Phonetica *30:* 129–148 (1974)

# On the Domain of Tone in Mandarin[1]

## Some Acoustical Evidence

JOHN M. HOWIE

Indiana University, Bloomington, Ind.

*Abstract.* Measurements of the fundamental frequency of 136 Mandarin citation syllables have yielded average curves for the four distinctive tones in nine types of syllable. It appears that the basic contours of Mandarin tones are coextensive only with the syllabic vowel and any voiced segment that may follow it, while those portions of the pitch pattern that occur with an initial voiced consonant or a non-syllabic vowel are merely anticipatory adjustments of the voice.

## *The Tones of Mandarin*

The four tone contours which can be observed in citation forms of monosyllabic morphemes in Mandarin are described by CHAO [4] as follows (pp. 24, 25):

| Tone | Chinese name | Description | Pitch |
|---|---|---|---|
| 1 | Yīnpíng Shēng | high-level | 55 |
| 2 | Yángpíng Shēng | high-rising | 35 |
| 3 | Shǎng Shēng | low-dipping | 214 |
| 4 | Qù Shēng | high-falling | 51 |

[1] This paper is an expanded version of a paper read at the Annual Meeting of the Modern Language Association of America, December, 1969. Most of the work reported here forms part of the author's Ph.D. dissertation, submitted to Indiana University in October, 1970. (The dissertation is being published as *Acoustical Studies of Mandarin Vowels and Tones* in Princeton-Cambridge Studies in Chinese Linguistics. The author wishes to thank the Syndics of the Cambridge University Press for their kind permission to include parts of this monograph in the present article.) The author wishes to acknowledge his indebtedness to his professors, F. W. HOUSEHOLDER and WU-CHI LIU, and to A. S. ABRAMSON. This research has been supported in part by the Department of Romance Languages, University of Missouri-Columbia, and the author wishes to express his gratitude for this support.

'If the average range of the speaker's voice is divided into four equal intervals separated by five points: 1 low, 2 half-low, 3 middle, 4 half-high, and 5 high, any tone can be fairly well represented by giving its starting and ending pitch, and, in the case of circumflex tones, the turning point.' The *Pinyin* romanization, used in this study, indicates the tones by means of diacritic marks placed over the syllabic vowel. (See table I for examples of tonal distinctions in Mandarin.)

## *The Aims of this Study*

The distinctive patterns of voice pitch obviously are restricted to the voiced part of the syllable. This includes not only vowels, but voiced consonants, as in [mɑ] and [lɑ]. The question arises whether regular variations can be observed in the patterns of voice pitch in syllables with a voiced consonant, compared to syllables with vowels only or with a voiceless consonant[2]. It seems possible that the variations in the patterns are such that the portion occurring with an initial voiced consonant could be disregarded as irrelevant to the tone contours, and that the tone contours could be viewed as occurring only with the vocalic (and final nasal) segments of the syllable. Very little comment on the question of the domain of tone is to be found in the published literature on Mandarin. While conflicting views are expressed concerning the relevance of an initial voiced consonant, no author takes the position that the pitch pattern of an initial non-syllabic vowel, as in [$^{i}\alpha^{U}$] and [$^{u}$ən], is irrelevant to the tone contours.

SHEN *et al.* [12] choose to include a non-syllabic vowel in their measurements but to exclude a voiced consonant, explaining: 'In other words, for both tone-2 and tone-3, the pitch pattern extended over all of any preceding on-glide (semi-vowel) and all of any following off-glide (semi-vowel) or nasal. On the other hand, it did not extend over a preceding /r, l, m, n/; nor of course a preceding stop, affricate, or fricative...'

BROTZMAN [2] presents pitch patterns of the four basic tones of Mandarin, as spoken in isolated syllables with a voiceless initial consonant, which he says 'approximate the characteristic pattern of the tone'. And, 'because of the slightly different pitch pattern which is produced', he presents separately the patterns from syllables having a

[2] CHAO suggested investigating this question in a personal interview with the author, September, 1963, Bloomington, Ind.

voiced initial consonant, i.e. [mα], 'with these voiced initials included in the pitch patterns'. BROTZMAN observes: 'The voicing during the consonant is approximately level, and tends toward a middle position in the pitch range. It was assumed that the boundaries between consonants and vowels would be known by other methods, and in this study no attempt has been made to specify the pitch contour during the period of voicing for consonants. Naturally this must be done before one can completely specify a pitch curve.'

