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'*; [ʔɬœʸ] : *'to pick'*, e.g. *'the nose'*;
[tɕɪ̆ŋ] : *'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';*		[sʅʔ] : *'water';*	
[kɑ̆ʔ] : *'river';*		[tɐ̆ŋʔ] : *'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';*
[mɐ̳̆ŋ]	*'door';*	[d̥ʑ̥ɪ̳̆ŋ]	*'town';*
[ʑ̥ʅ̳]	*'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.

[ʑ̥ʅ̳]	*'word';*	[d̥ɪ̳ʔ]	*'ground';*
[d̥ʑ̥ɪ̳̆ŋ]	*'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:

[fɐ̆ʔ]	*'law';*	[tɪ̆ʔ]	*'target';*
[tɕjɐ̆ʔ]	*'foot';*	[tɕhyɔ̆ʔ]	*'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:

[ɣ̥ě̳?] : *'buddha'*; [d̥ʑ̥ɥɒ̳̌?] : *'office'*;
[d̥ʑ̥ɛ̳̌?] : *'straight'*; [ɲɥɒ̳̌?] : *'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. [p^h 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, ɣ̥, ɡ̊, d̥ʑ̥] 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. [tsʰɑ, pʰi], or complex vowel finals, e.g. [wei, iju, ja, 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. [d̥z̥ɔ̱̌ŋ] : *'heavy'*, and C_sLq stands for any type 6 syllable with sonorant Initial, e.g. [mɔ̱̌ʔ] : *'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_o 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_o shapes are shown synchronised, in order to facilitate accurate comparison between the two parameters at any point in time. F_o 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 CǓ and CL̆ 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 CǓ 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 CL̆ syllables is 116 Hz. The total mean F_o range in CV(q) syllables is 169 Hz (CUq, at 0%DF) minus 105 Hz (CL̆, at 30%DP) = 64 Hz. The F_o register value for CǓ syllables is then

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

and for CL̆ syllables

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

Therefore, CL̆ and CǓ 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 (10	165 6	162 7	159 7	156 8	151 7	146 7	139 7	132 7	121 8	110 9)
Amplitude	17.9 (2.3	22.1 2.7	22.4 2.8	22.6 2.5	22.5 2.3	22.4 2.4	22.3 2.3	21.5 2.0	20.5 2.2	18.6 1.9	14.9 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 (10	181 9	176 8	172 8	168 8		161 9	153 10	145 9	136 9	126 8	116 7)
Amplitude	19.8 (3.4	25.3 2.5	24.6 1.7	23.5 1.4	23.0 1.3	22.6 1.5	22.7 1.3	22.3 1.1	21.1 1.2	19.8 1.3	17.2 1.2	13.8 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 (CǓ).

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 (CǓŋ).

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_sL̂).

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

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

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 (CĽŋ).

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).

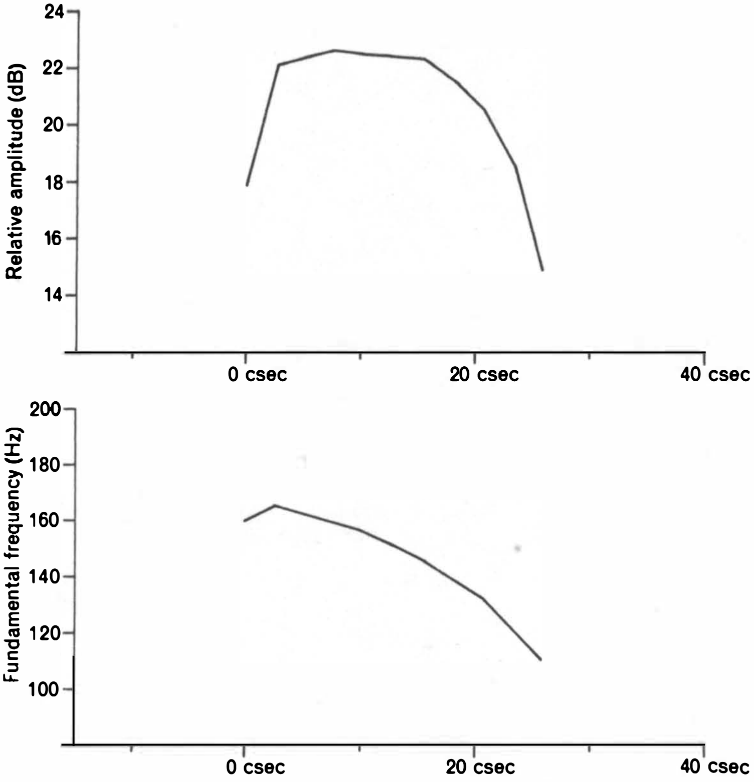


FIGURE 2

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 (CŬŋ). 'N' Indicates Point of Onset of Velar Nasal

