FREQUENCY

Basic Symbol		Tone I	Tone II	Tone III	Tone IV
/i/	F1	302	328	328	470
	F2	2334	2393	2419	2200
	F3	3095	3106	3078	2847
/e/	F1	372	410	424	
	F2	2278	2324	2315	
	F3	2880	2894	2880	
/ɛ/	F1	545	573	599	675
	F2	2093	2135	2111	1994
	F3	2721	2794	2761	2705
/a/	F1	833	876	876	847
	F2	1361	1411	1523	1529
	F3	2515	2570	2475	2492
/ɔ/	F1	632	639	665	
	F2	1014	1052	1074	
	F3	2527	2561	2541	
/o/	F1	412	435	481	
	F2	836	893	919	
	F3	2633	2660	2640	
/u/	F1	339	325	352	497
	F2	812	870	896	1107
	F3	2541	2574	2544	2502

8.4. *NASALISED MONOPHTHONGAL NUCLEI*

The three nasalised monophthongal nuclei (/ĩ, ã, ũ/) occur only in the first three tones. Each of their steady state positions, in any of the three tones, has its own wide area of random scattering when plotted on the graph for the twelve subjects. In spite of these large variations, they still keep their own distinctive area in all the three tones in which they occur. This is possible, mainly because there are only three nasalised monophthongal nuclei and thus there is a great amount of potential range for variation.

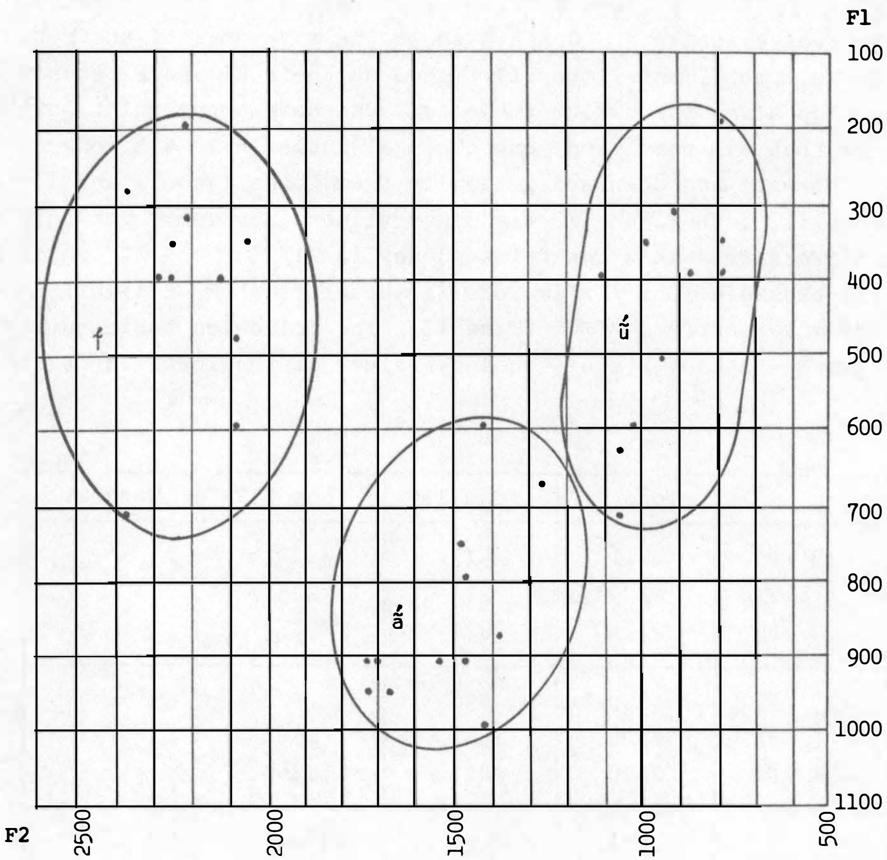
Graphs 2, 3, and 4 show their steady state positions in Tones I, II and III respectively, traced on their F1 and F2 frequency values, for all the twelve subjects. Graph 5 shows the movements of their positions from Tones I to II and II to III traced on their F1 and F2 steady state frequencies among the twelve subjects. One noteworthy point on this graph is that all nasalised monophthongal nuclei have a strong tendency to move forward and downward gradually from Tones I to II and II to III; that is, both F1 and F2 frequency values increase. The only case which differs from this general tendency is /ĩ/ in Tone III which moves slightly backward from /ĩ/ in Tone II though it shows a significant downward movement from /ĩ/ in Tone II. The following table summarises their average steady state frequency values of the lower three formants.

TABLE 9

		Tone I	Tone II	Tone III	Tone IV
[ĩ]	F1	411	415	441	
	F2	2209	2274	2257	
	F3	2821	2920	2877	
[ã]	F1	850	890	922	
	F2	1522	1558	1568	
	F3	2479	2471	2545	
[ũ]	F1	434	451	451	
	F2	929	1028	1067	
	F3	2620	2593	2554	

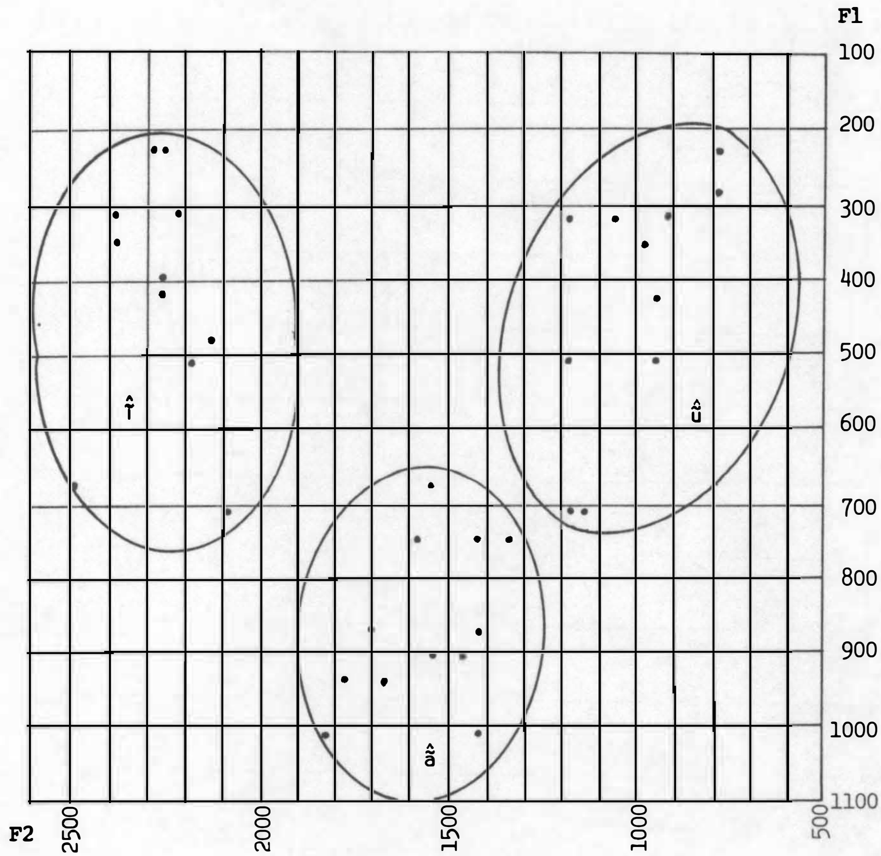
GRAPH 2

The Three Nasalised Vowels in Tone I



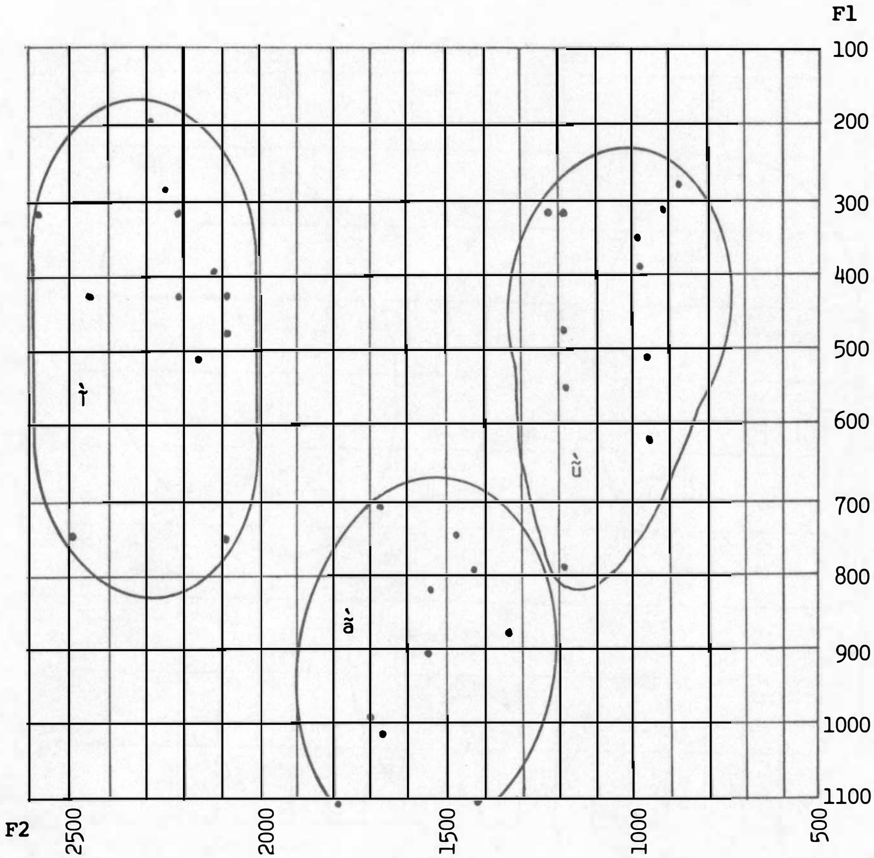
GRAPH 3

The Three Nasalised Vowels in Tone II



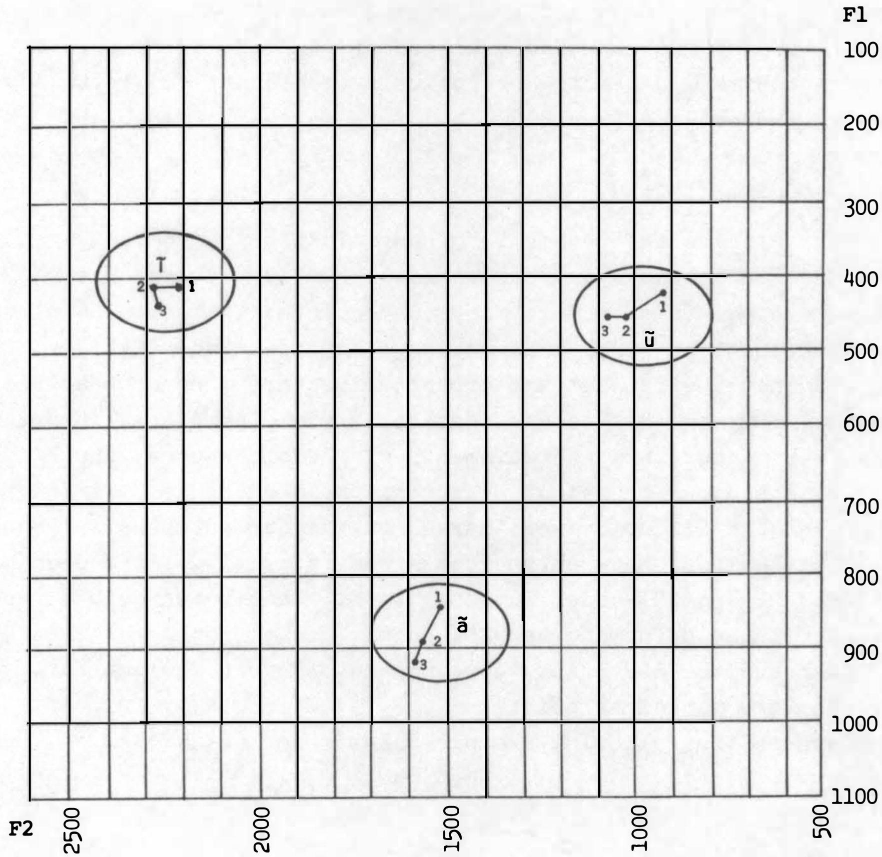
GRAPH 4

The Three Nasalised Vowels in Tone III



GRAPH 5

Movements of the Three Nasalised Vowels,
in Different Tones, Traced on their Average
F1 and F2 SS Values Calculated Among
the Twelve Subjects



8.5. *DIPHTHONGAL NUCLEI*

One remarkable feature shown by the Burmese diphthongs on the sonagrams is that they hardly have any onglide and offglide whether they are glottalised or nasalised in any tone. If the diphthongs show occasional onglides and offglides, their lengths are usually very limited.

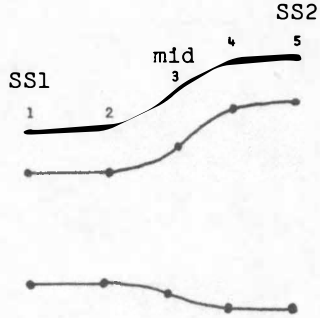
Another remarkable feature of the diphthongs is that their SS1 and SS2 are very distinct from one another both in nasalised and final-stop environments. In some cases the SS lengths may be extremely short, but they still maintain their own distinctive positions on the formant bars.

These two features show that unlike the monophthongal nuclei, the vowel qualities of the diphthongs are so strong that there is very little effect of the preceding or following consonant or glottalisation, on them, in [hṽvndá] or [hvvˀdá] (/hṼdá/ or /hVˀdá/) frame.

8.6. *METHOD OF MEASURING*

In order to trace the movements of the diphthongs from one steady state to the other, acoustic phoneticians usually measure onset, SS1, SS2 and offset points of the formants under investigation. As almost all the diphthongs in Burmese do not have onglides and offglides, it can be said that their onset and offset values are more or less the same as their SS1 and SS2 values respectively. Therefore, in order to investigate accurately the movements of the diphthongs, the writer had to devise his own method of measuring the diphthongs. Firstly the SS1 point and the SS2 point were marked on the formant bars. Then a point exactly halfway between the two steady states was marked on every formant (F1, F2 and F3) under investigation. Lastly two points, one between the middle point and the SS1 point, and the other between the middle point and the SS2 point, were marked on every formant bar. The five points were marked as point 1, 2, 3, 4, 5, serially from SS1 to SS2 as shown in the following example diagram for [ei].

[ei]



All the four diphthongs ([eiʔ], [auʔ], [ouʔ], [aiʔ]; and [ẽĩ], [ãĩ], [ãũ], [õũ]) occur in both nasalised and final-stop environments. In the nasalised environment, they occur only in the first three tones. What follows is an explanation of their movements from one SS to the other, traced on their average F1 and F2 frequencies measured at the five selected points.