CHENG [6] takes into consideration, in assigning structural relationships among phonemes, 'the notion of rhyming which has been the central motif of traditional Chinese phonology'. This notion may have some application to the question of the domain of tone in Mandarin. CHENG discusses the first four basic structure rules of his syllable grammar as follows:

*Rule 1*
Our first division of the syllable into the tone and segment is based on the tonal morphophonemic transformations in which the tone changes independently of the rest of the syllable which we call segment ...

*Rule 2*
The segment is divided into the initial and the final because it is the final that carries the relevant tone pitch. Some initials m, n, l, r may carry pitches, but their values are irrelevant in determining the tone.

*Rule 3*
The final is divided into the medial (i, u, ü) and the rhyming part because it is the rhyming part as a whole not the final that participates in rhyming. Some rhyming schemes require identity of the rhyming part plus the tone for rhyming; some require identity of the rhyming part only ...

*Rule 4*
The rhyming part consists of a nucleus with or without a terminal phoneme (n, ng, i, u, r).

WANG [13] points out the lack of interdependence between the behavior of tones and segments in the Chinese languages. In discussing the need to regard the tone features as features of individual syllables, WANG writes, 'Phonetically, of course, the domain of tone is over the entire voiced portion of the syllable.'

CHAO [3, pp. 19, 25] writes, 'In addition to an initial and a final, every syllable has as an essential component a tone, which is primarily the pitch pattern of the voiced part of the syllable, so that, if the initial is voiced, the tone begins with the initial and spreads over the

whole syllable, while, if the initial is voiceless, the tone is spread over the final only... For example, in *nian* the tone begins with *n*, but in *tyan*, it does not begin until the *y*.'

KRATOCHVIL [10] objects to CHAO's [3] statement, just cited: 'This may have a kind of logic to it, and it is presumably substantiated by subjective kinesthetic observations, but it is quite untrue from the strictly phonetic viewpoint: instrumental analysis of the fundamental frequency patterns of Mandarin syllables commonly shows that voiced initial consonants have erratic $F_0$ characteristics which obviously do not contribute to the general tonal tendency of the given syllables.'

KRATOCHVIL [9, p. 36] states his position in these words: 'Although tones are, structurally speaking, IC's of the whole syllables, their actual physical shape is the way in which they modify the only obligatory segmental component of the syllable, that is the vowel.' But since he includes diphthongs and triphthongs in his definition of the term *vowel*, as used for describing Mandarin, KRATOCHVIL apparently does not mean to exclude an initial non-syllabic vowel from the domain of tone.

DOW [7, p. 102] views the pitch pattern of an initial voiced consonant as irrelevant to the tone contours, writing: 'The tone in Chinese is carried by the vocalic part of the syllable.' Like KRATOCHVIL [9], however, he seems to be excluding a final nasal consonant, which nevertheless should probably be included in the domain of tone.

The quotations given above reveal the majority opinion excluding an initial voiced consonant – a claim to be confirmed by an acoustical investigation. One of the principal aims of this study, then, is to discover variations in the pitch pattern of each tone associated with differences in syllable structure: in particular, the kind of segment that precedes the syllabic vowel.

## *The Syllable in Mandarin*

Except for very rare dissyllabic morphemes, the syllable in Mandarin corresponds to the morpheme or to two morphemes when the morpheme *er* is suffixed to another morpheme. Homophonous morphemes are quite numerous in Mandarin because the constraints on the combination of segments in the syllable admit only about 400 different segmental syllables (without the *er* suffix). Taking into

account the factor of tone, there would be about 1,600 tonally differentiated syllables, but – with the accidental gaps in the distribution of the four tones – fewer than 1,200 syllables actually occur as unit morphemes.

In the traditional Chinese writing system, one morpheme is represented by one logograph, or character. Since there is a one-to-one correspondence between morpheme and syllable in all the language material used for the present study, it follows that the character corresponds to the syllable in this material.

The segmental structure of Mandarin citation syllables has been adequately treated elsewhere [3, 6]. The following syllable structure index should suffice as a linguistic framework in which to view the acoustical investigation of the tone contours. The index (which does not cover *er* or morphemically complex monosyllables) is consistent with CHENG's [6] basic structure rules, cited above. Optional segments are indicated by parentheses; only the syllabic vowel is obligatory.

| Segments | | | |
|---|---|---|---|
| (1) | (2) | 3 | (4) |
| (Initial consonant) | (Non-syllabic vowel) | Syllabic vowel | (Non-syllabic vowel or nasal consonant) |
| | | Nucleus | Terminal |
| | Medial | ← Rhyming part → | |
| Initial | ← Final → | | |
| ← Segmental syllable → | | | |

The optional non-syllabic vowels are high vowels; they occur only in syllables with a non-high syllabic vowel; and two non-syllabic vowels in a syllable (i.e., segments 2 and 4) cannot be the same. In terms of this structure index, the examples given above would be described as follows: *yao* [ˈɑᵾ] 234, *ma* [m*a*] 13.