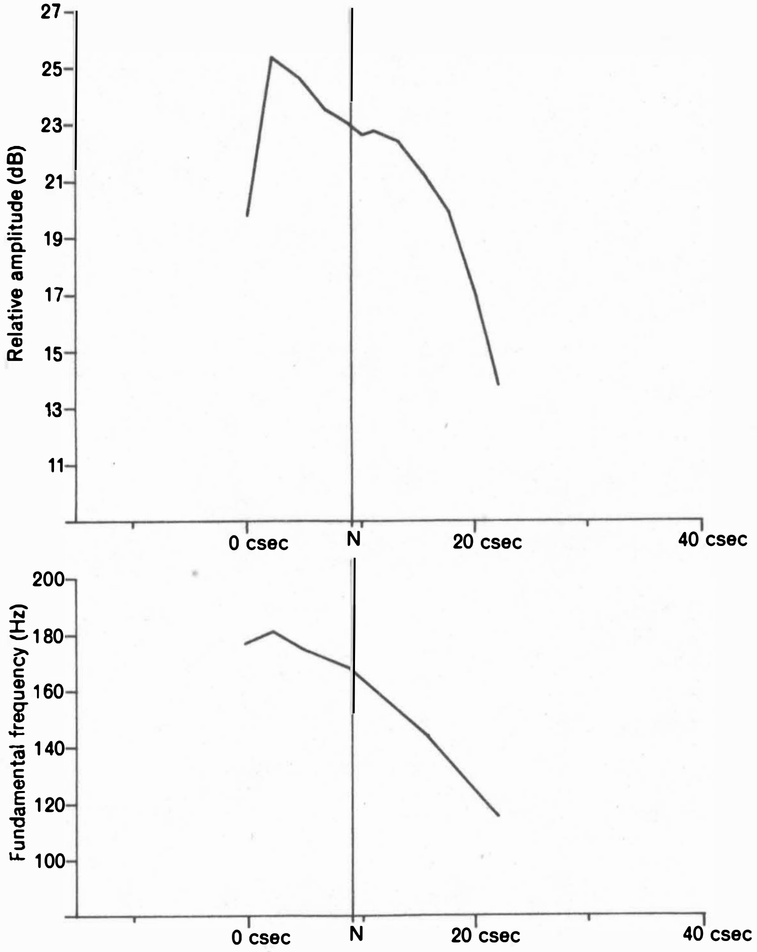


FIGURE 3

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 (CŬ).

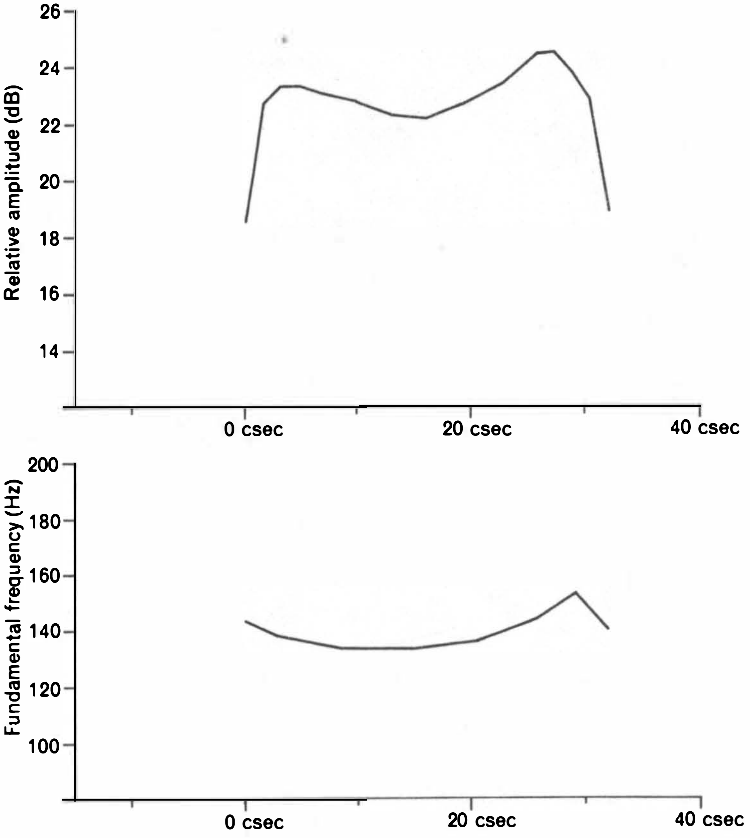


FIGURE 4

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.