[ẽĩ] and [eiʔ]

The movements of [ẽĩ] in Tone I and Tone II are more or less the same except that SS1 in Tone I is slightly lower than that in Tone II. The movement of [ẽĩ] in Tone III is lower (i.e. F1 increases) than those in Tones I and II especially at points 1, 2, 3, and 4. The movement of [eiʔ] is almost the same as [ẽĩ] in Tone III. (See Graphs 6 and 7).

[ãĩ] and [aiʔ]

The movements of [ãĩ] from SS1 to SS2 in the three tones and that of [aiʔ] are more or less the same (see Graphs 6 and 8).

[ãũ] and [auʔ]

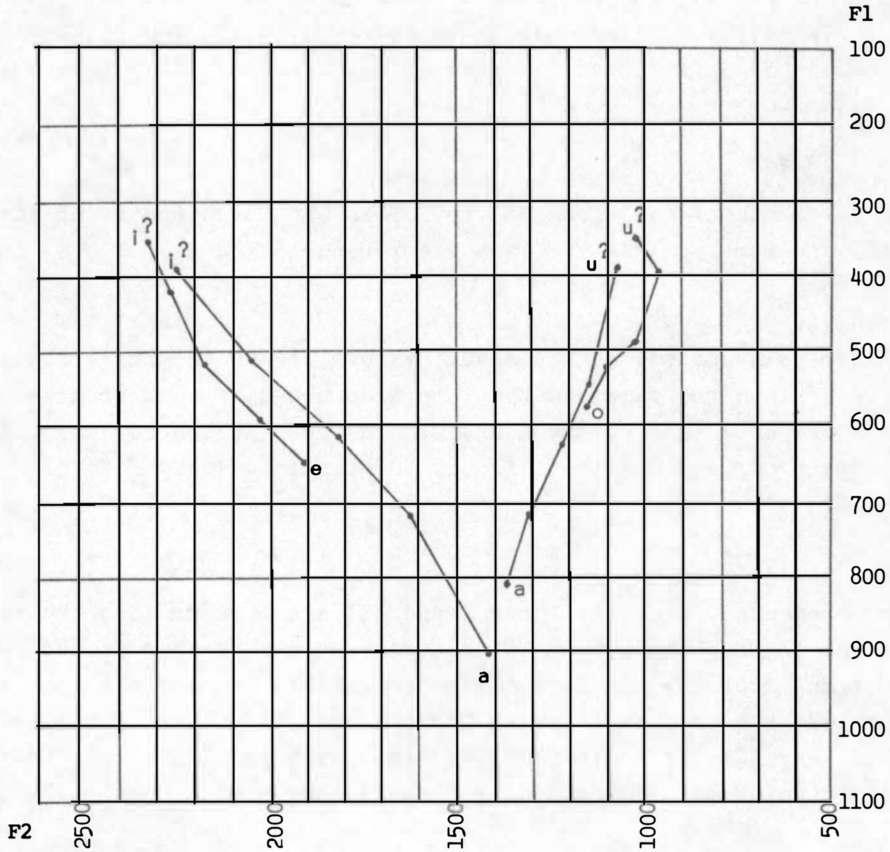
As with [ãĩ] and [aiʔ], the movements of [ãũ] in the three tones are more or less the same and they are also not different from the movement of [auʔ] (see Graphs 6 and 9). However it can be said that the movement of [ãũ] becomes slightly fronter from Tones I to II and II to III.

[õũ] and [ouʔ]

The movements of [õũ] in Tones I and III are more or less the same except for the fact that Tone III has some points further front. Points 1 and 2 of [ãũ] in Tone II are remarkably low and the rest are more or less the same as those in Tones I and III. The movement of [ouʔ] is very similar to that of [ãũ] in Tone I and III except that the SS2 point (point 5) moves slightly forward (see Graphs 6 and 10).

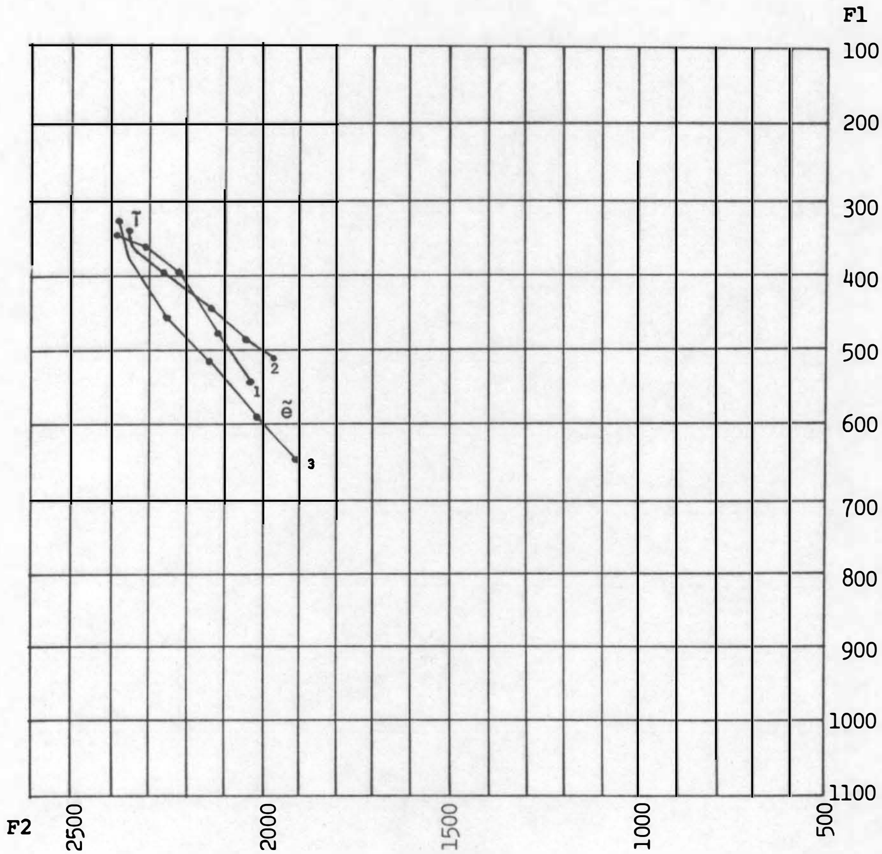
GRAPH 6

Movements of Glottalised Diphthongs (Diphthongs in Tone IV), Traced on the Average F1 and F2 Values at the Five Selected Points



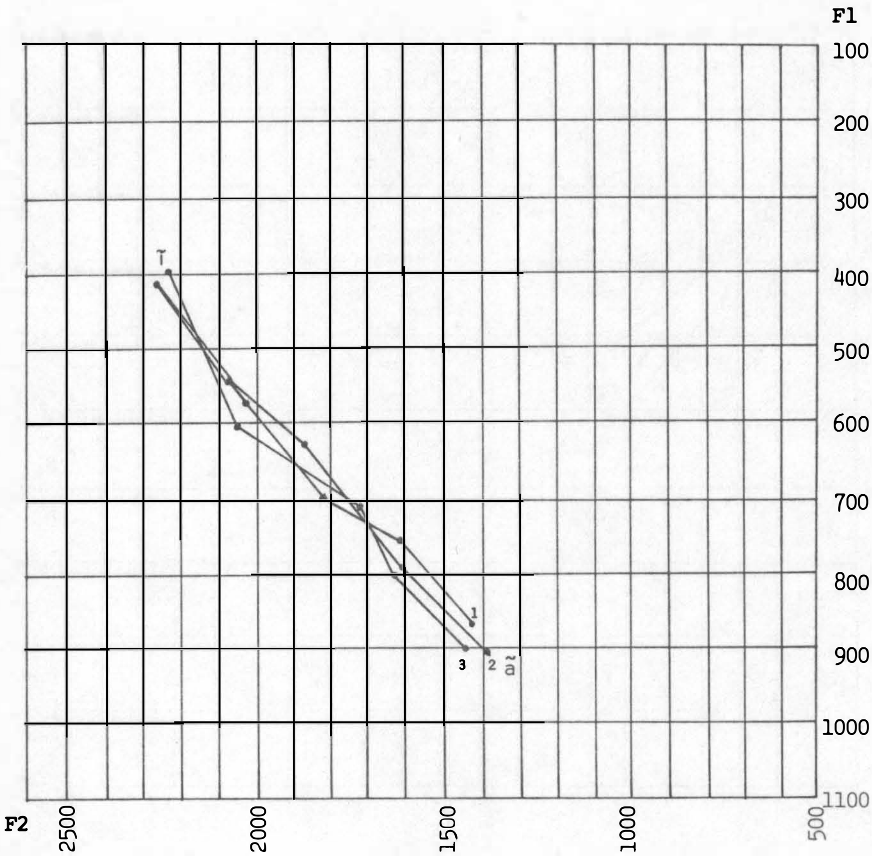
GRAPH 7

Movements of the Nasalised Diphthong [ẽĩ] in Different Tones, Traced on the Average F1 and F2 Values at the Five Selected Points



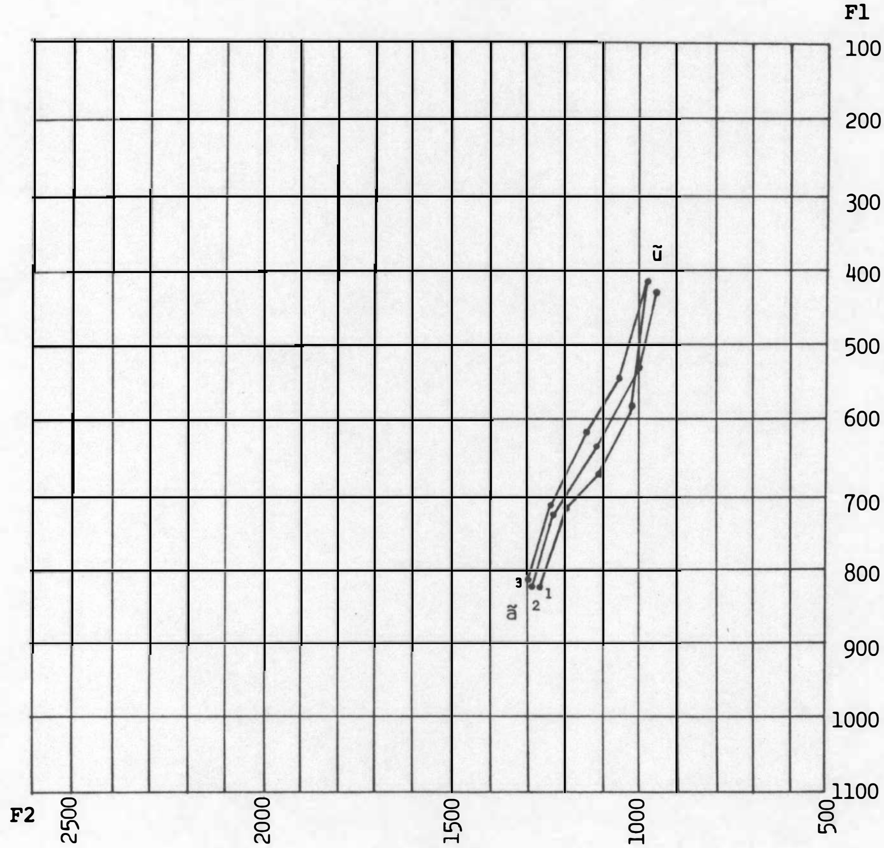
GRAPH 8

Movements of the Nasalised Diphthong [ãĩ] in Different Tones, Traced on the Average F1 and F2 Values at the Five Selected Points



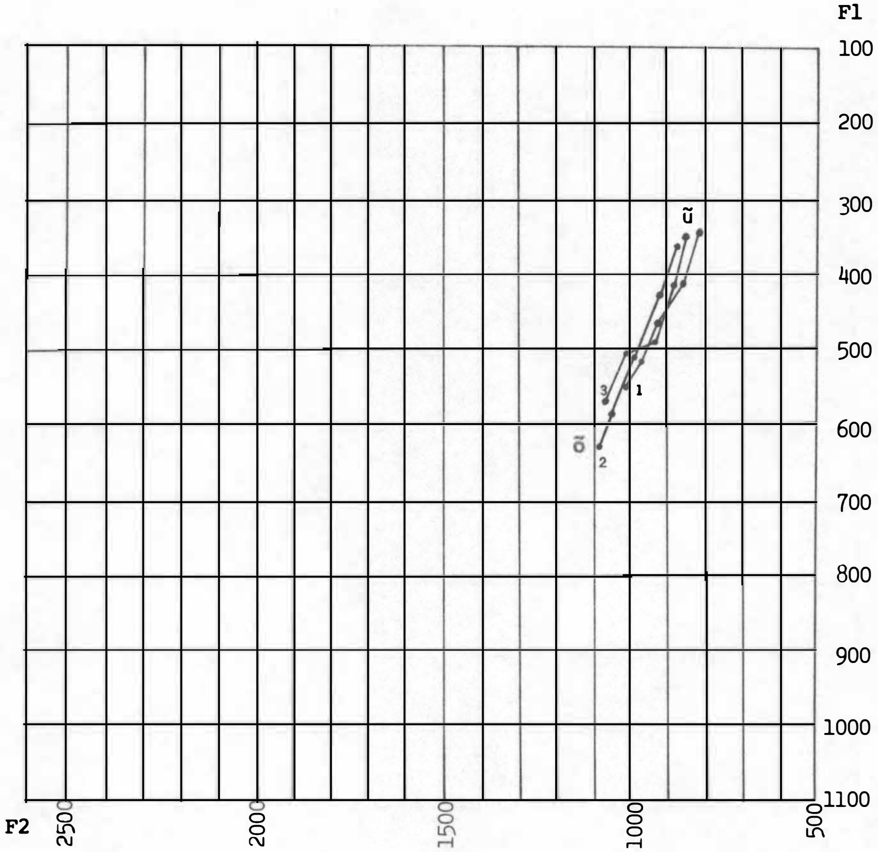
GRAPH 9

Movements of the Nasalised Diphthong [ãũ] in Different Tones, Traced on the Average F1 and F2 Values at the Five Selected Points



GRAPH 10

Movements of the Nasalised Diphthong [õũ] in Different Tones, Traced on the Average F1 and F2 Values at the Five Selected Points



Although the non-nasalised and nasalised monophthongal nuclei generally show the general tendency of moving forward and downward from Tones I to II, II to III and III to IV, the diphthongal nuclei do not. It is true that the movement of [e͠i] in Tone III is more fronted and lower than those of [e͠i] in Tones I and II and the movement of [a͠u] becomes more fronted from Tones I to II, and II to III. The movements of [a͠i] and [o͠u] from Tones I to II and II to III do not show any uniformity or consistency at all. Of the four stop-final diphthongal nuclei [auʔ] is the only one whose movement is more fronted than that of its nasalised counterpart [a͠u] in the first three tones. The following tables summarise the average frequencies of the stop-final and nasalised diphthongs in the three tones, measured at the five selected points for the twelve subjects.