The four tones occur simultaneously with segment 3, and with the sequence 34. From the results of this study, it is doubtful that the domain of tone also includes either segments 1 or 2. Following CHENG's [6] designation of segments 3 and 4 as the *rhyming part*, the domain of tone might therefore be identified with the *rhyming part* of the syllable.

## *Plan of Research*

This study has been limited to variations in the shape of the tonal contours found within citation syllables. Accordingly, a restricted list of Mandarin citation syllables was selected so that several variants of each tone, occurring in syllables of various segmental structures, were included. The prosodic features called tones are implemented mainly by differences in the perceived pitch of the voice. The physical dimension of sound most closely correlated with perceived pitch is fundamental frequency. The course of the fundamental frequency during the entire voiced part of the citation syllable was measured from spectrograms of the recorded utterances of one male speaker of Peking Mandarin. The frequency values for each syllable were then plotted logarithmically against duration, yielding a number of separate curves for each tone. After average curves had been drawn of each tone in each of the different syllable types, variations in shape emerged that are clearly associated with the segment preceding the syllabic vowel. Of greatest interest are the anticipatory movements of the fundamental seen during initial voiced consonants and non-syllabic vowels.

## *Word List*

It seemed desirable to construct the list of citation syllables from complete sets of monosyllabic morphemes minimally distinguished by tone, such as *mā* 'mama', *má* 'hemp', *mǎ* 'horse', *mà* 'to scold'. Fortunately, this goal was easy to achieve, as 68 such sets were found among the 1,192 citation syllables in the Inventory of Mandarin Syllables compiled by WANG *et al.* [15]. Because of the excessive amount of time required for the making and measuring of spectrograms, however, only half that number was included in the word list for this study. These 34 sets of tonally differentiated morphemes, comprising 136 morphemes, are given in table I, together with brief English glosses[3] to show that the morphemes are indeed all different.

As one of the objectives of this study was to discover variations in the pitch pattern of each tone associated with differences in syllable structure, the 34 sets of tonally differentiated morphemes were chosen so as to represent a wide variety of syllable structures – within the constraints that exist in Mandarin on the permissible sequences of vowels and consonants. These 34 sequences of vowels and consonants can be classified into nine types of syllable structures. This classification will be seen in table I, which presents the 136 Mandarin citation syllables in the quasi-phonemic *Pinyin* romanization, with a normalized phonetic notation given once for every set of monosyllabic morphemes minimally distinguished by tone. Type 1 syllables have an initial syllabic vowel; type 2 syllables have an initial non-

[3] The English equivalents of the Mandarin morphemes were found in CHAO and YANG [5].

*Table I.* Mandarin monosyllabic morphemes in 9 types of syllable structures, with concise English glosses

| | |
|---|---|
| *Type 1, with initial syllabic vowel* | |
| yī [i] | clothes |
| yí | doubt |
| yǐ | to take |
| yì | easy |
| yū [ü] | literal-minded |
| yú | remainder |
| yǔ | speech |
| yù | beforehand |
| wū [u] | house, room |
| wú | (a proper name) |
| wǔ | five |
| wù | affair |
| āi [$\alpha^{I}$] | sorrowful |
| ái | to suffer |
| ǎi | short, low |
| ài | to love |
| *Type 2, with initial non-syllabic vowel* | |
| yā [$^{i}\alpha$] | slave girl |
| yá | tooth |
| yǎ | elegant |
| yà | to be astounded |
| yāo [$^{i}\alpha^{U}$] | phantom |
| yáo | to shake |
| yǎo | to bite |
| yào | to want |
| yōu [$^{i}o^{U}$] | to be worried |
| yóu | to follow |
| yǒu | friend |
| yòu | again |
| *Type 3, with initial non-syllabic vowel and final nasal consonant* | |
| yāng [$^{i}\alpha\eta$] | center; to beg |
| yáng | sheep |
| yǎng | to support |
| yàng | manner |
| yuān [$^{\ddot{u}}$æn] | to wrong |
| yuán | round |
| yuǎn | far |
| yuàn | institution |
| yīng [ɪŋ] | brave |
| yíng | camp |
| yǐng | image |
| yìng | hard |
| wēn [$^{u}$ən] | to review |
| wén | writing |
| wěn | steady |
| wèn | to ask |
| *Type 4, with initial voiceless fricative* | |
| shī [ʂi] | teacher |
| shí | ten |
| shǐ | history |
| shì | market |
| fā [fɑ] | to send out |
| fá | punishment |
| fǎ | method |
| fà | France |
| xū [ɕü] | must |
| xú | slow |
| xǔ | to permit |
| xù | clues |
| hū [xu] | (interr. particle) |
| hú | irrelevantly |
| hǔ | tiger |
| hù | door |