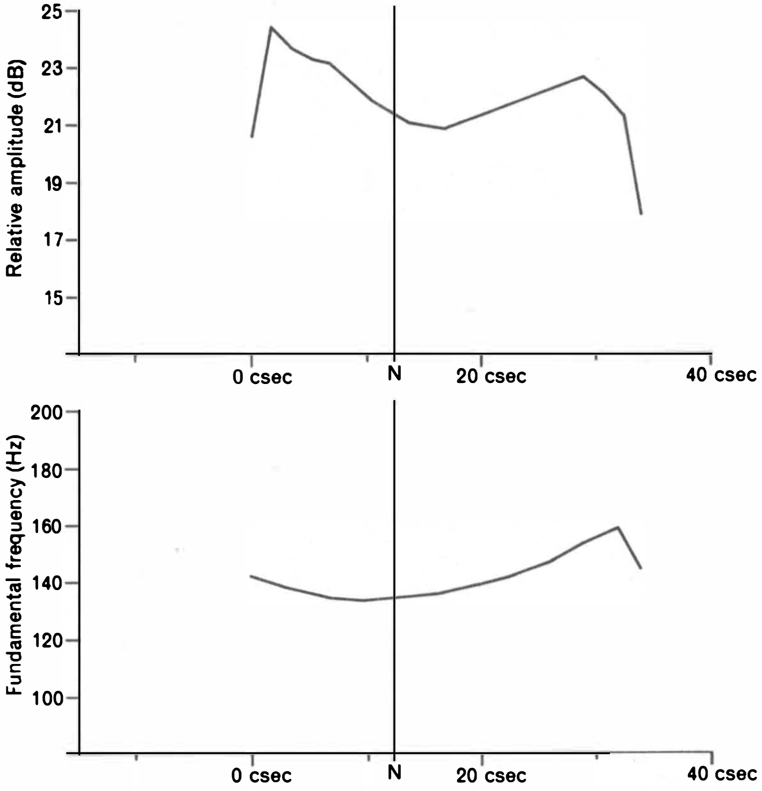


FIGURE 5

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

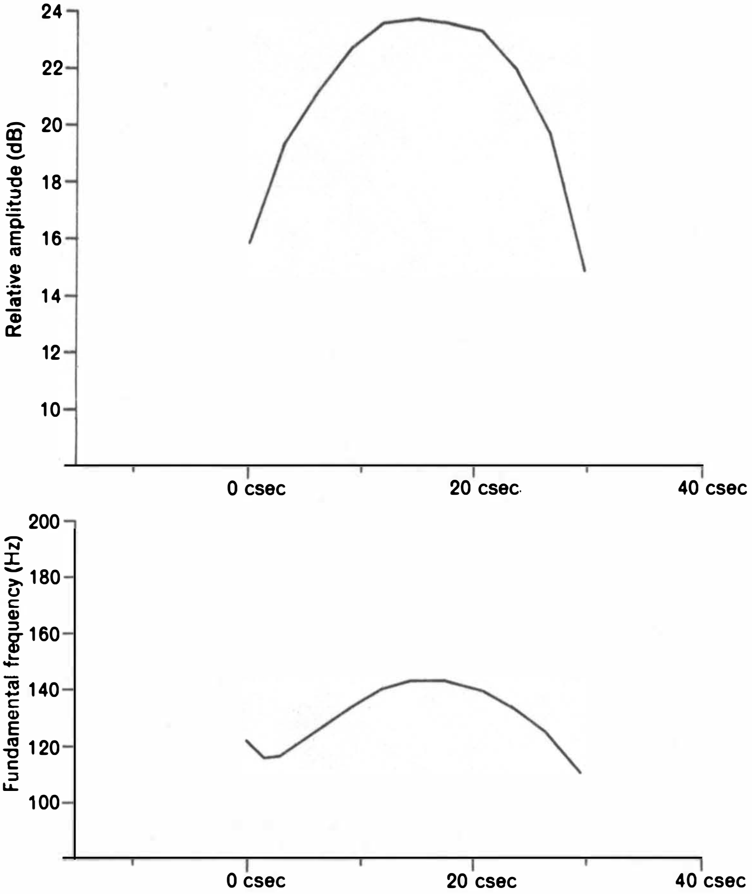


FIGURE 6

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.

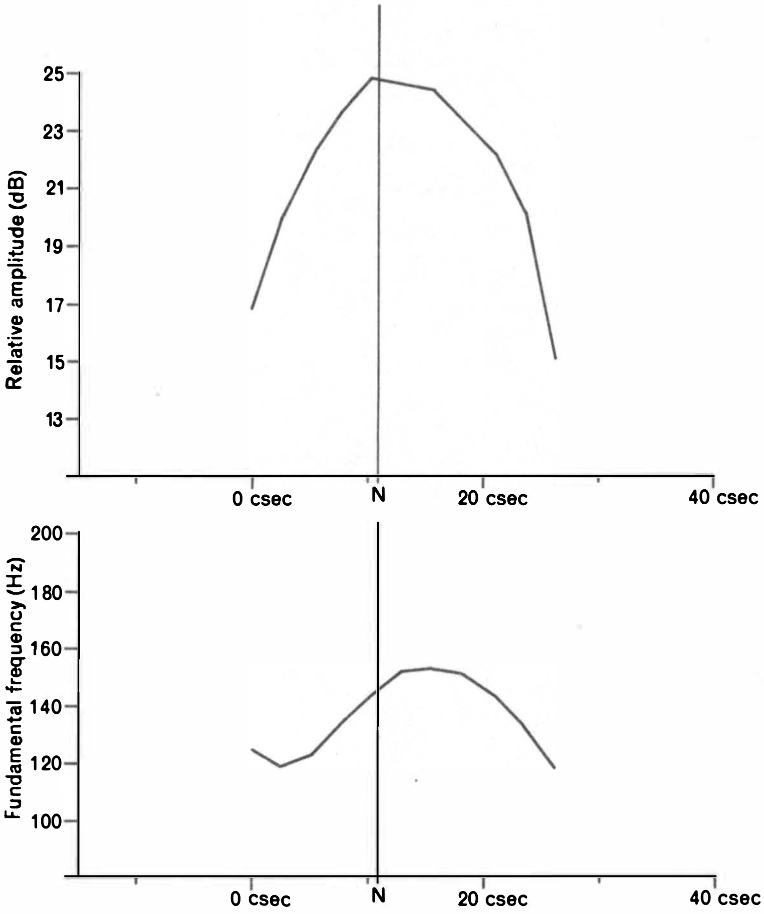


FIGURE 7

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.