TABLE 10

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[eiʔ]	F1	652	596	515	415	341
	F2	1908	2039	2173	2267	2320
	F3	2711	2781	2827	2883	2893
[aiʔ]	F1	903	728	625	510	382
	F2	1427	1638	1829	2059	2256
	F3	2376	2472	2548	2643	2729
[auʔ]	F1	576	715	633	558	385
	F2	1377	1315	1229	1166	1087
	F3	2534	2531	2544	2530	2482
[ouʔ]	F1	576	527	491	395	349
	F2	1171	1107	1022	962	1038
	F3	2577	2601	2611	2558	2531

TABLE 11

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[éi]	F1	542	473	397	365	349
	F2	2043	2130	2224	2315	2396
	F3	2715	2798	2800	2913	2897
[êi]	F1	516	480	454	395	345
	F2	1961	2056	2168	2267	2360
	F3	2712	2758	2798	2851	2930
[èi]	F1	652	583	524	461	339
	F2	1904	2036	2151	2264	2383
	F3	2442	2758	2898	2841	2893

TABLE 12

TABLE OF DEVIATION FROM TONE I

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[êi]	F1	-26	+7	+57	+30	-4
	F2	-82	-74	-56	-48	-36
	F3	-3	-40	-2	-62	+33
[èi]	F1	+110	+110	+127	+96	-10
	F2	-139	-94	-73	-51	-13
	F3	-273	-40	+98	-72	-4

TABLE 13

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[ái]	F1	876	754	692	567	408
	F2	1427	1618	1832	2033	2260
	F3	2419	2518	2600	2712	2829
[âi]	F1	919	797	629	563	402
	F2	1393	1605	1881	2089	2262
	F3	2455	2508	2574	2670	2818
[ài]	F1	899	790	711	602	395
	F2	1440	1625	1738	2050	2237
	F3	2366	2507	2570	2649	2787

TABLE 14

TABLE OF DEVIATION FROM TONE I

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[âi]	F1	+43	+43	-63	-4	-6
	F2	-34	-13	+49	+56	+2
	F3	+36	-10	-26	-42	-11
[ài]	F1	+23	+36	+19	+35	-13
	F2	+13	+7	-94	+17	-23
	F3	-53	-11	-30	-63	-42

TABLE 15

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[áu]	F1	833	714	675	576	428
	F2	1268	1197	1110	1037	981
	F3	2508	2514	2540	2534	2593
[âu]	F1	830	734	648	536	444
	F2	1285	1215	1123	1008	968
	F3	2557	2606	2649	2679	2692
[àu]	F1	826	714	619	550	421
	F2	1288	1232	1143	1059	978
	F3	2491	2534	2580	2603	2471

TABLE 16

TABLE OF DEVIATION FROM TONE I

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[âu]	F1	-3	+20	-27	-40	+16
	F2	+17	+18	+13	-29	-13
	F3	+49	+92	+109	+145	+99
[àu]	F1	-7	nil	-56	-26	-7
	F2	+20	+35	+33	+22	-3
	F3	-17	+20	+40	+69	-122

TABLE 17

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[óu]	F1	556	523	471	414	348
	F2	1034	978	929	866	823
	F3	2626	2610	2611	2629	2656
[ôu]	F1	569	507	484	421	355
	F2	1064	1011	942	889	856
	F3	2623	2671	2698	2698	2686
[òu]	F1	642	569	517	438	365
	F2	1090	1050	984	912	869
	F3	2560	2596	2613	2656	2609

TABLE 18

TABLE OF DEVIATION FROM TONE I

		SS1	TRANSITION			SS2
		Point 1	Point 2	Point 3	Point 4	Point 5
[ôu]	F1	+13	-16	+13	+7	+7
	F2	+30	+33	+13	+23	+33
	F3	-3	+36	+60	+69	+30
[òu]	F1	+86	+46	+46	+24	+17
	F2	+56	+72	+55	+46	+46
	F3	-66	-14	+2	+27	-47

9. GENERAL CONCLUSION FROM THE THREE EXPERIMENTS

Experiments 1 and 2 reveal that fundamental frequency and total length of the tones can be regarded as remarkably consistent features.

Experiment 3 shows that all monophthongal nuclei (nasalised and non-nasalised) show a strong tendency to have lower and more fronted allophones from Tones I to II and II to III, that is, F1 and F2 values increase. The diphthongal nuclei, however, do not show this general tendency, at least not as significantly as the monophthongal nuclei do. In other words, of the fourteen vocalic nuclei (seven basic vowels, three nasalised vowels and four diphthongs), ten (all monophthongal nuclei, that is, seven basic vowels and three nasalised vowels) show the tendency and four (all diphthongs), do not. As ten out of fourteen vocalic nuclei (71.4%) show the tendency, it is justified to regard this tendency of spectral quality as another consistent parameter of the tones.

Two other features of the tones, that is, fundamental frequency and total length, seem more consistent than the feature of spectral quality. Within the /hVdá/ test frame, the duration pattern of the tones seems to be more consistent than the fundamental frequency. Therefore it can be concluded that the duration pattern is the most consistent feature and the fundamental frequency is the second most consistent feature of the Burmese tones.

BIBLIOGRAPHY

BECKER, A.S.

1964 Burmese tonemics: a discussion of various points of view. In J.K. Yamagiwa, ed., *Papers of the CIC Far Eastern Language Institute*. Ann Arbor: University of Michigan.

BURLING, R.

1967 *Proto Lolo-Burmese*. Publication 43 of Indiana University Publications in Anthropology and Linguistics. Bloomington.

CORNYN, W.S.

1944 *Outline of Burmese grammar*. Language 20/4, supplement. Language dissertation No.38.

CORNYN, W.S. and D.H. ROOP

1968 *Beginning Burmese*. New Haven: Yale University Press.

FIRTH, J.R.

1933 Notes on the transcription of Burmese. *Bulletin of the School of Oriental Studies* 7/1:137-140.

1936 Alphabets and phonology in India and Burma. *Bulletin of the School of Oriental Studies* (Grierson memorial volume) 8/2:517-546.

McDAVID, R.

1945 Burmese phonemics. *Studies in Linguistics* 3/1:6-17.

OKELL, J.A.O.

1969 *A reference grammar of colloquial Burmese*. 2 vols. London: Oxford University Press.

STEWART, J.A.

1955 *A manual of colloquial Burmese*. 2nd edn. London: Luzac.

TAYLOR, L.F.

1920 On the tones of certain languages of Burma. *Bulletin of the School of Oriental Studies* 1/4:91-106.

THEIN TUN, U

1976 A survey and critique of previous studies on Burmese phonology. *Working papers of the Speech and Language Research Centre, Macquarie University*. January, 95-124.

TRUBETZKOY, N.S.

1968 *Introduction to the principles of phonological description.* Translated by L.A. Murray. The Hague: Martinus Nijhoff.

1969 *Principles of phonology.* Translated by Christaine A.M. Baltaxe. Berkeley and Los Angeles: University of California Press.

REGISTER IN BURMESE

David Bradley

0. INTRODUCTION[1]

It has been 'traditional' in Tibeto-Burman (TB) linguistics to describe languages in terms of a suprasegmental opposition of **tones** which have been talked about mainly or exclusively in terms of fundamental frequency (or pitch). The literature on the development of these tones refers to reconstructed characteristics of consonant segments in the syllable: initial consonants conditioning pitch height, and certain 'laryngeal' final consonants often conditioning the development of pitch contour.

To be exact, proto-voiced initials have been found to condition the development of a lower pitch, and proto-voiceless initials may condition the development of a higher pitch; there are also cases in which three categories of initials - such as voiceless aspirated, unaspirated, and voiced - may condition the development of three different relative pitches. A final -ʔ may condition the development of a rising contour, and a final -h, a falling contour. These proposals are discussed succinctly in Matisoff (1973); an example of the development of higher versus lower tones conditioned by *voiceless versus *voiced initials which subsequently themselves changed is seen in Matisoff (1972), and an example of a rising tone which developed in the environment of a recently-developed final glottal stop (as well as a difference in initial) is seen in Matisoff (1970). These two examples from Loloish are examples of excellent comparative work which demonstrate that recent developments in the suprasegmental systems of these languages have been conditioned by segmental differences.

In Austroasiatic (AA) linguistics, the focus of descriptive and comparative efforts has instead been on suprasegmentals of voice quality or phonation, which have been termed register by Henderson (1952).

Bradley, D. "Register in Burmese". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation*.
A-62:117-132. Pacific Linguistics, The Australian National University, 1982. DOI:10.15144/PL-A62.117

These differences between breathy, 'normal', and creaky voicing have been the main parameter considered in many descriptions, though there are often other characteristics noted as well. The conditioning factor in the development of the register opposition, according to the usual reconstruction, is the voicing characteristic of the initial consonant in the syllable; if the initial was voiced, a breathy phonation might develop for the entire syllable; if voiceless, a creaky phonation might develop. It seems much less typical of AA languages to have more than two contrasting suprasegmentals of register than it is of TB languages to have more than two tones conditioned by initial characteristics. Moreover, it is often the case that Austroasiatic languages with a register contrast have only one 'marked' phonation; i.e. there are systems with breathy versus 'normal' register, or 'normal' versus creaky register, but it is less frequent to find a system with breathy vs. creaky register.

It is interesting to note that the principal conditioning factor adduced in many instances of tonogenesis or tone splitting in TB languages is identical to that for the development of register in AA languages: voiceless versus voiced initial consonants. Perhaps the differences in description are partly due to the fact that TB and AA linguists don't talk to each other enough.

As Henderson has also noted, it is hardly ever the case, in Southeast Asia at least, that a suprasegmental system can be described in terms of one parameter only. Fundamental frequency or pitch is one parameter only; so is phonation. Other parameters of suprasegmentals whose domain is the syllable (syllable prosodies) may include duration, intensity, and characteristics of the segments such as vowel quality. Gregerson (1976) suggests that the position of the tongue root may be a major factor in several of these parameters including especially voice quality overall, and vowel quality. Other articulatory characteristics involved would include larynx adjustments: raising or lowering as well as differences in vocal cord tension and so on; airstream differences (more or less subglottal pressure); pharynx shape differences, partly but not only related to the position of the tongue root; and possibly more. The principal differences in articulation involve the larynx, but the shape of the vocal tract above the larynx can also contribute significantly.

There have been a couple of studies within TB which propose to account for suprasegmental phenomena in terms of register; the earliest that I know of is Bradley (1969), which describes the Loloish TB language Akha in terms of a laryngealised versus 'normal' register which crosscuts the high versus mid versus low pitch tone system; the laryngealised register occurs mainly with mid and low tones. It is a true syllable prosody in that initials are unaspirated in laryngealised register and aspirated (or slightly breathy if voiced) in 'normal' register; there is also variably a final glottal-stop in laryngealised register syllables. The diachronic sources of this prosody are final stops: *-p, *-t, or *-k, which had probably merged to a final glottal stop whose characteristics spread into the preceding initial and vowel and produced the laryngealised register. For a further description of Akha register, see Thurgood (1980).

Another proposal of register in TB is Glover (1971), who accounts for the suprasegmental systems of three related languages in Nepal (Gurung, Tamang, and Thakali) in terms of two crosscutting parameters: phonation, clear versus breathy; and effort, more versus less intense. The register contrast is related by Glover to a preliminary version of Gregerson's (1976) advanced tongue root hypothesis; Glover reports that vowels are sometimes higher in 'normal' than in breathy register in Gurung; and that pitch is affected by both suprasegmentals, with higher fundamental frequency in the more intense syllables, and lower in the less intense syllables; and relatively lower fundamental frequency overall in the breathy syllables. The three languages differ somewhat in the exact details of pitch characteristics, but there is an exact correspondence between them. This correspondence pattern has important implications for the reconstruction of Proto-TB suprasegmentals, as the three languages are genetically quite close to Tibetan. Shafer's argument against the reconstruction of tones is based on the secondary nature of Tibetan tones, but if the Gurung/Tamang/Thakali system shows regular correspondences to Benedict's Tones *A and *B, then it must be the case that Tibetan has **lost** the Proto-TB suprasegmental opposition relatively recently. Mazaudon (1977) discusses this and other factors in the history of Tamang and other TB languages.

Weidert (1979) proposes to reconstruct final laryngeal segments rather than tones for Proto-TB based on data from Kuki-Chin, Naga, Baris (Bodo-Garo), and Jinghpaw. What he is suggesting amounts to a claim that Tone *A was unmarked, Tone *B had a final glottal-stop, and that other secondary tones that developed within these subgroups of TB arose from finals *-h, *-s, and *-p/ *-t/ *-k. There is some evidence of -ʔ in the tone which is the reflex of *B in Garo, Bodo, Lotha Naga, Mikir, Nocte, and Tangsa, especially in forms which occur in isolation; in fact this could also be related to a register difference. Unfortunately for those who would like to reconstruct register for Proto-TB, it is precisely the opposite phonation, **breathy**, which occurs in the Burmese reflex of TB Tone *B.

In an extremely important paper, Egerod (1971) proposes a similar segment-derived register origin for the tone system of Chinese; Chinese is of course the other major group within Sino-Tibetan (ST), along with TB. The even (ping) tone is reconstructed as 'normal' register, the rising (shang) tone is reconstructed as creaky in phonation, with a final *-ʔ which, in accord with the usual TB tonogenesis principles, had also a rising contour. Similarly the going (qu) tone is postulated as having had breathy phonation, developed from a final *-h which also conditioned a falling contour. The entering (ru) tone with final *-p/ *-t/ *-k was not opposed to the first three in earlier stages of Chinese when it was the only possibility in stop-final syllables; the ping, shang, and qu do not occur in stop-final syllables.

ST, and **a fortiori** TB, thus includes several subgroups which have, or have been reconstructed as formerly having, register-type systems. Sinitic, Bodic, Burmic, and Baric, all four of Shafer's major subgroups, include such cases. Within TB, it seems certain that one must reconstruct two suprasegmentals, *A and *B, based on widespread correspondence patterns; it is less clear what the possible realisations of the two may have been.

1. PROSODIES IN BURMESE

Burmese has been analysed as having up to five opposed suprasegmentals, realised as follows (data from personal observations; also Thein Tun (1982)):

name	pitch	contour	intensity	phonation	duration	vowel quality
'even'	low	level	low	normal	fairly long	intermediate
'creaky'	high	slight fall	very high	creaky	less long	higher,more fronted
'heavy'	fairly high	sharp fall	high	breathy	very long	lower,more backed
'killed'	very high	slight fall	high	normal	short	(different system)
(reduced)	variable	variable	very low	normal	very short	[ə] only

The 'reduced' possibility occurs with the so-called 'minor syllable' which is found in various Southeast Asian languages; it occurs only with the vowel [ə], which does not occur with the other suprasegmentals; so it is not opposed to the others and has been excluded from most analyses of Burmese 'tone'. For some details of the origins of this syllable type in Burmese, see Bradley (1980).