*Table I* (continued)

| | |
|---|---|
| *Type 5, with initial voiced continuant* | |
| mā [mɑ] | ma, mama |
| má | hemp |
| mǎ | horse |
| mà | to scold |
| nā [nɑ] | (a surname) |
| ná | to take |
| nǎ | which? |
| nà | to pay |
| lā [lɑ] | to pull |
| lá | to slash |
| lǎ | a bugle |
| là | hot, pungent |
| *Type 6, with initial aspirated stop* | |
| pāo [pʰɑʊ] | to throw |
| páo | gown |
| pǎo | to run |
| pào | gun |
| pī [pʰi] | to pass on |
| pí | skin |
| pǐ | single man |
| pì | secluded |
| tū [tʰu] | bald |
| tú | diagram |
| tǔ | earth |
| tù | rabbit |
| tuō [tʰuɔ] | to let drag |
| tuó | humpback |
| tuǒ | secure |
| tuò | saliva |
| kē [kʰɣʌ] | a subject |
| ké | to cough |
| kě | be permitted |
| kè | a lesson |
| *Type 7, with initial unaspirated stop* | |
| bāo [pɑʊ] | to wrap |
| báo | hailstone |
| bǎo | guarantee |
| bào | newspaper |
| bī [pi] | to compel |
| bí | nose |
| bǐ | to compare |
| bì | to shut |
| dū [tu] | to supervise |
| dú | alone |
| dǔ | to gamble |
| dù | measure; degree |
| duō [tuɔ] | much, many |
| duó | to rob |
| duǒ | flower (AN) |
| duò | lazy |
| gē [kɣʌ] | elder brother |
| gé | frame |
| gě | (a surname) |
| gè | (aux. noun) |
| *Type 8, with initial aspirated affricate* | |
| qiē [tɕʰiɛ] | to cut |
| qié | eggplant |
| qiě | moreover |
| qiè | to steal |
| cān [tsʰan] | a meal (AN) |
| cán | remnant |
| cǎn | tragic |
| càn | cowardly |
| chī [tʂʰi] | to eat |
| chí | late |
| chǐ | Chinese foot |

*Table I* (continued)

| | | | |
|---|---|---|---|
| chì | wing | zān [tsan] | hairpin |
| | | zán | we (incl.) |
| *Type 9, with initial unaspirated affricate* | | zǎn | to hoard |
| | | zàn | to approve |
| jiē [tɕiɛ] | street | zhí [ʈʂi] | to know |
| jié | hero | zhí | straight |
| jiě | to guard | zhǐ | only |
| jiè | realm | zhì | ambition |

syllabic vowel; type 3 syllables have an initial non-syllabic vowel and a final nasal consonant; type 4 syllables have an initial voiceless fricative; type 5 syllables have an initial voiced continuant; type 6 syllables have an initial aspirated stop; type 7 syllables have an initial unaspirated stop; type 8 syllables have an initial aspirated affricate; type 9 syllables have an initial unaspirated affricate.

A randomized version of the word list, written only in Chinese characters, was prepared for the informants to read. The informants would thus be prevented from anticipating items in a related series, as might be expected if, for example, the four tones appeared successively in a conventional order. So that the syllables would be spoken in citation form, with a constant rhythm and intonation – yet without the sentence-final intonation characteristic of isolated items in a list – the morphemes were to be inserted in the carrier sentence: *Zhèige... zì, shì Lǎo Lǐ xiě de.* ('This word, ..., was written by Mr. Li.')[4] The sentence is an appropriate substitution frame for the monosyllabic morphemes – which are often bound forms that are meaningless in isolation – since the sentence refers to the character used to write the morpheme whose spoken form is to be uttered by the informant.