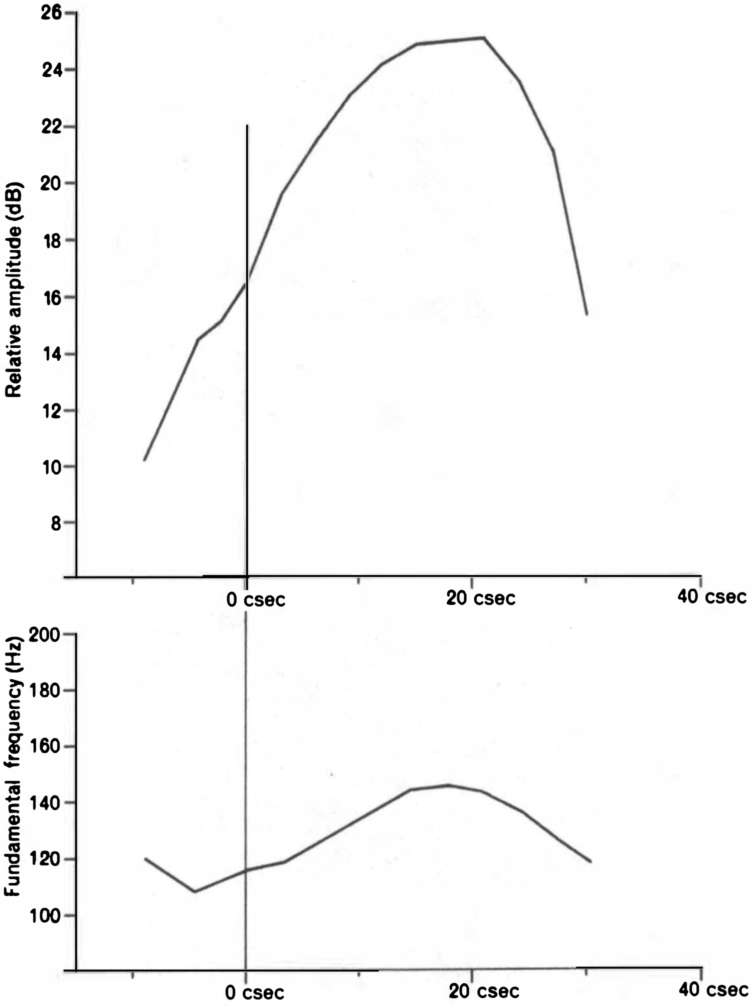


FIGURE 8

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

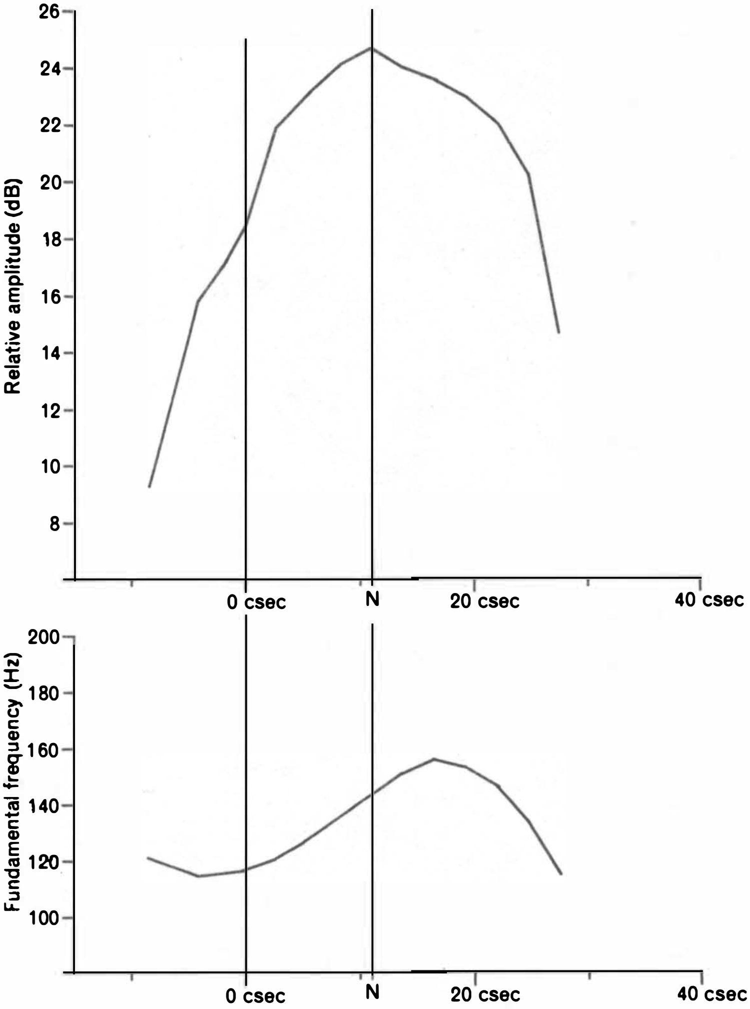


FIGURE 9

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).

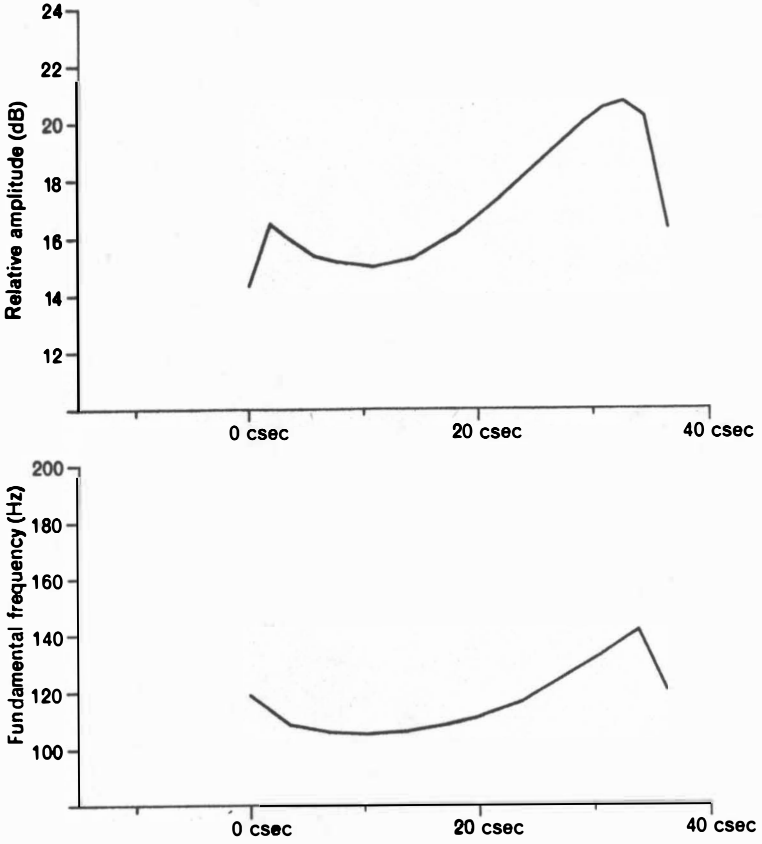


FIGURE 10

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.