The 'killed' type occurs only with a final stop; glottal in isolation, and homorganic to the initial consonant of the following syllable in close juncture. Also, the following consonant is not voiced -

unlike the initial of a syllable in close juncture after the 'even', 'heavy', or 'creaky' types. Like the Chinese ru tone, it is historically derived from syllables with a final stop; the Burmese orthography still represents the positions of these stops, but in modern spoken dialects the features of the stops have been 'shuffled' into the vocalic nuclei. As a result, the vowel system in 'killed' syllables (and in nasalised syllables, which occur 'even', 'heavy', or 'creaky'; and likewise reflect final nasals, etymologically and in spelling) is radically different from that of open syllables that are 'even', 'heavy', or 'creaky'. Because of the differences of juncture, it is possible to regard the 'killed' syllable type as non-contrastive: it is the **only** possibility in syllables with a final stop.

Thus we are down to three 'tones'; the parameters involved in their realisation, as set out in the above table, would in fact enable us to regard **any** of the six parameters as the contrastive one, if one had to choose only one.

Sprigg's analysis (1964) suggests instead a two-tone analysis, high versus low pitch, with a crosscutting register difference on the high tone: glottal (creaky) versus non-glottal. That is, 'even' is the low tone; 'heavy' and 'creaky' are the high tone, 'heavy' being non-creaky and 'creaky' being creaky.

The Indic-derived Burmese orthography treats 'even', 'heavy', and 'creaky', as opposed to each other, and so speakers usually regard these three as the contrasting suprasegmentals in the language. With two oral and all nasalised vowels, 'even' is unmarked; 'heavy' is marked by Indic visarga; and 'creaky' is marked by a subscript dot which originated as a final -ʔ in early inscriptions. With three vowels, 'creaky' is represented by an Indic short vowel; 'even' and 'heavy' are represented by an Indic long vowel, and 'heavy' has the visarga. With the remaining two oral vowels, there is a separate representation for the 'even' version; the separate 'heavy' version therefore does not require the visarga; and the 'creaky' version has the subscript dot added to the 'heavy' version of the vowel. It is interesting that the visarga, which represents a final (breathy) -ɦ in the Indic orthographies, is used to represent 'heavy' which is also breathy. Similarly, the use of a dot derived from a final glottal-stop for 'creaky' may reflect the phonation of this suprasegmental which is of course creaky.

The Burmese orthography was almost certainly devised by Mon monks, speakers of an AA language, about 1100AD. The Indic-derived Mon orthography uses voiced and voiceless initial consonants in cases where modern spoken Mon has breathy and 'normal' register syllables, but whether register was present in Mon about 1100AD is not clear. In any case the Burmese orthography could be reflecting the fact that phonation differences were the most salient features of the suprasegmentals to the Mon monks; or they could have simply been inventing combinations using Indic orthographic resources. It seems likely at least that the invented combinations would have some relation to the parameters of the Burmese suprasegmentals at the time.

In fact the earliest inscriptions are somewhat inconsistent; the use of a short vowel or of a final -ʔ for 'creaky' was very early, but the visarga for 'heavy' came much later, and did not become entirely regular for quite some time. It thus seems likely that the relatively shorter duration and creaky phonation of 'creaky' were both characteristic of Burmese about 1100. The early ambiguity between 'even' and 'heavy' certainly does not represent an absence of contrast, given the regular correspondences between modern Burmese and closely related languages. It is less certain that 'heavy' was already (or still) breathy in 1100, but it certainly is now.

In terms of the comparative picture, Burmese 'even' corresponds to Proto TB Tone *A; and 'breathy' corresponds to Proto TB Tone *B. For details of the system of closely-related Loloish, see Bradley (1977, 1979). Those sources also include some speculations about the phonetic parameters of the suprasegmentals at an earlier stage; in general, phonation differences are rarely found within Loloish in Tone *1 (from *A) nor in Tone *2 (from *B). Within the Burmese-Lolo (BL) family (which includes Burmish and Loloish languages) a third suprasegmental category, reconstructed as Tone *3, has developed; this has as its Burmese reflex the 'creaky', and often has creaky phonation elsewhere. Bradley (1971), relying on TB data in a preliminary version of Benedict (1972), proposes that this Tone *3 developed from certain *s- and *ʔ-prefixed etyma, mainly in Proto-TB Tone *B, via spreading of the prefix characteristic into a creaky phonation of the syllable which further conditioned a difference in pitch as well. The development of BL Tone *3 separates this family from quite closely related TB languages such as Naxi (Bradley (1975), Hsihsia, and Tosu.

2. BURMESE AS A REGISTER LANGUAGE

There are various reasons for regarding register, rather than tone, as the contrastive suprasegmental in Burmese. In terms of the orthographic and terminological tradition, it seems best; almost all of the acoustic parameters involved show characteristics that fit as secondary results of a register contrast, but not of a pitch-based tone system. Changes in several dialects, such as Arakanese, give even more evidence.

Thus it seems that there has been a re-analysis of the Proto-BL pitch-based tone system into the modern Burmese phonation-based register system; this may in part have been triggered by the development of the creaky Proto-BL Tone *3, and was certainly furthered by the assimilation into the population of Burmese speakers of very large numbers of Mons, speakers of an AA language with a register contrast.

The traditional orthography, as noted above, uses -h for the breathy register; leaves the 'normal' register unmarked; and uses a dot derived from -ʔ or a short vowel for the creaky register. This clearly supports the register analysis, as do the linguistic terms for the categories: 'creaky' for the creaky register, 'even' for the 'normal' register, and 'heavy' for the breathy register. Finally, the orthography-based order of citation of the three registers is hierarchically arranged in terms of degree of vocal cord structure: most to least, viz. 'creaky', 'even', 'heavy'.

Perhaps the most telling factor in favour of the register analysis is the fact that vowel height and position, as measured by Formant 1 and Formant 2, correlate quite strongly with the suprasegmentals of Burmese. It has often been stated that tones, i.e. pitch differences, do not have any effect on vowel quality; exceptions are explained in Thurgood (1980) as due to phonation as a concomitant parameter included in the 'tone'. The pervasive, systematic differences in vowel quality measured by Thein Tun (1982) in the Burmese of four speakers, two male and two female, force the conclusion that register must be involved.

Nearly all oral non-stop-final and nasalised vowels of Burmese are highest and most backed in 'creaky', lowest and most fronted in 'heavy' i.e. breathy register, and intermediate in 'even' i.e. normal register. This difference is particularly great for /a/, which is far more fronted in the breathy register, but is significant for all except /u/ which has a slightly lower (14Hz difference) F1 in 'even' than in 'creaky' and may thus be slightly higher.

In the Arakanese dialect spoken in western Burma, the vowel quality difference seems impressionistically to be even greater; there are very large differences, especially of vowel height. On the whole, both of the 'marked' phonations, creaky and breathy, tend to condition higher vowel allophones; while 'normal' phonation tends to condition lower vowel allophones. Thus Arakanese differs from Burmese in that breathy register, like creaky register, seems to be 'advanced tongue root' [+ATR], while in Burmese only the creaky register is [+ATR], and there seems rather to be a continuum of degrees of [ATR] from creaky [+ATR] to normal [0 ATR] to breathy [-ATR].

The extreme differences of vowel quality in Arakanese have had several repercussions. One is that the vowel corresponding to Burmese /e/ has split and merged with /i/ in breathy and creaky register, and with /ɛ/ in 'normal' register; this process is confused by the continuing contact with, and prestige of, 'standard' Burmese in Arakan. Further, the new /i/ of Arakanese which corresponds to Burmese /e/ is nasalised after a nasal initial, as is the original /i/ which corresponds to Burmese /i/. Another example is that the extremely back /a/ in breathy register has fused with a preceding /w/ into a new vowel nucleus, Arakanese /ɒ/, which occurs almost exclusively in breathy register, but in all three registers when nasalised. These kinds of vowel differences conditioned by tone are unheard of; hence Arakanese and Burmese seem to have the phonation difference as the primary one; the secondary Arakanese developments suggest that this has been the case for a long time.

Another parameter which supports register as the contrastive suprasegmental is intensity. If the system were tonal, one might expect a linear relationship between pitch and loudness, such that the 'creaky', which is the highest in pitch, would be the loudest; the 'even', which is the lowest in pitch, would be the softest. If the system is a register one, then the loudest would be the 'heavy' which has greater airflow because of the nature of the phonation; the least loud would be the 'creaky' again due to the phonation characteristics. In fact the 'heavy' is the loudest, not the 'creaky', supporting the register theory. However, the 'even' is the lowest in **peak**

Order of Intensity if		**Actual Intensity**	
register → intensity	tone → intensity	(peak)	(total)
'heavy'	'creaky'	'heavy'	'heavy'
'even'	'heavy'	'creaky'	'even'
'creaky'	'even'	'even'	'creaky'

intensity, perhaps due to the fact that inherently less effort is involved in 'normal' phonation, so the two registers with 'marked' phonation both have greater peak intensity. Of course, because of its shorter duration the 'creaky' has less **total** intensity than the 'even'. Thein Tun (1982) does not measure intensity, as his data were nearly all two-syllable words in isolation and the result was unnatural variation in intensity which he feels not to be characteristic of natural speech. The intensities considered here are based on my measurements of two speakers from Mandalay, one male and one female, using a Frokjaer-Jensen Intensity Meter and recording the output with an Elema mingograph.

Duration is a parameter which is often independent of tone and/or register: there are tone languages and register languages with contrastive length as well. Burmese has no contrastive length; the diphthongal nasalised vowels are somewhat longer than the monophthongal nasalised vowels in non-stop-final syllables; and syllables with a /w/ onglide may be longer still, especially if the vowel nucleus is a nasalised diphthong.

oral		nasalised		stop-final	
i	u	(w)ĩ	ɷ̃	ι	ɷ
(w)ẹ	ọ	ẹ̃ι	ọ̃ɷ	ẹι	ọɷ
(w)ɛ̣	ɔ̣	ãι	ãɷ	(w)ɛ̣	
(w)a		ã		aι	aɷ
				a	

In Thein Tun (1982) there are various interesting findings concerning the vowel durations; the relative durations are summarised in the parameters chart above. Overall, 'heavy' syllables are the longest, 21 centiseconds on average; 'even' syllables average 18.5 csec; and 'creaky' syllables average only 15.4 csec in duration. The stop-final syllables, by contrast, average about 10.3 csec in duration; this figure is raised by the fact that half of them are diphthongs, while only four of the fourteen syllables with the three-way suprasegmental opposition are.

Interesting comparisons can also be made between oral non-stop-final vowels and other monophthongs: in 'even' syllables, for example, the oral vowels average 20.5 csec; the three nasalised monophthongs average just under 14 csec; and the four stop-final monophthongs average 9.9 csec in duration; a factor of more than 2:1 length difference between the extremes. Another intriguing finding is that the duration of vowels is also related to their height: opener (lower) vowels have longer durations, other things being equal. In summary, it seems there are various factors involved in duration which may simply occur together with the other parameters. It could perhaps be argued that the shorter duration of 'creaky' relates to its tighter muscle constriction in the vocal cords; and conversely that 'heavy' is longer because of slacker vocal cord constriction, so duration too could be used as evidence in favour of register than tone in Burmese.

One might think that fundamental frequency *per se* is unlikely to provide evidence for a register analysis, but even here some useful parallels can be drawn. For example, the fact that the 'heavy' has a sharply falling contour could be related to its breathy nature; after all, -h often conditions a falling contour. Similarly, the relative absence of contour in the 'even' argues for its unmarked quality. The slight fall of the 'creaky' is harder to fit into this picture.

To a similar degree, relative pitch of the suprasegmentals could be adduced in support of the register hypothesis. Of the three, 'creaky' is the highest; this is exactly what one would expect of a creaky register. However, for most speakers, 'heavy' is higher than 'even'; though it may **end** lower. This is not what one would expect **a priori**; the breathy register might be expected to have lower pitch. For most speakers, it ends lower at least; I have found a few speakers, such as one lady from Mandalay, whose 'heavy' **was** in fact lower in pitch even at the beginning than her 'even' - in accord with the predictions concerning pitch if register is contrastive.

There are often segmental differences associated with suprasegmentals, such as a syllable-final glottal-stop. In Burmese, the 'creaky' sometimes in isolation ends in a weakly-articulated glottal-stop, while keeping its longer duration and different vowel quality from the 'killed' syllable type. This kind of segmental manifestation of a prosody is however more widespread in languages other than Burmese.

The obvious manifestation of register, apart from all the above parameters reflecting it, is in phonation or voice quality. I have measured this in two ways: indirectly in the acoustic output, and directly by measuring resistance across the glottis. Narrow band spectrograms show the phonation difference in an obvious but difficult to quantify mode: irregular, striated harmonics for 'creaky', solid harmonics for 'even', and blurry harmonics for 'heavy' types; the difference is less obvious in wide band spectrograms.

The only really direct observation of phonation is by laryngoscopy, either via a mirror or with fiberoptics and high-speed photography. Since this was unavailable, I instead used a Frokjaer-Jensen electroglottograph to measure resistance across the vocal cords, and hence degree and area of contact between the vocal cords. The output as displayed with a mingograph showed distinctly different characteristics for the three suprasegmentals in the speech of the one male informant I convinced to chance the electrodes; though again the differences are difficult to quantify. In a forthcoming article I will provide details of this and other methods such as photoglottography and FFT analysis to describe the register phenomena of Burmese.

In summary, orthography, vowel quality, and intensity support the analysis of Burmese as having register. Duration, contour, pitch, and segmental factors partly support the analysis, or at least provide no contradictions that would instead require a tonal analysis. Phonation factors themselves are also present, though more difficult than some other parameters to measure and quantify.

3. BURMESE AND AUSTROASIATIC

In lower Burma, there are millions of Mon speakers; there are many more who are aware of their Mon family background, but speak little or no Mon. Further, there are certainly very many Burmans who are descendants of Mon speakers but are unaware of it. These facts are a consequence of the gradual conquest (or rather repeated conquests) of the Mons by a series of Burmese kingdoms. Because of its location in the area of several former Mon kingdoms, Burma's current capital Rangoon is probably populated to a large extent by ex-Mons.

Therefore, it would hardly be surprising if there were some effect on the development of the Burmese language from this pervasive and long-lasting contact; especially since such important areas of Burmese culture as Theravada Buddhism and the writing system were directly received from the Mons.

In a short article (Bradley (1980)) I have shown several areas in which Burmese phonology appears to have diverged from a typical TB pattern in the direction of a more AA-like pattern. These include the very fundamental shift from a tone to a register system discussed above; vowel system developments to an eight vowel system lacking only /ɨ/ from a typical AA vowel system; adjustment of TB phonotactic patterns towards AA ones, such as the presence, in the orthography at least, of final palatals -c and -ɲ; and the development by non-final syllable reduction and other processes of non-monosyllabic words with a 'minor syllable' in non-final position. All are rather basic changes, and in many cases have no parallels within TB.