When the word list was recorded, it was noted that the syllable *zì* ('word'), which immediately follows the citation syllable with the tone to be measured, was spoken unstressed in every utterance of the carrier sentence. Because unstressed syllables have the 'neutral tone' in Mandarin, each speech sample to be analyzed consisted of a sequence of citation tone followed by neutral tone. This is a pattern typical of syntactic words in Mandarin – very frequently of words formed with the nominalizing suffix *-zi*. By this coincidence, many of the sequences in the word list turned out to be homophonous with dissyllabic nouns. CHAO [3, p. 154] gives a succinct explanation of this phenomenon: 'If a neutral tone occurs, it will form part of a word with the preceding syllable....'

## *Informant*

The Mandarin-speaking informant for this study was born in Peking and completed elementary school there. His high school years were spent in Peking and Chunking, and his university study was done on Taiwan. In 1965, when he contributed the voice recording of the word list, he was a Lecturer in Chinese at Indiana University in Bloomington. The

[4] The same carrier sentence was used in a study of vowel formant values by WANG *et al.* [15].

randomized word list (in characters only) was recorded by the informant, who pronounced each monosyllabic morpheme in citation form in the carrier sentence. The recording was later checked for normality by three other Mandarin speakers.

## *Instruments*

The voice recording was made in a sound-treated studio, with a Shure model 55-S dynamic microphone, at a tape speed of 7.5 in/sec on an Ampex model 351 tape recorder. The recording was subsequently played into the spectrograph from a Viking *Studio 95* tape transport with the appropriate preamplifier. The sound spectrograph used to analyze the informant's speech was a modified Kay *Sona-graph* (one of the early models). Wide- and narrow-band spectrograms with the standard 2,000 Hz/in scale were used as an aid in segmenting the utterance to locate the boundaries of the citation syllable or of its voiced segment. Narrow-band spectrograms with the expanded 200 Hz/in scale were used to make measurements of fundamental frequency.

## *Fundamental Frequency Measurements*

When the boundaries of the voiced segment had been marked on the spectrogram, the duration of the entire voiced segment was measured and specified in milliseconds (to the nearest even number). The points where the fundamental frequency was to be measured were then determined as percentages of this duration, allowing the segments to be normalized as to duration in order to make the shapes of the pitch curves directly comparable. Lines perpendicular to the baseline were constructed at the points so chosen.

Next, lines were drawn by eye through the centers of the first two or three harmonic streaks on the spectrogram. Since the frequencies of the harmonics are integral multiples of the fundamental frequency, the difference between any two adjacent harmonics is equal to the fundamental. Therefore, the distance between the centers of the first and second harmonics was measured along each perpendicular line. This distance, measured in inches and converted to herz, gives the fundamental frequency of the vocal cord tone at each point in time. Wherever the fundamental streak was obscure or absent, the second and third harmonics were measured instead[5].

A pair of dividers was used to transfer the fundamental frequency measurements from the perpendicular lines on the 200 Hz/in spectrograms to a scale having twenty divisions to the inch. Between divisions marked on the scale, five further divisions were estimated by eye, allowing the measured frequencies to be expressed to the nearest even number. It is felt that these round-number values reflect adequately the precision with which the measurements of the fundamental were made.

The points in the duration of the voiced segment at which the fundamental frequency was measured always included the beginning and the end of the voiced segment. Depending

[5] This procedure is described by ABRAMSON [1, p. 115], who measured the fundamental frequency in 88 Thai syllables (spoken by one male informant) at intervals of 25 or 50 msec. ABRAMSON [1, p. 119] explains, 'In order to arrive at average tonal curves, it was first necessary to normalize the raw curves as to duration. This was done by shrinking or stretching the curves proportionally to a convenient length.' The method of normalizing the curves that was used in the present study was suggested by F. W. HOUSEHOLDER.

on the duration of the voiced segment, from 5 to 21 measurements were made in each citation syllable, at intervals of some multiple of 5% of the duration. Very roughly, the measurements were made at equally spaced points, with more points being included as the course of the fundamental moved more steeply up or down. Care was taken to assure that measurements were made at absolute time intervals of no more than 25 msec during steeply moving stretches and of about 50 msec during less steeply moving and level stretches.

## *Average Curves*

The values obtained in this way for the fundamental frequency were plotted on a separate graph – using tracing paper – for each of the 136 syllables. The graphs were all of identical dimensions, with frequency represented on a logarithmic scale on the vertical axis, and percentage of duration on a linear scale on the horizontal axis. The plotted points on the graph were connected to form a curve, which can be said to portray the acoustic equivalent of the pitch pattern of the syllable in that graph. Because the voiced segments had been normalized as to duration before measurements were taken, all the plotted curves have, in effect, been 'stretched' or 'shrunk' to the same length.