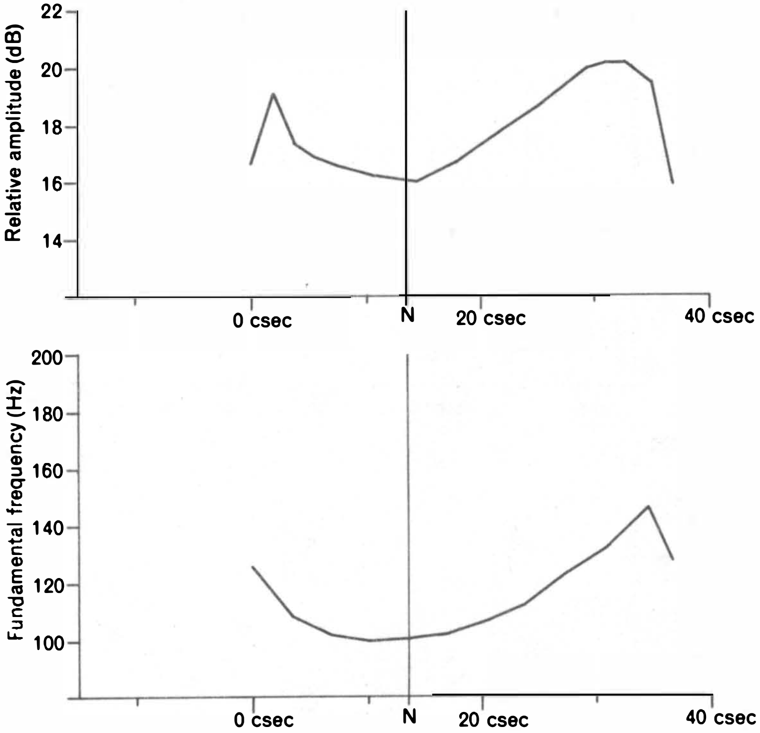


FIGURE 11

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 ($C_s\check{L}$). The Point of Onset of Vowel Final is Located at csec 0.

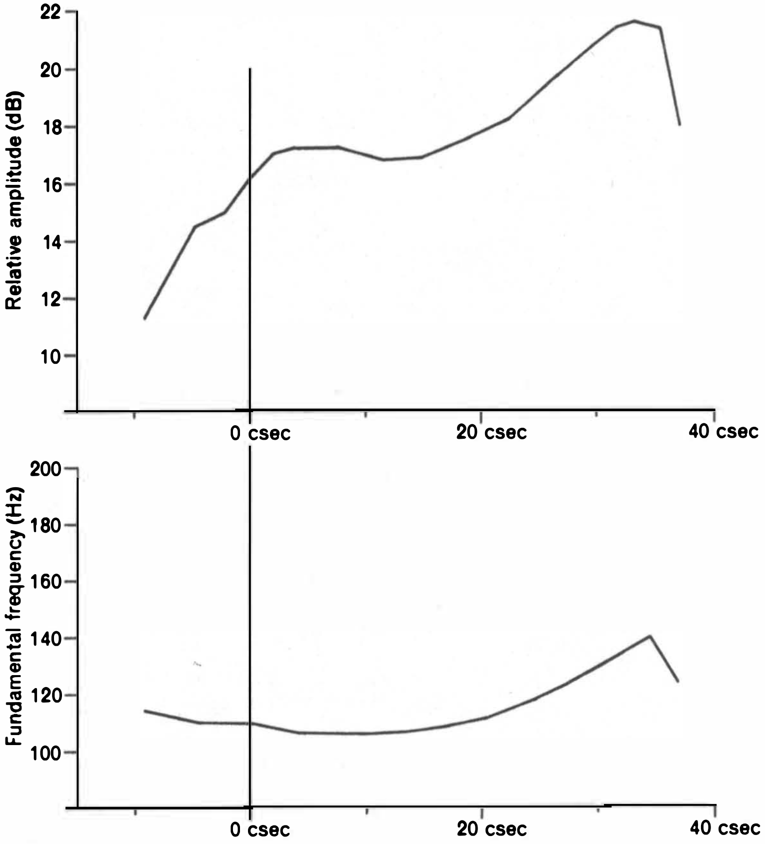


FIGURE 12

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).

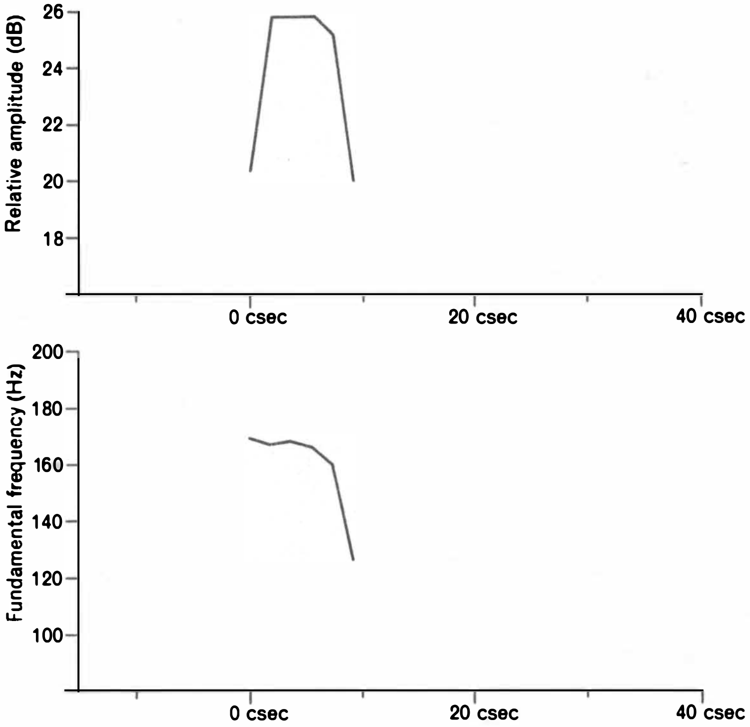


FIGURE 13

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).