There are also less basic and/or less unparalleled changes, such as the diphthongisation of various stop- and nasal-final *rhymes and the addition of [aɩ] to the nasalised and stop-final rhyme inventory in Mon and other loans. Burmese has merged Proto-Bl affricates and fricatives reconstructed as *ts, *tš; *dz, *dž; *s, *š and so on; this opposition is typical of TB languages, but usually absent in non-northern AA languages. Subsequent changes in Burmese have led to the re-development of alveopalatal affricates and fricatives, however. Many other phonological changes could be cited. In addition to these fundamental phonological changes, there has also been substantial lexical borrowing; first from Mon into Burmese, and lately from Burmese into Mon.

Is it unreasonable to connect these instances of structural convergence with the historical facts noted above? I think not.

The Burmese suprasegmental system involves a three-way distinction whose primary parameter is phonation, with a variety of secondary parameters synchronically relatable to the phonation differences.

Historically it seems more likely that pitch was the primary parameter at an earlier stage (though Mazaudon (1977) and Weidert (1979) have argued otherwise). Thus it seems that an earlier tone system has become a register system in Burmese; that is in striking contrast to the developments in the tone systems of very closely related Burmish languages such as Atsi and Maru, which have remained 'tonal'. It is even more different from the further developments in several Loloish BL languages such as Lisu and Lahu, in which tones have proliferated (Bradley 1977, 1979).

In Burmese, the other parameters involved in the realisations of the registers include fundamental frequency; this may represent a persistence of the former primary parameter of Tones *A and *B. It would be very interesting to conduct a synthetic speech experiment with speakers of Burmese, to compare the relative weight of the parameters involved in the suprasegmental system. The danger of such a procedure would be the assumption that register, which is represented by a variety of factors, is decomposable into those separate factors rather than being an overall phonatory-articulatory set. Gregerson has suggested the position of the tongue root as the primary characteristic of this phonatory-articulatory set, while others concentrate on the action of the vocal cords. Again, the articulatory questions involved could be observed with studies of muscle activity using electromyography, or directly in terms of laryngoscopic or other articulatory observations, or indirectly by other methods.

It is to be noted that the Burmese register system makes a **three** way contrast between breathy, 'normal', and creaky phonation; so it is dissimilar to the frequent AA **two** way contrast within any one language. If one looks at AA languages more generally, one finds all three phonation types; though typically comparative evidence shows that contrastive use of register is a secondary phenomenon. Thus, Burmese has not developed an AA-like two-term register system; rather, it has become typologically more AA-like without completely converging.

Mon, which has a breathy versus 'normal' phonation opposition - with of course many other parameters involved, as ably described by Shorto and others - was the AA language in the most intimate contact with Burmese. The Pre-Burmese Tone *3 probably had creaky phonation; it developed in Proto-Tb etyma with *s or *? prefix or suffix (Bradley 1971). Thus, both the internal factor of creaky phonation in one tone and the contact factor of Mons, with a breathy versus 'normal' phonation opposition, becoming speakers of Burmese would have favoured the development of register in Burmese.

The conclusion from these facts about Burmese is that one should not assume phonological similarity between genetically related languages. Rather, one should describe all the phenomena in a language, and see how the system of that language functions. If it differs from the systems of closely related languages, as Burmese does, one possible explanation is areal convergence between unrelated languages in contact, such as Burmese and Mon.

TB and ST linguists should not presuppose that the prosodic systems of all TB or ST languages will be tonal; nor should AA linguists presuppose that AA languages will tend to develop register systems. These are general tendencies within the respective families, and may reflect characteristics of the prosidic systems of the respective proto-languages; but they have exceptions. For example, Vietnamese has developed a fairly complex tonal system, along lines more typical of ST languages. However, these developments do not mean that Vietnamese is not genetically related to the rest of AA; they simply make it typologically less AA-like in terms of its prosodies. Likewise, though Burmese has acquired some AA-like characteristics, it remains genetically TB and ST.

N O T E S

1. I am glad to acknowledge very useful comments by Gregerson and by Benedict on a preliminary version of this paper, given at the Austro-Asiatic Symposium, Helsingør 1979; and later comments from Matisoff, Haas, Diffloth, Thurgood, and Okell - which is not to blame them for any errors I have made.

BIBLIOGRAPHY

BENEDICT, P.K.

1972 *Sino-Tibetan: a conspectus*. Cambridge: Cambridge University Press.

BRADLEY, D.

1969 Akha dialects: synchronic and diachronic. AB thesis, Columbia University, New York.

1971 Prefixes and suffixes in Burmese-Lolo. Paper presented at the Fourth International Sino-Tibetan Conference, Bloomington, Indiana.

1975 Nahsi and Proto-Burmese-Lolo. *LTBA* 1/2:93-150.

1977 Proto-Loloish tones. *PL*, A-49:1-23.

1979 *Proto-Loloish*. SIAS Monograph Series No.39. London/Malmo: Curzon Press.

1980 Phonological convergence between languages in contact: Mon-Khmer structural borrowing in Burmese. *Berkeley Linguistics Society* 6:259-267.

EGEROD, S.

1971 Phonation types in Chinese and South East Asian languages. *Acta Linguistica Hafniensa* 13:159-171. Copenhagen.

GLOVER, W.W.

1971 Register in Tibeto-Burman languages of Nepal: a comparison with Mon-Khmer. *PL*, A-29:1-22.

GREGERSON, K.

1976 Tongue root and register in Mon-Khmer. In Jenner et al. 1976, vol.1: 323-369.

HENDERSON, E.J.A.

1952 The main features of Cambodian pronunciation. *BSOAS* 14:149-174.

HYMAN, L.M. ed.

1973 *Consonant types and tone*. Southern California Occasional Papers in Linguistics No.1.

JENNER, P.N., L.M. THOMPSON and S. STAROSTA, eds

1976 *Austro-Asiatic Studies*. *Oceanic Linguistics* Special Publication No.13.

MATISOFF, J.A.

1970 Glottal dissimilation and the Lahu high rising tone: a tonogenetic case study. *JAOS* 90/1:13-44.

1972 *The Loloish tonal split revisited*. Center for South and Southeast Asian Studies, Berkeley. Monograph No.7.

MATISOFF, J.A.

1973 Tonogenesis in Southeast Asia. In Hyman, 1973:75-95.

MAZAUDON, M.

1977 Tibeto-Burman tonogenetics. *LTBA* 3/2.

SPRIGG, R.K.

1964 Burmese orthography and the tonal classification of Burmese lexical items. *Journal of the Burma Research Society* 48/2:415-444. Rangoon.

THEIN TUN

1982 Some acoustic properties of tones in Burmese. In this volume.

THURGOOD, G.

1980 Consonants, phonation types, and pitch height. *CAAAL* 13:207-219.

WEIDERT, A.

1979 The Sino-Tibetan tonogenetic laryngeal reconstruction theory. *LTBA* 5/1:49-127.

Bradley, D. "Register in Burmese". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation.*
A-62:117-132. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.117*

A NEW HIGH TONE IN SOUTHERN THAI

Anthony Diller

0. INTRODUCTION

For some speakers of Southern Thai, lenition of final velar and glottal stops is beginning to distinguish a new class of open-syllable high-tone items. This 'new tone' arises in problematic circumstances which are discussed. Reference is made to more general contemporary theories of tonal development.

1. KIHT: TONOGENESIS-IN-PROGRESS

The Thai geographical metaphor for 'isthmus' is khɔɔ khɔ̂ɔt *'constricted neck'*. For some Southern Thai speakers in the Kra Isthmus region *'constriction'* is also more literal: it is an articulatory feature associated with the differentiation of a new phonological tone. This 'Kra-Isthmus High Tone', henceforth KIHT, is a rather limited case of tone development in progress and its evolution is along lines now familiar (cf. Haudricourt 1954; Jones 1960; Matisoff 1970); yet there are additional structural and sociolinguistic factors in the precarious survival of the tone which merit examination. Below we review the phonological setting for KIHT's appearance, its sociolinguistic 'protagonist', and some questions raised for general issues in tonal development.[1]

Evidence for neolithic settlement in the Kra region is firm. For the first millennium A.D. Chinese records suggest several Indianised trading outposts, and these are confirmed by local archaeological evidence at Takuapa, Chaiya and Kanchanadit (Wheatley 1961). When the Kra region became substantially Tai-speaking is a matter of conjecture. Local historical legends (Wyatt 1975) and other testimony make it plausible that Tai varieties have been spoken in the area for some six or seven hundred years. Apart from a brief Burmese invasion in 1785

Diller, A. "A New High Tone in Southern Thai". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation*.
A-62:133-154. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.133*

TABLE 1

TAI TONE CATEGORIES

	open, unmarked	open, ˀèek	open, thoo	closed, long	closed, short	
High class	A1	B1	C1	D1-L	D1-S	yin
Mid class	A2	B2	C2	D2-L	D2-S	yin
Low class	A3	B3	C3	D3-L	D3-S	yang
	ping	qu	shang	ru	ru	

historical sources mention little population discruption; there also has been less contact with non-Tai speakers than would be the case further south.

2. KRA ISTHMUS TONAL SYSTEMS

The Southern Thai dialect group extends from somewhat north of the Kra Isthmus proper (perhaps from Bang Saphan Yai, 11°17'N, where isoglosses converge) to south of 7°N. Tone systems in the Kra area have been described in surveys by Jones (1965) and Brown (1965). These surveys were concerned with wider issues of comparative Tai reconstruction, and it is not strange that they gave no specific attention to the rather parochial facts discussed here. KIHT occurs on items in a limited correspondence category which is usually of little more than marginal interest in comparative Tai, and furthermore occurs mainly among uneducated rural speakers.

'Tai' here refers to the large language family of which varieties of Central Thai (or Siamese), Southern Thai, etc., are members. It is convenient to discuss tones by referring to Tai tonal categories as labelled by F.-K. Li (1977), slightly modified as shown in Table 1. Etymological correspondence is reflected moderately well by spelling conventions in the traditional Thai orthography; these are shown above and to the left in Table 1. Unaspirated stops are confined to the Mid orthographic class; other initial consonants are presently represented by letters in the High and Low classes. There is also enough regularity in relationships with traditional Chinese etymological classes to postulate normal borrowing patterns for early-strate Chinese loans into Proto-Tai; these are shown below and to the right in Table 1.

Table 2 summarises Southern Thai tonal systems common in the Kra Isthmus region. Local varieties are shown in a north-to-south sequence from Chumphon to Phanom, the latter an isolated inland settlement on an old trans-isthmus route. The nearby Kanchanadit dialect on Bandon

TABLE 2

KRA ISTHMUS TONAL SYSTEMS

Tai tone category	D1-L	C1	C2	D2-L	B3	D3-L	C3	A3	A2	B2	B1	A1
1. Chumphon	/ 55 /	33	/		24		/ 21 /	232	/		52	/
2. Ranong	/ " /	"	/		"		/ " /	"		/	"	/
3. Langsuan	/ " /	"	/		"		/ " /	"		/	"	/
4. Tha Chang	/ 54 /	"	/		"		/ " /	232	(343)	/	"	/
5. Phanom	/ 44	/	33 /		"		/ " /	" /	343	/	452	/
6. Kanchanadit	/ "	/			"		/ " /	32 /	43	/	"	/

Bay is cited for comparison. Pitch and contour are indicated by the low-to-high 1-to-5 numbering system (see Bradley 1977:1 for more detail). For the closed D categories, short- and long-vowel items have rather similar characteristics, particularly in the High and Mid orthographic classes, and only long-vowel items are shown in Table 2.

TABLE 3

URBAN CHUMPHON SOUTHERN THAI EXAMPLES

	Tai tone category		tonal pitch and contour
1. *'leg'*	A1	khaa	high falling
2. *'crow'*	A2	kaa	mid-low rising-falling
3. *'thatch grass'*	A3	khaa	mid-low rising-falling
4. *'sp. rhizome'*	B1	khaa	high falling (=1)
5. *'jungle'*	B2	paa	high falling (=1)
6. *'value'*	B3	khaa	low rising
7. *'to kill'*	C1	khaa	mid level
8. *'aunt'*	C2	paa	mid level
9. *'to do trading'*	C3	khaa	low slightly falling
10. *'to lack'*	D1-L	khaat	high level
11. *'mouth'*	D2-L	paak	low rising
12. *'land leech'*	D3-L	thaak	low rising
13. *'to polish'*	D1-S	khat	high slightly rising
14. *'to bite'*	D2-S	kat	mid slightly rising
15. *'to think'*	D3-S	khit	low slightly rising

Table 3 illustrates the preceding issues for the case of Chumphon with data from an urban speaker. For comparison with Central Thai and also with a variety south of the Kra Isthmus, see Diller 1979:61-65.

As Tables 2 and 3 indicate, Kra Isthmus systems merge Tai tone categories into six or seven discrete tonal shapes. Several qualifications are necessary. Table 2 is based on citation-form pronunciation and questions of tonal sandhi are not considered. In fact in pretonic environments there is considerable levelling of contour in most varieties (see Thongkum 1978:27,47 for details). Also, Table 2 should be interpreted as a 'digitalisation' of what is really an areal continuum. Tonal shapes gradually shift to produce the systems reported, and there is a certain arbitrariness in deciding exactly where to report two discrete tonal shapes rather than one. This is particularly a problem in the case of the category A2 merging with A3. At the extreme north of the continuum the two are clearly identical; at the south, clearly separate. In the intermediate area a given speaker may show variation, sometimes - or for some items - making a distinction, but elsewhere not. Thongkum (1978:8) has recognised the difficulty for the Suratthani varieties she reviews and reports that speakers have a 'feeling' of separateness even when only slight register differences are discernible in sound spectrograms. It must be kept in mind that the consonant initials in A2 and A3 are in complementary distribution. In structuralist terms then we are confronted with an areally distributed separation of tonal allophones by a register distinction.

Another question of tonal allophone differentiation is directly relevant to the origin of KIHT. This concerns the tonal status of long-vowel closed-syllable items, particularly those with initial consonants associated with the High or -l class, such as item 10 in Table 3.

In virtually all Tai varieties any long-vowel item terminating in -p, -t, -k will coincide tonally with other non-stop-final items, that is, D will merge suprasegmentally with either A, B or C, with only the final consonant, and not tonal features, distinguishing a given D item from similar open-syllable ones. A common pattern is for D to merge with B. This is the case for Central Thai, Northern Thai, Shan, Khamti and for most varieties of Zhuang. Mergers of D with A are unusual, e.g. White Tai D3-L = A3. D and C merge in most Lao varieties, although D2-L usually joins D1-L in merging with C1. Southern Thai south of the Kra Isthmus area under study here shows a mixed system, with D1-L and D2-L merging with respective C categories, while D3-L merges with B. The situation is similar to, but rather more tidy than, the shifting about of ru-class items in Chinese.