The separate graphs of the syllables of one tone and one type were superimposed under a fresh sheet of tracing paper, the axes were drawn on the new graph, and an average curve was drawn through the plotted curves[6]. The resulting 36 average curves (4 tones in 9 types) are all the same length, and can be readily compared in shape.

The average curves of one tone are displayed in each one of the figures 1–4, with the curves assigned to graphs so that the same syllable types appear together in every figure. It will be seen that the syllable types have been grouped in such a way that the most similar curves are usually displayed in the same graph for easy comparison of minor variations in shape. Thus, each tone has similar curves in syllable types 1, 6, and 7. (Certain variations are pointed out below.) Even more pronounced are the similarities in types 2, 3, and 5, and in types 4, 8, and 9.

## *Descriptions of Tone Contours*

Before passing on to the variations among the average curves of various syllable types – and their implications for the question of the domain of tone – it is worth-while to consider these average curves as acoustic descriptions of the tonal features in Mandarin. From an inspection of the curves in figures 1–4, the contours of the four citation tones of Mandarin can be described as quite distinct, with an overall pitch range of approximately ten semitones. Ignoring for the moment the variations in shape, and observing only what the nine curves of each tone have in common, one can make the following statements about Mandarin tone contours: tone 1 is high in the upper half of the

[6] The same procedure for deriving average curves was used by ABRAMSON [1, p. 119], who states, 'This was done by eye as an easy and reasonable way of achieving weighting. It is believed that repetitions of the procedure would yield curves of essentially the same shape with only minor variations in absolute value.'

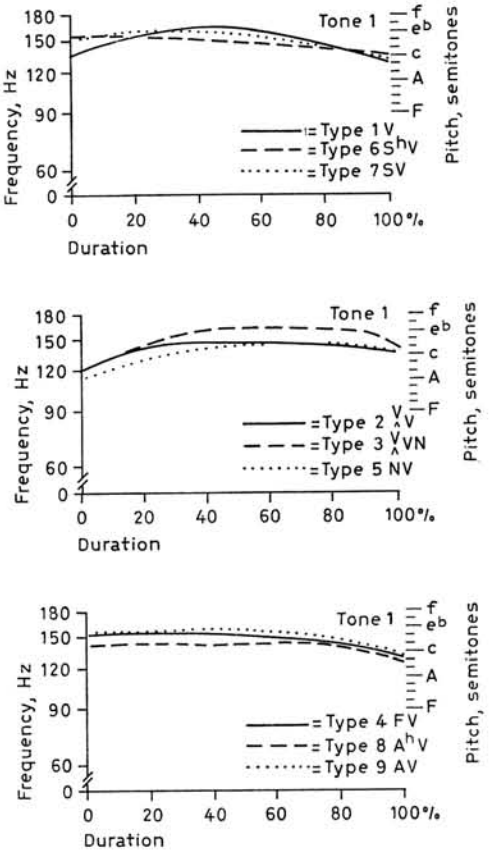


*Fig. 1.* Average curves of Mandarin tone 1: citation syllables in pre-neutral position.

pitch range; tone 2 rises through the upper half of the range; tone 3 remains within the lower half of the range; tone 4 falls through the upper half of the range.

These statements are in good agreement with CHAO's [4] familiar descriptions, given above. But a more detailed comparison is in order here, using the curves of syllable type 4, which might be regarded as 'basic' in that they seem to exemplify what all the average curves have in common for each tone. If the average curves of syllable type 4 are represented on a graph with CHAO's five-point pitch scale, as seen in figure 5, the four contours may be compared directly with CHAO's contours in terms of the points on the scale (pp. 24, 25, 85, 90).