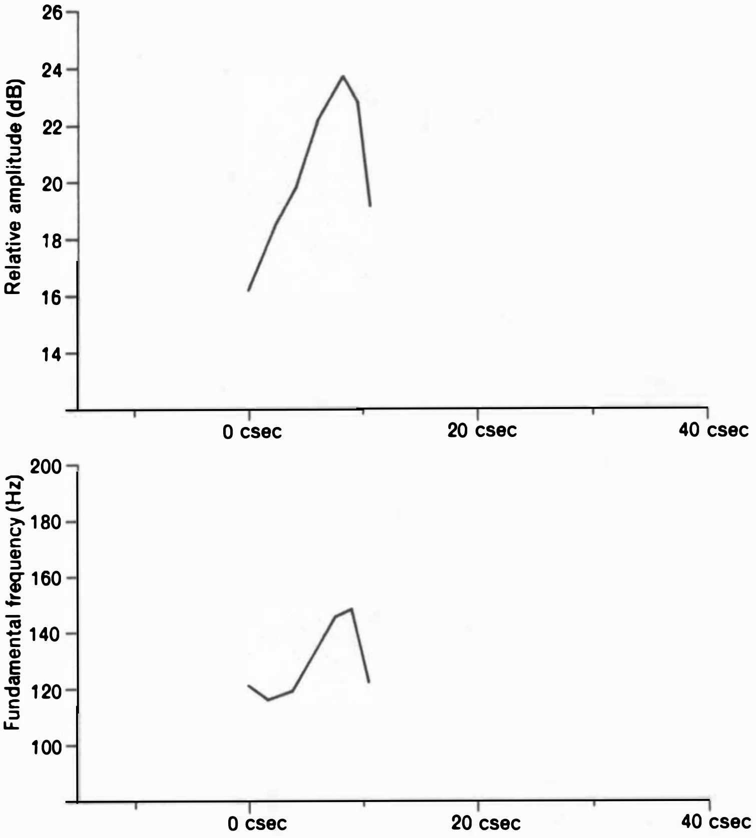


FIGURE 14

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.

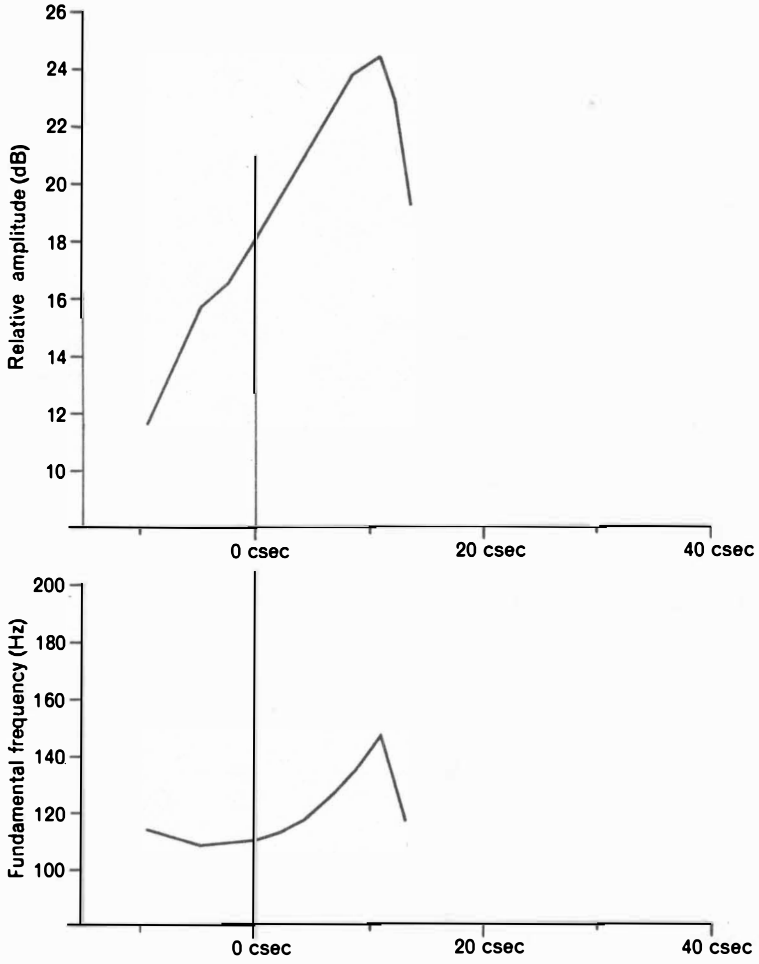


FIGURE 15

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.