Now the Kra Isthmus systems show a rare exception to this tendency to merge. D2-L and D3-L indeed join tonally with B3. D1-L however resists the tendency, especially as one moves northward up the isthmus.

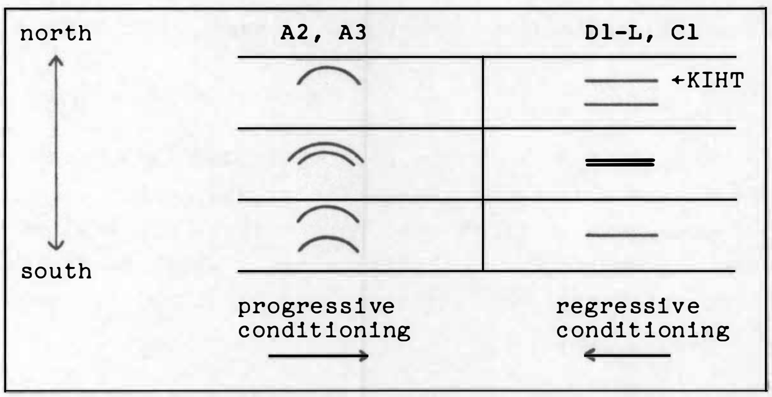
Starting from just north of Phanom - Kanchanadit, where D1-L = C1, the closed syllables (D1-L) begin to select a slightly higher allophone, the open ones (C1) a slightly lower one. At first the separation may be sporadic and in flux, but as one passes north of Chaiya to Tha Chang the difference is clearly systematic and becomes more pronounced farther north.

From a rather strict structuralist point of view this tonal separation still is of an allophonic nature as long as there remain final stops -p, -t, -k to stand in complimentary distribution with other finals in the C1 class. That is, we would still need to treat D1-L as a conditioned higher-register 'variant' of C1, with parallelism of contour - in this case level - plausibly constituting the basis for grouping. We simply recognise higher and lower allophones of the same tone, with stop-final items selecting the former one.

At this point we can remember that a similar allophonic differentiation process is occurring in the area with respect to A2 and A3, as mentioned above. Interestingly, the areal distributions are exactly opposite; this is shown schematically in Table 4.

TABLE 4

AREAL PATTERNING OF KRA ISTHMUS REGISTER SEPARATIONS



It is tempting to see more than a chance relationship in Table 4. Could the systems in the Kra Isthmus be 'seeking' a six-tone equilibrium, with register differentiation controlled by more general pressures from systematic phonology? If this is so, then it is of interest that the register differentiation called for by such a programme is achieved through two very different mechanisms. On the one hand, the A2/A3 separation is a matter of **progressive** conditioning on the basis of **initial** unaspirated stops (vs. aspirated stops, nasals and sonorants); on the other hand, the D1-L/C1 separation is a matter of **regressive** conditioning on the basis of **final** stops (vs. final nasals, semi-vowels, or Ø). Both of these processes appear to have the same result: register differentiation without any significant contour shift. For the time being the interrelationship of these register differences and exactly how they pertain to diachronic and areal factors must remain speculative, although the basic synchronic facts are not in question.

3. THE ORIGIN OF KIHT

Now we are in a position to see how KIHT arises. As noted in the preceding section, as long as D1-L is confined to stop-final items and C1 to others, the parallelism of contour along with complementarity of initials indicates that /D1-L + C1/ is a 'single tone' from a lexical or abstract phonological point of view. Thus an item with this lexical tone would receive surface pitch-contour features through a rule such as:

RULE 1 $\text{tone}_{\chi} \rightarrow \begin{bmatrix}\alpha\ \text{high} \\ -\ \text{contour}\end{bmatrix} \ / \ \_\_\_ \ [-\alpha\ \text{continuant}]$

As it happens, rural varieties of Southern Thai in the Kra Isthmus area are subject to lenition of -k after long vowels - a change found in other Tai varieties and discussed below in 4. The change occurs on the western coast of the isthmus to below the island of Phuket, where it is frequently heard in urban speech as well. The shift is diachronic:

RULE 2 *k → ʔ / [+ long] ___ #

The net result of Rules 1 and 2 is to render more tenuous the status of /D1-L + C1/ as a 'single' tone. The grosser articulatory organs of lips and tongue have given over to more subtle laryngeal components the task of keeping apart the two syllabic types which condition Rule 1. In fact a further stage of lenition would be quite plausible.

RULE 3 *ʔ → Ø / [+ long] ___ #

Where Rules 2 and 3 have applied in sequence a phonological reassessment of Rule 1 is necessary. It is no longer possible to condition allophonic assignment of tone on the basis of final continuant, since -Ø, i.e. forms terminating in long vowels without final consonants, now occur in both pitch registers covered by Rule 1.

Such a change is presently underway for rural speakers in various Kra Isthmus communities. Rural speakers in Langsuan District of Chumphon Province were found to be particularly thorough-going in making the change.[2] For them the new tone KIHT must be recognised as one with independent phonological status. Contrasts like the following occur:

ITEM		TONAL PITCH-CONTOUR	TONE CATEGORY
'mist'	mɔɔ	high level	KIHT, i.e. D1-L with application of Rules 2 and 3 (< *mɔɔk)
'doctor'	mɔɔ	high falling	A1
'pot'	mɔɔ	mid level	C1
'full'	mɔɔ	mid-low rising-falling	A3
'father' }	phɔɔ	low rising	B3
'stomach' }	phɔɔ	low rising	D3-L with application of Rule 3 (< *phɔɔʔ)
'sp. small palm'	phɔɔ	low slightly falling	C3

Clearly, for these speakers there can be no question of phonological reduction; they operate with a six-tone system. It must quickly be added however that KIHT, the high level tone, is presently very restricted as to permissible finals: it requires -p, -t, or -Ø. It is perhaps a 'minor' tone.

The preceding list reveals another diachronic process relevant to the origin of KIHT. The fact that *'father'* and *'stomach'* have merged on the low rising tone shows another potential source for KIHT items. A slight excursus into tone-conditioned vowel lengthening is necessary before taking up this matter.

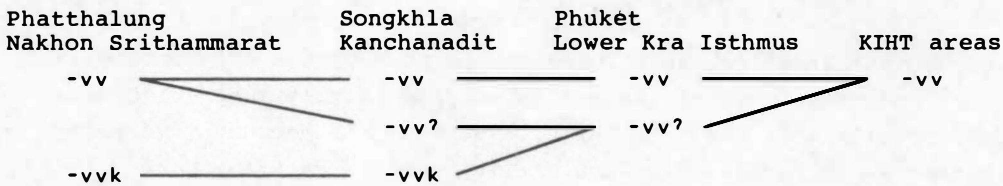
Egerod (1961), Hartmann (1976:147) and Gandour (1977:58) have called attention to Southern Thai lengthening of *-aw, *-ay in C tones, and occasionally in B. This lengthening is probably of long standing, since Wyatt (1975:15) has described an early Southern Thai spelling convention that seems to have represented it. Gandour (1977) has called attention to a Phuket variety in which tone-conditioned length distinctions apply for all non-low vowels; again C along with B3 is associated with long vowels.[3]

Southern Thai south of the Kra Isthmus differs from most other Tai varieties in allowing two types of glottal-stop-final syllables; e.g. for Ranot in Songkhla Province:

(emphatic particle)	naʔ	high rising
'to cough up'	naaʔ	high level
(derogatory particle)	waʔ	high rising
'to divide proceeds'	waaʔ	high level
(connective adverb)	kaʔ	mid rising
'to estimate'	kaaʔ	mid level

The sequence short vowel plus glottal stop is mainly restricted to particles and adverbs; long vowel plus glottal stop, to nouns and verbs. Unlike Kra Isthmus items which owe their glottal stops (and where KIHT occurs, the proximate source of their open finals) to Rule 2, the long-vowel glottal-stop items above show no evidence of once having ended in -k. On the contrary, many of these items have cognates in Central Thai and other dialects in short vowel plus glottal stop. It appears likely that a pre-glottal-stop lengthening rule has applied in these cases. For Songkhla varieties, this results in -vv, -vvʔ, and -vvk as possible syllable types (with vv representing a long vowel).

As one moves north of Songkhla this liberality in the final component of the syllable is not tolerated and mergers in syllable type involving Rules 2 and 3 occur as follows:



We are now in a position to understand, e.g. how *'father'* and *'stomach'* have fallen together in Langsuan as phɔɔ$_{24}$: the former is an unmodified item in -vv; the latter is cognate with Central Tai kraphɔʔ$_{5}$ and with Songkhla phɔɔʔ$_{23}$ (showing typical Southern Thai prefixal deletion and pre-glottal-stop lengthening) and Rule 3 has applied to give -vvʔ → -vv.[4]

When the process above occurs with High-class or -1 tone-category items, then for the relevant speakers they fall into KIHT -vv items along with those originally in -vvk. Thus there are two etymological sources for KIHT open syllables.

Table 5 illustrates in an areal way the diachronic processes involved in KIHT's origin. Kanchanadit forms are cited for their presumed conservatism. Tha Chang represents the Lower-Kra-Isthmus varieties which have undergone Rule 2 but not Rule 3. Finally, the rural Langsuan variety cited has undergone full lenition (save after diphthongs) and the items in Table 5 are evidence for recognising KIHT as a phonologically independent tone. Items 1-40 are derived from *-vvk; the remaining ones, from glottal-stop finals, although in some cases there are no apparent Central Thai cognates.

Not shown in Table 5 is possible laryngeal constriction which can mark transition from -vvʔ to -vv. For items derived from -k and -ʔ a few speakers in the KIHT area produce such constriction regularly.

TABLE 5

LENITION OF KRA ISTHMUS HIGH-TONE FINALS

	tone	Kanchanadit	Tha Chang	Langsuan
1. *'dried up'*	D1-L	phaak	phaaʔ	phaa
2. *'to tie'*	"	phuuk	phooʔ	phoo
3. *'albino'*	"	phɯək	phɯəʔ	phɯəʔ
4. *'to hew'*	"	thaak	thaaʔ	thaa
5. *'to shove'*	"	thɛɛk	thɛɛʔ	thɛɛ
6. *'to be afflicted'*	"	thuuk	thooʔ	thoo
7. *'to pull back'*	"	thɔɔk	thɔɔʔ	thɔɔ
8. *'screen'*	"	chaak	chaaʔ	chaa
9. *'to tear'*	"	cheek	cheeʔ	chee
10. *'to cough up'*	"	khaak	khaaʔ	khaa
11. *'to rap'*	"	kheek	kheeʔ	khee
12. *'guest; Malay'*	"	khɛɛk	khɛɛʔ	khɛɛ
13. *'small unit'*	"	khook	khooʔ	x
14. *'to entrust'*	"	khwaak	khwaaʔ	faa
15. *'elephant grass'*	"	khwɛɛk	khwɛɛʔ	fɛɛ
16. *'splint'*	"	khwɯək	khwɯəʔ	fɯəʔ
17. *'teak; pestle'*	"	saak	saaʔ	saa
18. *'to pour holy water'*	"	seek	seeʔ	see
19. *'tp. suspended basket'*	"	sɛɛk	sɛɛʔ	sɛɛ
20. *'sadness'*	"	sook	sooʔ	soo
21. *'lower arm'*	"	sɔɔk	sɔɔʔ	sɔɔ
22. *'separate'*	"	haak	haaʔ	haa
23. *'to break apart'*	"	hɛɛk	hɛɛʔ	hɛɛ
24. *'grey hair'*	"	hɔɔk	ŋɔɔʔ	ŋɔɔ
25. *'spear'*	"	hɔɔk	hɔɔʔ	hɔɔ
26. *'loom'*	"	huuk	hooʔ	huu
27. *'gill'*	"	hɯək	ŋɯəʔ	ŋɯəʔ
28. *'areca'*	"	maak	maaʔ	maa
29. *'mist'*	"	mɔɔk	mɔɔʔ	mɔɔ
30. *'nose'*	"	muuk	muuʔ	muu
31. *'hat'*	"	muək	muəʔ	muəʔ
32. *'ox hump'*	"	nook	nooʔ	noo
33. *'strange'*	"	laak	x	laa
34. *'to avoid'*	"	leek	leeʔ	lee
35. *'to frighten'*	"	lɔɔk	lɔɔʔ	lɔɔ
36. *'to raise the eyes'*	"	lɯək	lɯəʔ	lɯəʔ
37. *'sugar-palm beer'*	"	waak	waaʔ	waa

Table 5 cont'd

	tone	Kanchanadit	Tha Chang	Langsuan
38. *'to push apart'*	D1-L	wɛɛk	wɛɛʔ	wɛɛ
39. *'to want'*	"	yaak	yaaʔ	yaa
40. *'to tease'*	"	yɔɔk	yɔɔʔ	yɔɔ
41. *'unit of thatch'*	"	phlaaʔ	x	phlaa
42. *'rotten'*	"	phuuʔ	phooʔ	phoo
43. *'wet'*	"	chɛɛʔ	chɛɛʔ	chɛɛ
44. *'to take advantage'*	"	khɛɛʔ	khɛɛʔ	khɛɛ
45. *'mixed up'*	"	khləəʔ	khləəʔ	khləə
46. *'indistinct'*	"	khruəʔ	khruəʔ	khruəʔ
47. *'pond'*	"	saaʔ	saaʔ	saa
48. *'tp. fishtrap; nearly ripe'*	"	sooʔ	sooʔ	soo
49. *'backwash'*	"	sɔɔʔ	sɔɔʔ	sɔɔ
50. *'falling apart'*	"	hooʔ	x	hoo
51. *'to carve'*	"	lɛɛʔ	mlɛɛʔ	mlɛɛ
52. *'ugly'*	"	mrooʔ	mrooʔ	mloo
53. *'tp. skin disease'*	"	laaʔ	laaʔ	laa
54. *'to lure with light'*	"	looʔ	looʔ	loo
55. *'tp. wood worm'*	"	raaʔ	raaʔ	raa
56. *'tightening stick'*	"	rɛɛʔ	rɛɛʔ	rɛɛ
57. *'rubbish'*	"	yaaʔ	yaaʔ	yaa

For such speakers items such as phɔɔ *'father'* (B3) are rather tenuously distinguished from others such as phɔɔ [+constriction] *'stomach'* (< *D3, coinciding with B3). The constriction feature would also apply to relevant KIHT items. The feature appears more frequently as a sporadic item-by-item adjunct and in any event tends to be lost in connected speech.

Also, a remark on the change -vvk → -vvʔ of Rule 2 is in order. For Southern Thai varieties preserving finals -p, -t, -k, it is often possible to detect along with final oral stop terminus a coarticulated glottal stop. One might symbolise this by -ʔp, -ʔt, -ʔk. This suggests a plausible articulatory clarification of Rule 2: what occurs is perhaps more a gradual weakening of dorso-velar contact rather than an abrupt 'jump' of point of closure.