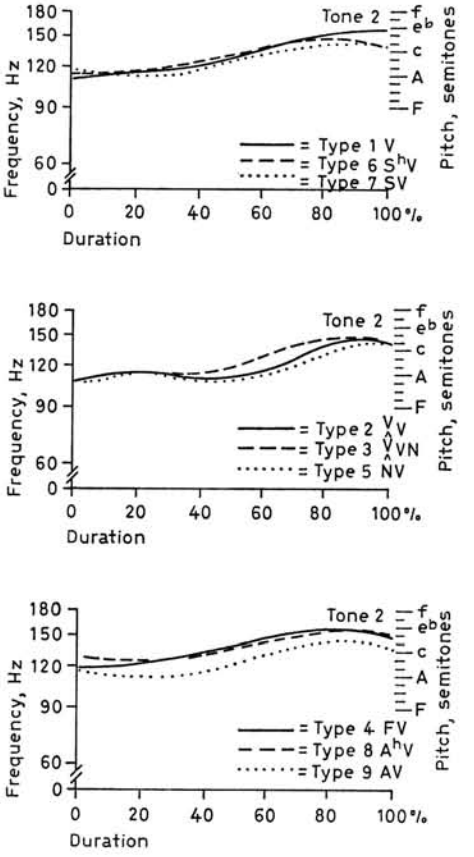


*Fig. 2.* Average curves of Mandarin tone 2: citation syllables in pre-neutral position.

| Chao's contours | |
|---|---|
| in isolation | before neutral tone |
| Tone 1 55 | Tone 1 55 or 53 |
| Tone 2 35 | Tone 2 35 |
| Tone 3 214 | Tone 3 21 |
| Tone 4 51 | Tone 4 51 |

Discrepancies between the type 4 average curves and Chao's contours – which may be seen mainly in the endpoints of tones 3 and 4 – will be discussed below.

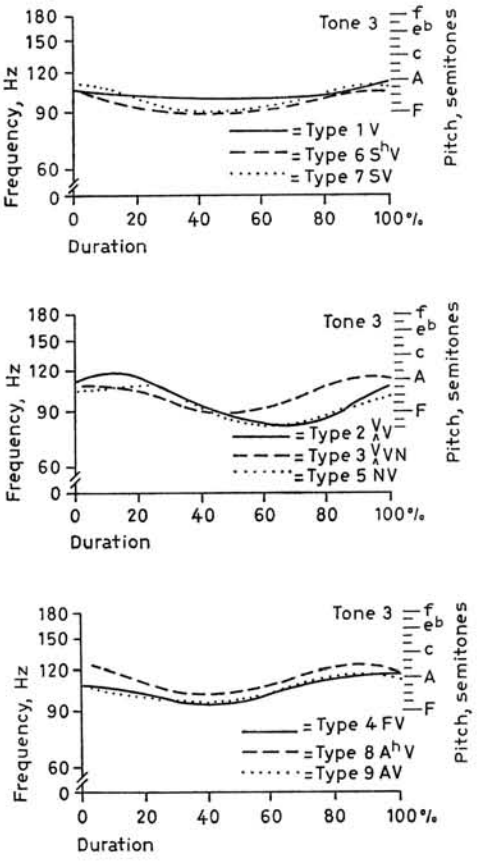


*Fig. 3.* Average curves of Mandarin tone 3: citation syllables in pre-neutral position.

## *Variations among Syllable Types*

There are variations in shape among the nine average curves for each tone, shown in figures 1–4. For example, there are variations between type 1, with initial syllabic vowels, and types 6 and 7, with initial stops. Tones 1 and 4 with initial vowels begin with a prolonged rise, and tones 2 and 3 with initial vowels do not dip as low as with initial stops. There are also slight variations between type 6, with aspirated stops, and type 7, with unaspirated stops. Tone 1 with aspirated stops begins quite level, while tone 1 with unaspirated stops

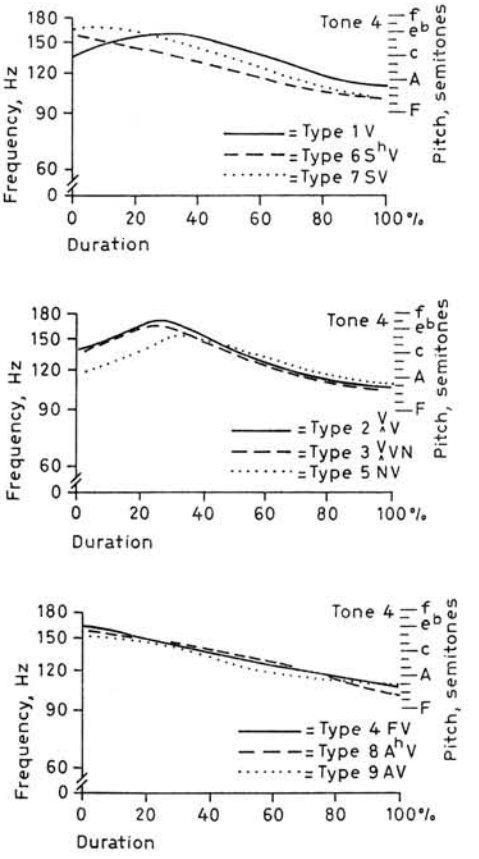


*Fig. 4.* Average curves of Mandarin tone 4: citation syllables in pre-neutral position.