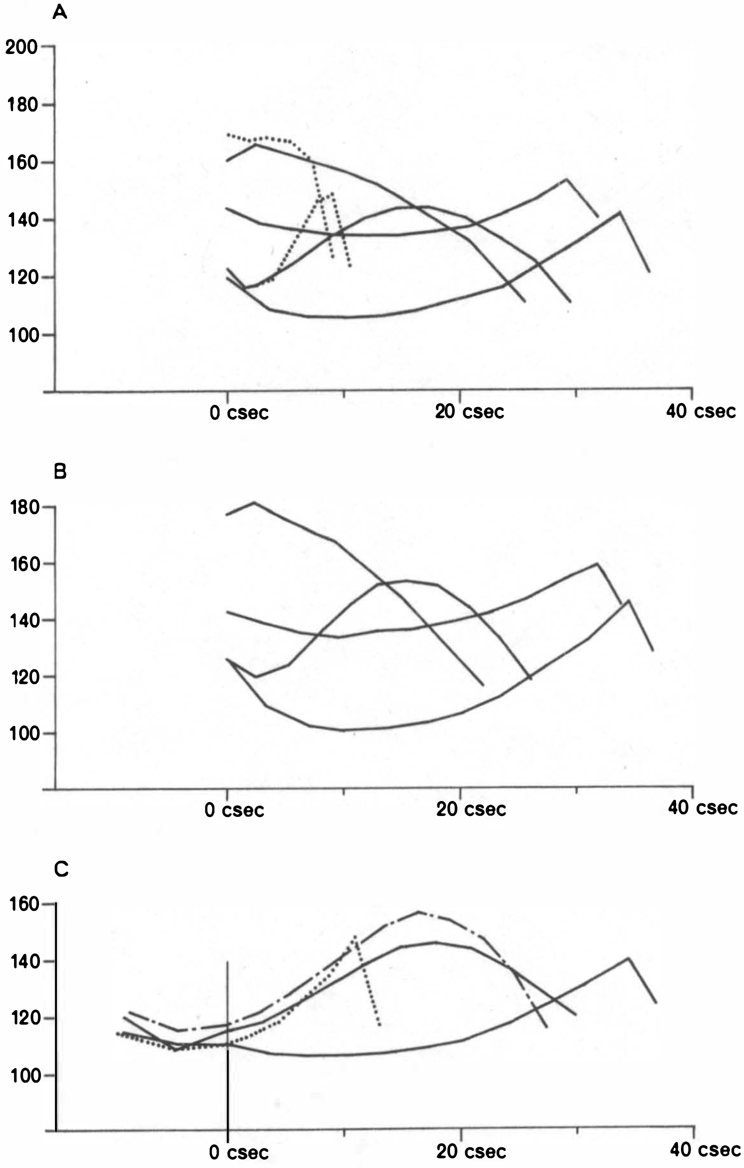
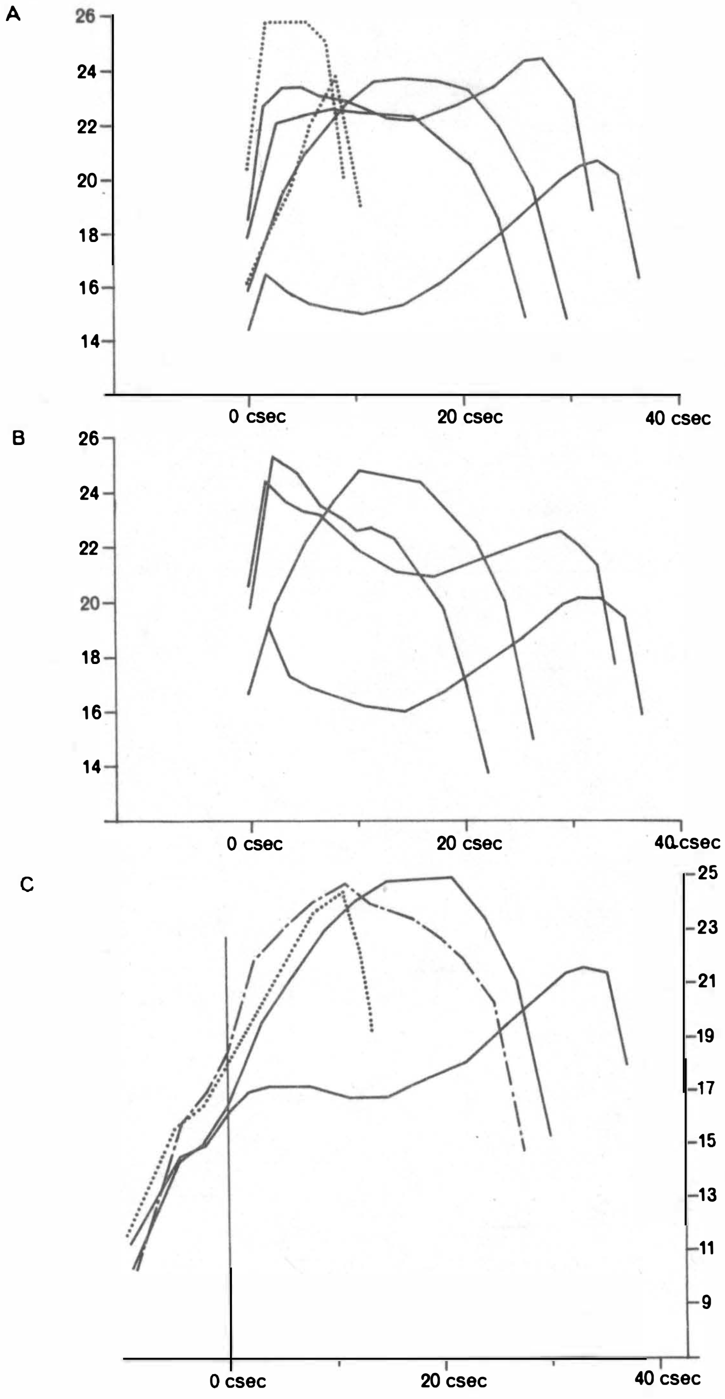


FIGURE 16

Relative Amplitude Values Compared for (A) CV(q), (B) CVŋ, and (C) C_sL(q) / $C_s\hat{L}$ŋ Syllables; ...... = $C_{(s)}$Vq Syllables; —·—·— = $C_s\hat{L}$ŋ 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 CǓ 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 CǓ and CL̆ 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 CL̆ contour falls quicker and starts to rise earlier than the CǓ 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 CǓ AND CL̆ F_o CONTOURS

Source	**Sums of Squares of Residuals (3 degree polynomial)**	**df**	**Mean Squares**	**Variance Ratio**
Combined regression, CǓ, CL̆	731.18	18	40.62	
Sum of regressions, CǓ, CL̆	148.58	14	10.61	2.83
Difference, combined regression - sum of regressions		4	30.01	

(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 CǓ and CL̆ 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. CǓ syllables have an Ao register value of 72%, compared with 24% for CL̆ syllables. There is therefore an Ao register difference between CL̆ and CǓ syllables of 48% (cf. 36% for F_o register).