4. RETENTION AND LOSS OF FINAL STOPS IN TAI

The history proposed for the origin of KIHT is a matter of rather general phonetic processes which have applied at one time or another in numerous languages in the area. We might expect that the case at hand in Southern Thai is but one of many among languages of the Tai family. This question is examined briefly in the present section.

The general Tai pattern is one of conservatism with respect to final stops -p, -t, -k. Only on the extreme northern fringe of the Tai area in Guizhou Province has the Bu-i dialect of Shuicheng reportedly lost final stops completely. In other Bu-i dialects loss is not complete and one finds various stages of attrition related to syllabic conditions, particularly to vowel length and height. These changes have been discussed in some detail by Li (1971, 1977:53-55), Dell (1969) and Sarawit (1975).[5]

White Tai and so-called 'Phu-thai' varieties, including many of Black Tai, regularly retain -p and -t, but there is loss of -k after vowels which were once (at least) long. Homophony has not been avoided, as we see in these examples from Donaldson and Dieu (1970):

'field'	nã	(< *naa (A3))
'otter'	nã	(< *naak (D3-L))
'heavy'	năk	(< *ŋ̊ak (D1-S))

and similarly tã *'to paint'* (A3), *'land leech'* (< D3-L), but tăk *'to greet'* (D3-S). This White Tai situation is aberrant in two respects: it involves the merger, unusual as noted above, of a D-L category with an A one, and also the resulting 'shared' tone is glottalised, quite abnormal and diachronically puzzling for an A tone. (Tai comparativists would rather find glottalisation on C or possibly on B tones; Lao provides good examples.) Perhaps a laryngeal contagion - the traces of -k lenition - has spread as the tone D3-L merged with A3.

Cases of potential and real homophony like those above set the White Tai situation off from that of KIHT. Clearly in White Tai there has been suprasegmental **merger** and then final -k lenition after long vowels; at the Kra Isthmus there has been suprasegmental **split** with subsequent lenition. In the first case the result is by no means a 'new tone' but rather total merger and homophony; in the latter case prior register separation assures new open-final items an 'independent' status.

On the other hand, the Bu-i dialect of Tushan described by Li (1977: 55) seems to have acquired a 'new' low falling-rising tone through a set of changes very much like what happened to produce KIHT. In Tushan both D1-L and D1-S are combined in the new tone (with other categories of formerly closed syllables merging with open categories). The forms rɔɔ *'six'*, paa *'mouth'*, and ʔɔɔ *'to go out'* are examples. The latter two would be D2-L in the terminology used here; in KIHT dialects they would merge suprasegmentally with B3. As with KIHT, the Tushan D1-derived low falling-rising tone is 'new' in the sense that it accomodates open syllables but cannot be accounted for by the normal Tai open-syllable categories A, B, or C.

Before leaving the question of final stops and their lenition, we should call attention to a reverse case. In some Southern Thai communities well south of the Kra Isthmus the final glottal stop is replaced by -k:

RULE 4 $\text{ʔ} \rightarrow \text{k} \;/\; \begin{bmatrix} +\text{ long} \\ -\text{ high} \end{bmatrix}$ ___

The following forms from Ban Nam Phrai in Trang Province illustrate the phenomenon. Songkhla forms, which are conservative on comparative grounds, are cited to the left.

	Songkhla	Ban Nam Phrai
'to divide'	waaʔ$_{44}$	waak$_{44}$ }
'sugar-palm beer'	waak$_{44}$	waak$_{44}$ }
'falling apart'	hooʔ$_{44}$	hook$_{44}$
'to lure with light'	looʔ$_{44}$	look$_{44}$
'table'	tooʔ$_{33}$	took$_{33}$

Clearly we must guard against proposing hard-and-fast Tai-wide rules of the sort which have produced KIHT and the Tushan low falling-rising tone. Although Rule 4 has apparently not been reported elsewhere among Tai speakers, when more local rural data become available it may well appear.

5. SOCIOLINGUISTIC CONFRONTATION

KIHT's survival as a phonologically discrete tone is presently rather problematic. It is threatened by sociolinguistic pressures described in this section.

To see why KIHT may soon be a matter of past history, some consideration of socially constrained dialect mixing and switching among Tai varieties is in order (for a more detailed account, see Diller 1979).

Prince Damrong, pre-eminent among Thai scholars in the early decades of the present century, wrote the following in the essay 'An Explanation of Tones':

> These days there are government schools being established and spreading more and more. Wherever the schools reach, the Bangkok sound system is transmitted. New generations of citizens will increasingly use the Bangkok sound system, and former systems in local use will probably deteriorate in stages. In another fifty years or so the tone marks in the writing system will probably be realised in the same way by everyone ....[6]

By now compulsory Central Thai education is nearly universal in Thailand, and the mass media and increased areal mobility among the rural working class have further strengthened the position of the standard language. However the 'stages of deterioration' of local varieties predicted by Prince Damrong have proved to be somewhat longer-term and more complex than was supposed. KIHT's fate is wrapped up in these issues.

The dynamics of Central Thai spreading is somewhat different among the various non-Central communities in Thailand. The North for example appears more recalcitrant than the South in admitting Central Thai replacements, a condition perhaps connected to somewhat different political histories.

Of particular importance are intermediate stages of Central-local **hybridisation**, that is, of Central influence on local varieties that stops short of complete supplanting. One particularly common and systematic form of such dialect mixing consists in **preservation of a local tonal system but with segmental and lexico-semantic replacements from Central Thai.** There may be a graduated scale of such replacements, depending on social backgrounds of speakers or on contextual factors relating to speech acts or conditions of social interaction. Shifts up the scale toward total Central segmental and lexico-semantic replacement may signal decreasing intimacy and increasing formality and officiousness. Eventually Central replacements may become more-or-less 'standard' in the local variety and the more authentic local segmentals and lexemes may become archaic and disappear, all without any necessary modification in the tonal system. The final result would be two varieties differentiated entirely by tone, and probably used by bidialectal speakers mainly as a social resource for purposes of stylistic switching and related social manoeuvering. It might be that in speech communities where this had happened the social utility of such tonal switching might contribute to the preservation of a 'segmentally subjugated' local variety well beyond the time span envisaged by Prince Damrong.

In any event the process sketched above is underway in many regional urban areas and it reaches into village speech as well. A particularly relevant case concerns the replacement of -ʔ by -k in a Northeastern Phu-thai variety described by Khanitthanan (1977:44), which is undoubtedly influenced by surrounding Northeastern Thai (Lao) as well as by Central Thai.

	Central Thai	'Authentic' Phu-thai		Hybrid Phu-thai
'bark, peeling'	plɯək$_{21}$	pəʔ$_{31}$	→	pəək$_{31}$ ∿ pɯək$_{31}$
'to select'	lɯək$_{52}$	ləʔ	→	ləək$_{31}$ ∿ lɯək$_{31}$
'mouth'	paak$_{21}$	paʔ$_{31}$	→	paak$_{31}$
'offspring'	luuk$_{52}$	luʔ$_{31}$	→	luuk$_{31}$
'to expose'	taak$_{21}$	taʔ$_{31}$	→	taak$_{31}$
'wing'	piik$_{21}$	piʔ$_{31}$	→	piik$_{31}$

It is important to draw a distinction between this sociolinguistic hybridisation process and -ʔ → -k described for Ban Nam Phrai, Trang, in the preceding section. In the latter case Central Thai applies no pressure for (or indeed applies negative pressure against) the change.

The Chumphon variety cited in Table 3 has gone through the same replacement process as Phu-thai above in the case of items 11 and 12, paak$_{24}$ *'mouth'* and thaak$_{24}$ *'land leech'*. Rural Kra Isthmus varieties would show either -ʔ or, in KIHT areas, -Ø. Although not directly relevant to tonal questions, lexico-semantic differences can also be illustrated at this point to place the segmental substitutions in their wider systematic context.

	Central Thai	'Authentic' Kra Isthmus	Hybrid Kra Isthmus
'mouth'	paak$_{21}$	paaʔ$_{24}$ ∿ paa$_{24}$	→ paak$_{24}$
'land leech'	thaak$_{52}$	thaaʔ$_{24}$ ∿ thaa$_{24}$	→ thaak$_{24}$
'seal'	traa$_{33}$	kraa$_{232}$	→ traa$_{232}$
'muddy'	mɯək$_{52}$	mlɯəʔ$_{24}$	→ mɯək$_{24}$
'to ache'	mɯəy$_{52}$	mlɯəy$_{24}$ ∿ lɯəy$_{24}$	→ mɯəy$_{24}$
'to throw'	paa$_{33}$	liw$_{232}$	→ paa$_{232}$
'fast'	riip$_{52}$	khɛɛp$_{55}$	→ riip$_{24}$
'thick (of liquids)'	khon$_{52}$	khɛɛn$_{33}$	→ khon$_{33}$
'basin'	kalamaŋ$_{33}$	khoom$_{232}$	→ kalamaŋ$_{232}$

As for KIHT, two Central Thai pressures are set against it, corresponding to the two Central Thai syllable types:

(1) C(C)vvk

(2) C(C)vʔ

That is, for a given KIHT item (e.g. in the right-hand column of Table 5) the hybrid tendency is to add -k if there is a Central Thai cognate in (1), or to treat the KIHT item as D1-S with -ʔ if there is a cognate in (2). Homophony may be disambiguated in this way.[7]

	Central Thai	'Authentic' Kra Isthmus (KIHT areas)		Hybrid Kra Isthmus
'pestle'	$saak_{21}$	saa_{55}	→	$saak_{55}$
'pond'	saʔ	saa_{55}	→	$saʔ_{45}$
'mist'	$mɔɔk_{21}$	$mɔɔ_{55}$	→	$mɔɔk_{55}$
'suitable'	$mɔʔ_{21}$	$mɔɔ_{55}$	→	$mɔʔ_{45}$

Where there is no clear Central Thai cognate, lexical replacement tends to occur instead:

'ugly'	$naa_{52}kliat_{21}$	$mloo_{55}$	→	$naa_{33}kliat_{24}$

In this way the KIHT situation is reduced to complementary distribution with the tone C1 (33) as described in 2, and KIHT can no longer claim status as an independent new tone. It remains however as a register-differentiated allophone.

What we find in the case of KIHT then is a 'conflict of interest': in terms of purely internal-systematic factors, Rules 2 and 3 - lenition of final velar and glottal stops after long vowels - seem to present a phonologically reasonable simplification of the syllable-final component with resulting albeit rather marginal new-tone independence in the D1-L case; on the other hand, competing Rule 4, fostered by external factors along with the Central Thai syllabic target -vʔ, tends to undermine KIHT and return it to merely allophonic status. One expects that the indomitable superimposition of Central Thai predicted by Prince Damrong may do away with this fledgling tone in the near future.

6. TAI TONES AND TONOGENETIC THEORY

Several issues above have some relevance for wider concerns and theories in the development of tones (see Bradley 1977:1-4 for a concise summary of recent discussions). It is perhaps more useful to represent these issues somewhat negatively in the form of caveats to oversimplified theory construction, although of course many aspects of the Tai data discussed above fit comfortably into the wider framework now generally accepted.

1. It is not generally true for Tai that tones and segmentals rearrange themselves in such a way as to maximise distinctive phonological space and thus avoid homophony. Cases of homophony abound in Tai and the need to use tones and segments distinctively must be balanced against other opposite pressures, some perhaps from larger phonological 'conspiracies' or from other abstract diachronic trends. Tais are not afraid of manageable lexical ambiguity.

2. One might set up a 'standard theory of tonogenesis' (with admittedly straw-man characteristics):

> Tones develop in two ways, progressively (with conditioning from initial consonants) and regressively (linked to finals). Progressive tonal development is a matter of register separation caused by voicing, which lowers pitch, or other initial features such as preglottalisation or aspiration. Regressive tonal development is clearest in the case of final laryngeal components: final glottal constriction or glottal (and perhaps oral) stops cause rising contour; final -h (and perhaps -s, -x) cause falling contour.

Whatever merit this 'standard theory' may have as a point of theoretical departure, the Tai languages are difficult to force into such a diachronic schema. Brown (1975) and others have shown the complexity of apparent progressive tonal development in Tai, and although Southern Thai in general may provide some support for the progressive half of the 'standard theory', many other varieties in the Tai family do not. Note that the A2/A3 register difference discussed above for lower Kra Isthmus varieties (see Tables 2, 4) provides a counter-example: in this case voiced initial stops are not causing a lower register but, along with other Mid or -2 class items, are actually selecting a higher pitch level. Kra Isthmus data calls into question the 'standard theory's' regressive development programme as well. There is no significant contour difference in the final-stop-associated separation of D1-L and C1; instead contours are even and the difference is one of pitch register alone. Nor can we even conclude that final stops raise pitch: they do for D1-L but they do not for D2-L or for D3-L.[8]

3. Related to matters raised above is an unspoken assumption that final laryngeal constriction, final glottal stops and especially final oral stops are by nature conservative clues to the arising of tonal phenomena from an earlier tone-free state. Above we noted that White

Tai A3 items have now become laryngealised, perhaps through 'contagion' as formerly stopped items merged with them suprasegmentally, and in the case of -vvk items, completely. The Southern Thai examples of -vvʔ → -vvk show even more clearly the danger of taking an uncritical attitude toward laryngeal and oral finals.[9]

4. Sociolinguistic factors in tonal development may not always be negligible. Very few Tai varieties (Ahom being an example that comes to mind) have developed in relative isolation. For a millennium Tai groups have interacted among themselves and their neighbours in varying configurations of political alliance, domination and subjugation. Probably significant groups of speakers have been partially bidialectal, and tones and other phonological features have probably been caught up in wider patterns of socioeconomic absorption and integration. On the one hand Tai as a whole provides a near paradigmatic example of the 'Neogrammarian Hypothesis', with quite regular sound changes and a plausibly reconstructable proto-language; on the other hand, wherever one turns attention to lexical or phonological detail, inconsistencies and peculiarities arise which can only be elucidated by examining inter-dialectal processes and in some cases contact with non-Tai neighbours. Discussion of the suppression of KIHT open syllables by phonological targets in the locally prestigious Southern-Central hybrid has suggested one way sociolinguistic constraints on tonal development can operate.

5. Methodologically, the KIHT situation raises two important issues. One is the need to consider phonological systems in their entirety and to be aware of problems in segmental-suprasegmental distribution, particularly when complimentarity is involved. There have been recent attempts to systematise and compare 'tonal systems' across many languages with a view to arriving at structural generalities. But we may be comparing oranges with bananas if we are oblivious to how specific phonetic tonal contours function within phonologies as a whole. KIHT, for example, shows that whether or not to 'count' an extra new tone should be taken as a **problematic** issue, partly a matter of structural theoretical interpretation, partly a matter of fact in the field, and partly a recognition of conflicting psycho-articulatory pressures.