The difference in Ao contour between CǓ and CL̆ 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 CǓ syllables: note that, in CǓ 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 CǓ 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 CǓ syllables is therefore a function of both register and contour differences.

CǓ 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 CǓ 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 CǓ (——) and CĽ (—·—·—) Syllables

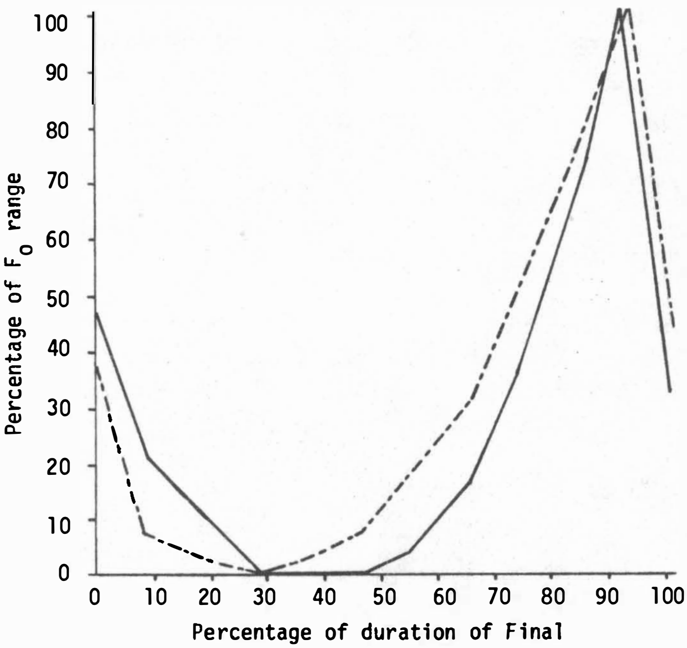
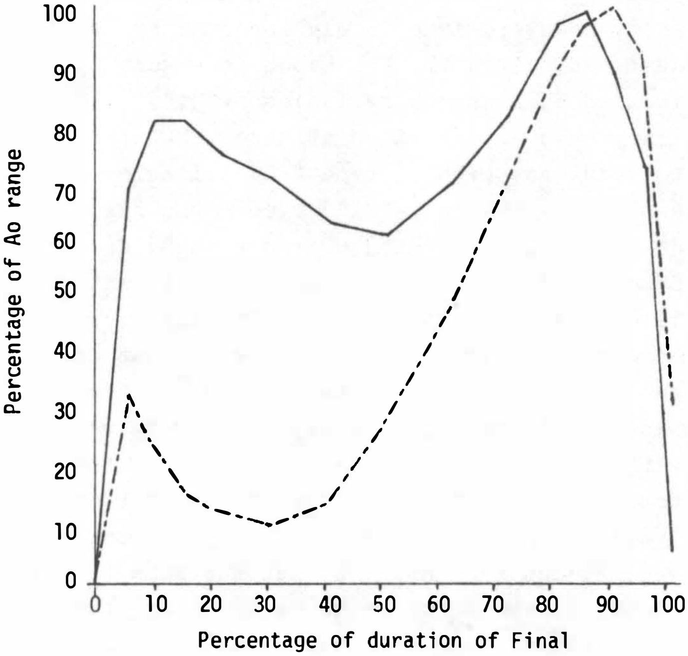


FIGURE 18

Comparison of Ao Contours in CŬ (——) and CL̆ (—·—·—) 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?); CL̂: 232; CL̆: 213, CUq: $\underline{5}$; CLq: $\underline{23}$ ($\underline{24}$?).

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%:CÙŋ, 64%:CÙ; 53%:CǓ; 52%:CǓŋ; 46%:$\mathrm{C_s}$L̂ŋ, 44%:CL̂ŋ,
41%:$\mathrm{C_s}$L̂, 38%:CL̂; 39%:CLq, 31%:$\mathrm{C_s}$Lq; 19%:CĽŋ, 17%:CĽ, 16%:$\mathrm{C_s}$Ľ.

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_o contour has been posited. Ohala (1978:31) notes that speakers appear to be able to produce falling F_o contours faster than rising ones, which explains why tones with rising F_o have a longer duration than tones with falling F_o. Also, correlations have been reported between duration and F_o register, and duration and F_o range: in syllables with falling F_o, duration decreases with F_o register, whereas in syllables with rising F_o, duration depends on the range of F_o 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_o range, and is therefore a plausible extension of the previously noted correlation for syllables with rising F_o : syllables of type 4 and type 2 both have the same concave F_o contour, but the mean F_o 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_o 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_o contours, are shorter than types 2 and 4, which have falling-rising contours, and type 6 syllables, with rising F_o contours, are shorter than type 1 syllables, which have falling contours.

Now a convex F_o contour, like that of Sh-Zh type 3 syllables, is very rare in Chinese dialects: of the 3433 F_o 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($^{\text{q}}_{\text{ŋ}}$) 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_o, (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_o/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_o, 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_o 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 [dʑɥɔ̌ŋ], t and dʑ are Initials, and i and ɥɔ̌ŋ are Finals.

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

5. For typographical convenience, the symbol [ɩ] 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 [± 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_o, the JND is apparently ± 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 sιstəm 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 *Nasálfest: 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.

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*