Finally, discussion of hybridised varieties above suggests a caution in organising field research in Tai dialectology and elsewhere. If one relies, for example, on university students as (very convenient) informants, one will almost certainly be up against an issue of hybridisation plus, perhaps, 'educated' stereotyping of 'uneducated' rural speech. Although data acquired in such a way is of undoubted value and may show structural consistency in its own way, urbanised and educated speakers of local Tai varieties are almost certain to be 'lames' in Labov's sense, and their hybrids may be concealing interesting differences in local rural speech. KIHT is one such difference.

NOTES

1. The Thai National Research Council kindly arranged for permission to conduct field research, which was supported by a Ford Foundation Southeast Asia Fellowship Program grant. The faculty and students of Sri Nakharinwirot University, Songkhla, have been very helpful, and special thanks are due to Acan Suthiwong Phongphaibun and to Acan Banyat Ruangsri, now of Silpakorn University. Acan Theraphan L. Thongkum and her colleagues in the Indigenous Languages Research Project organised a field seminar at Chaiya, Suratthani Province, in September 1977, to which they kindly invited me. The Thai proceedings of the seminar (Thongkum 1978) are the basis for the Tha Chang, Phanom and Kanchanadit data reported in this study. Participants Jimmy G. Harris and Jerry Gainey kindly discussed problems in Southern Thai data with me.

2. Mr Manun Phayakamas and family provided kind hospitality at Langsuan and invaluable assistance in rural explorations to locate 'missing' glottal stops.

3. Tone A is associated with lengthening for some items in East-coast Southern varieties, e.g. prenasally with -ii-:

	Central Thai	Songkhla
'ginger'	$khiŋ_{24}$	$khiiŋ_{45}$
'stone'	hin_{24}	$hiin_{45}$
'sea leech'	$pliŋ_{33}$	$pliiŋ_{34}$, etc.

Although superficially an exception to Gandour's observation, the above examples actually confirm it: length is again associated with rising contour (irrespective of etymological tone category).

4. In fact the items *'father'* and *'stomach'* tend to be kept apart by speakers who otherwise make similar mergers. In the course of field-work an elderly farmer and his (totally merging) wife fell into a heated argument on the point. It should be noted that throughout the South varieties preserving -ʔ in careful isolated pronunciation tend to lose it in connected speech.

5. Bee (1965) has noted the lenition process in Central Thai items like càʔ (< càk) 'incompletive marker', kàʔ (< kàp) *'with'* and ma- (< màak) 'prefixal syllable in names of fruits'.

6. The Thai text is reproduced without note of provenance in Na-Nakhon (1973:49-57); translation provided.

7. This is not to suggest that disambiguating homophony is in any way an impetus for hybridisation. In other cases the hybrid forms create homophony not present in conservative rural speech (Diller 1979:68-69). One sporadic exception to regular correspondence is *'Year of the Rabbit'*, thɔɔʔ_{44} in varieties south of the Kra Isthmus with regular cognates in other Tai languages, but thooŋ_{52} ∿ thoŋ_{52} in Kra Isthmus varieties. The source of this form is unknown, but it conveniently avoids confusion with a mild taboo form thɔɔʔ_{55} (or thɔɔk_{44} to the south) *'to draw back the foreskin'*. In other vocabulary homophony is widely tolerated.

8. Empirical evidence for a direction-of-contour difference based on type of final consonant is fickle. Note, for example, -h in Jeh, which has raised rather than lowered pitch (Gradin 1966).

9. For example, final -ʔ in Khmuʔ has led Haudricourt (1954) to postulate a process for Vietnamese which has become influential in Sino-Tibetan tonogenetic theorising. Haudricourt's model may be correct, but Khmuʔ phonology taken as a whole leads one to suspect that at least some Khmuʔ final glottal stops are a consequence of a syllable structure condition which dislikes open-vowel finals (see Smalley 1961). For other tonal effects on consonants, see Maddieson (1977).

BIBLIOGRAPHY

ABRAMSON, A.S.

1972 Word-final stops in Thai. In Harris and Noss, eds 1972:1-7.

ABRAMSON, A.S. and D.M. ERICSON

1978 Diachronic tone splits and voicing shifts in Thai: some perceptual data. *Haskins Laboratories status report of speech research* SR53/2:85-96.

BEE, P.

1975 Restricted phonology in certain Thai linker-syllables. In Harris and Chamberlain, eds 1975:17-32.

BRADLEY, David

1977 *Papers in South-east Asian linguistics* No. 5. *PL*, A-49.

BROWN, J.M.

1965 *From ancient Thai to modern dialects*. Bangkok: Social Science Association of Thailand Press.

1975 The great tone split: did it work in two opposite ways? In Harris and Chamberlain, eds 1975:33-48.

DELL, F.

1969 Langues en contact et changements phonétiques; notes sur l'histoire des finales en Bu-yi (Dioi). *Société de Linguistique de Paris, Bulletin* 64:184-220.

DILLER, Anthony

1979 Tones, segments and Thai regional society. In T.L. Thongkum, et al., eds *Studies in Tai and Mon Khmer phonetics and phonology in honour of Eugénie J.A. Henderson*, 60-93. Bangkok: Chulalongkorn University Press.

DONALDSON, J. and C.N. DIEU

1970 *Tai-Vietnamese-English dictionary*. Saigon: Bo Giao-Duc Xuat Ban.

EGEROD, Søren

1961 Studies in Thai dialectology. *Acta orientalia* 26:43-91.

GANDOUR, Jack

1977 On the interaction between tone and vowel length: evidence from Thai dialects. *Phonetica* 24:54-65.

GETHING, T.W., J.G. HARRIS and P. KULLAVANIJAYA, eds

1976 *Tai linguistics in honor of Fang-Kuei Li*. Bangkok: Chulalongkorn University Press.

GRADIN, Dwight

1966 Consonantal tone in Jeh phonemics. *Mon-Khmer studies* II:41-43. Saigon: The Linguistic Circle of Saigon.

HARRIS, J.G. and J.R. CHAMBERLAIN, eds

1975 *Studies in Tai linguistics in honor of William J. Gedney*. Bangkok: Central Institute of English Language, Mahidol University.

HARRIS, J.G. and R.B. NOSS, eds

1972 *Tai phonetics and phonology*. Bangkok: Central Institute of English Language, Mahidol University.

HARTMANN, J.R.

1976 The waxing and waning of vowel length in Tai dialects. In Gething, et al., eds 1976:142-159.

HAUDRICOURT, A.G.

1954 De l'origine des tons en viêtnamien. *Journal Asiatique* 242:68-82.

JONES, R.B.

1960 *Laryngeals and the development of tones in Karen*. *Burma research society, fiftieth anniversary publication* No.1.

1965 On the reconstruction of proto-Thai. *Lingua* 14:194-229.

KHANITTANAN, Wilaiwan

1977 *Phasa Phuthai*. Bangkok: Thammasat University Press.

LI, F.-K.

1971 The final stops in Tushan. *Bulletin of the Institute of History and Philology, Academia Sinica* 43:195-200.

1977 *A handbook of comparative Tai*. *Oceanic linguistics special publication* No.15.

MADDIESON, Ian

1977 Tone effects on consonants. *UCLA working papers in phonetics* 36:91-110.

MATISOFF, J.A.

1970 Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case study. *Journal of the American oriental society* 90:13-44.

NA-NAKHON, Plueang

1973 *Panha phasa Thai*. Bangkok: Bamrungsat Press.

SARAWIT, M.

1975 Some changes in the final component of the Tai syllable. In Harris and Chamberlain, eds 1975:316-328.

SMALLEY, William

1961 *An outline of Khmuʔ structure*. *American oriental society essay* No.2.

THONGKUM, T.L., ed.

1978 *Siang lae rabop siang nai phasa Thai thin tai Cangwat Suratthani 16 amphoe*. Bangkok: Chulalongkorn University Indigenous Languages Research Project.

WHEATLEY, Paul

1961 *The golden Khersonese*. Kuala Lumpur: University of Malaya Press.

WYATT, D.K.

1975 *The crystal sands: the chonicles of Nagara Sri Dharrmaraja*. Cornell University Southeast Asia Data Paper No.98.

Diller, A. "A New High Tone in Southern Thai". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation*. A-62:133-154. Pacific Linguistics, The Australian National University, 1982. DOI:10.15144/PL-A62.133

THE RELATIVE FREQUENCY OF TONES IN THAI[1]

Jack Gandour and Mary Jane Gandour

0. INTRODUCTION

The purpose of this study is to analyse the relative frequency of occurrence of tones in a representative sample of (Standard) Thai. Such quantitative data about Thai tones is of special importance for questions concerning tone acquisition, tonal inventories, historical development of tonal systems, errors in production ('slips of the tongue') and perception ('slips of the ear') of spontaneous speech, diagnostic procedures and therapy programs for aphasics and other speech-hearing impaired populations, and problems of pronunciation faced in the teaching of Thai as a foreign language. No previous statistical analyses of Thai tones (consonants and vowels) exist, in marked contrast, for example, to the numerous phoneme frequency studies which have been conducted on English (Dewey 1923, Hayden 1950, Tobias 1959, Wang and Crawford 1960). The study herein is a beginning attempt toward filling this void in Thai phonetics and phonology.

1. METHOD

The material selected for investigation consisted of 25,000+ entries from the Haas (1964) *Thai-English Students' Dictionary*. These entries were gleaned primarily from the context of modern periodicals, novels, short stories, newspapers, government reports, and advanced textbooks on literature, history, and the sciences. For each entry, the Thai pronunciation is shown in a phonemic transcription. All entries except examples (preceded by an open circle in the dictionary) were included in the analysis.

Five phonemic tones are posited for Thai in the Haas transcription: (1) Mid /khaa/ *'cogon, a species of grass'*, (2) Low /khàa/ *'galangal,*

Gandour, J. and Gandour, M.J. "The Relative Frequency of Tones in Thai". In Bradley, D. editor, *Papers in Southeast Asian Linguistics No. 8: Tonation.*
A-62:155-167. Pacific Linguistics, The Australian National University, 1982. DOI:*10.15144/PL-A62.155*

a rhizome' (3) Falling /khâa/ *'slave, servant'*, (4) High /kháa/ *'to engage in trade'*, (5) Rising /khǎa/ *'leg'*. The pitch contours on syllables ending in stop consonants are assigned to the low, falling, and high tones: /phèd/ *'to be hot (peppery)'*, /khìid/ *'line'*, /phûud/ *'to speak'*, /phéd/ *'diamond'*. The pitch on unstressed CV syllables with short /a/ is assigned to the mid tone: /thahaǎn/ *'soldier'*, /thanàd/ *'skillful'*. This phonemic analysis of the Thai tones was followed without exception in tabulating their frequency of occurrence.

From the phonemic transcriptions, the tones occurring on each syllable were counted, and the raw totals of the five Thai tones in all the entries were calculated in terms of their percentages of the grand total of tones.

2. RESULTS AND DISCUSSION

The number of occurrences and frequency percentage of each of the five Thai tones is shown in Table 1. The tones ranked in order of frequency of occurrence are (1) mid, (2) low, (3) falling, (4) high, and (5) rising. The mid tone is clearly predominant; it outnumbers the second-ranked low tone by nearly two-to-one. The statistical predominance of the mid tone may be partly attributed to the Haas transcription, in which short CV syllables with short /a/ are assigned to the mid tone.

TABLE 1

RELATIVE FREQUENCY OF OCCURRENCE OF THAI TONES
BASED ON HAAS DICTIONARY ENTRIES

Tone	Number of Occurrences	Frequency Percentage
(1) MID	24,479	39.98
(2) LOW	12,683	20.72
(3) FALLING	10,612	17.33
(4) HIGH	7,231	11.81
(5) RISING	6,217	10.16
Grand Total	61,222	

The distribution of tones in the adult language is an important factor to keep in mind in trying to make generalisations about the acquisition of tones in child language. As Li and Thompson (1978) pointed out, "A statistical survey of tone distribution in both the adult language and in the lexicons of the children's language may be relevant in explaining the order of acquisition of tones in a particular language." Tuaycharoen's (1977, 1979) findings in the acquisition of Thai; however, cannot be fully explained on the basis of the distribution of tones in the adult language. Although she found that her subject acquired the mid and low tones first, the falling and high tones were acquired last. The rising tone was acquired earlier than the falling and high tones. In terms of the chronological order in which the tones were acquired, the rising tone came **before** either the falling or high tones; in terms of tonal distribution in the adult language based on the results of this study, the rising tone comes **after** both the falling and high tones.

Finally, we wish to point out possible directions for future research on the relative frequency of occurrence of Thai phonemes. First, a similar frequency study of Thai consonants and vowels, again based on the Haas dictionary entries, would be a logical extension of the current study. Second, a frequency study of Thai phonemes within 'smooth' syllables (Proto-Tai tonal categories A, B, C) and 'checked' syllables (Proto-Tai tonal category D) should provide valuable quantitative information for both diachronic and synchronic investigations. Third, a frequency study of Thai phonemes based on a representative sample of spoken materials should enable us to determine to what extent the results of the current study, based on a sample of printed materials, are comparable. Fourth, a frequency study of Thai tones in connected speech, following an alternative method of transcription in which certain unstressed CV syllables with short /a/ (perhaps /i, u/) are treated as 'toneless' (cf. Abramson 1979), would probably yield a slightly different picture of the relative frequency of occurrence of the mid tone in comparison to the other four tones. These various possible extensions of the current study further indicate the need for one to take the following factors into consideration when interpreting the results of frequency studies: the size of the sample, the nature and content of the sample, and the method of transcription of the sample.

N O T E S

1. This material is based upon work supported by the National Science Foundation under Grant No. 7905854.

BIBLIOGRAPHY

ABRAMSON, A.S.

1979 Lexical tone and sentence prosody in Thai. *Proceedings of the 9th International Congress of Phonetic Sciences*, 380-387. Copenhagen: Institute of Phonetics, University of Copenhagen.

DEWEY, G.

1923 *Relative frequency of English speech sounds*. Cambridge, Mass.: Harvard University Press.

HAAS, M.R.

1964 *Thai-English student's dictionary*. Stanford, California: Stanford University Press.

HAYDEN, R.E.

1950 The relative frequency of phonemes in General-American English. *Word* 6/3:217-223.

LI, C.N. and S.A. THOMPSON

1978 The acquisition of tone. In V.A. Fromkin, ed. *Tone: a linguistic survey*, 271-284. New York: Academic Press.

TOBIAS, J.V.

1959 Relative occurrence of phonemes in American English. *Journal of the Acoustical Society of America* 31:631.

TUAYCHAROEN, P.

1977 The phonetic and phonological development of a Thai baby: from early communicative interaction to speech. Ph.D. thesis, School of Oriental and African Studies, University of London.

1979 An account of speech development of a Thai child: from babbling to speech. In T.L. Thongkum, V. Panupong, P. Kullavanijaya and M.R.K. Tingsabadh, eds *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie J.A. Henderson*, 261-277. Bangkok: Chulalongkorn University Press.

WANG, W. S-Y. and J. CRAWFORD

1960 Frequency studies of English consonants. *Language and speech* 3:131-139.