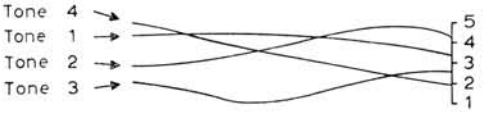


*Fig. 5.* Basic curves of Mandarin tones: citation syllables in pre-neutral position (i.e., average curves of syllable type 4, with initial voiceless fricative, *xu*, *hu*, *shi*, *fa*). The overall range of these curves is divided into four equal intervals of pitch separated by five points, indicated on the vertical scale.

*Table II.* Mean duration (in msec) of the voiced portion of 136 syllables, by tone and types of syllable structure

| | Tone 1 | Tone 2 | Tone 3 | Tone 4 |
|---|---|---|---|---|
| Types 1, 6, 7 | 187 | 211 | 228 | 218 |
| Types 2, 3, 5 | 301 | 334 | 347 | 320 |
| Types 4, 8, 9 | 186 | 215 | 231 | 198 |

rises gently at first. Tone 2 begins rising sooner with aspirated stops than with unaspirated stops. Tone 3 begins dipping sooner with aspirated stops than with unaspirated stops. And tone 4 begins falling much more abruptly with aspirated stops than with unaspirated stops[7].

The most striking variations in shape that may be observed in the average curves are the variations between types 2, 3, and 5, with initial voiced consonants or non-syllabic vowels, and all the other types. If the curves of types 2, 3, and 5 are compared tone for tone with the other average curves, they will appear to have humps or shoulders near the beginning. Moreover, the last part of the curves, following the turning point of the hump or shoulder, will appear to be similar in shape to the entire curve of the same tone in all the other types.

In making comparisons of average curves among the types, it must be borne in mind that the differences in actual duration of the voiced part of the syllable have not been preserved in these graphs. The mean durations of the syllables represented in each graph of figures 1–4 are given in table II. If the mean durations of the voiced part of the syllable are compared, it will be found that for every tone the mean duration of types 2, 3, and 5 is from 50 to 60% greater than the mean durations of the other types, while the mean durations of the other types are nearly identical for each tone. The significance of these comparative durations will be discussed below.

[7] These variations were apparently noted by SHEN *et al.* [12] in their spectrograms of tones 2 and 3. 'When a spectrogram began with a stop, there was a little cue indicating the type of stop at the beginning of the pitch curve. That cue was considered the consonantal cue and was eliminated in the measurements.' Comparable observations in the perceived pitch patterns of English monosyllables were made by REEDS and WANG [11], in a tape-cutting experiment with words like (1) *bat*, (2) *pat*, and (3) *spat*, where the initial [s] had been removed from each of the set (3) words. 'One of the authors, W. S.-Y. W., observed a relatively consistent cue which distinguishes the set (1) words from the truncated set (3) words in that the pitch drop appears to be more immediate and abrupt in the set (3) words.'

### *Citation Tones in Pre-Neutral Position*

The endpoints of all the average curves are no doubt influenced by the presence of the following syllable of the carrier sentence. Because the citation syllable and the unstressed syllable following it really amounted to a syntactic word as spoken in the carrier sentence, it seems reasonable to suppose that these citation tones before an evident neutral tone are typical examples of the tones in pre-neutral position. But the basic curves seen in figure 5 do not all correspond exactly either to CHAO's [4] contours before neutral tone or to his contours in isolation.

The basic curve of tone 4 (like all the average curves of tone 4) falls only to point 2, rather than to point 1 – as CHAO describes tone 4 both in isolation and before neutral tone. The same discrepancy is also seen in the results of another recent instrumental study of Mandarin tones. DREHER and LEE [8, p. 23] compared the normalized, simplified contours – derived from measurements of 18 Mandarin speakers (both men and women) – of tones spoken in isolation with those followed by a neutral tone, and they found that in pre-neutral position 'the tones are uniformly shorter (about 20%)....' In terms of the same five-point scale of equal intervals of pitch, DREHER and LEE's tone 4 (measured from the syllable *qù*) falls all the way to point 1 only in isolation and other final positions, and falls to point 2 before neutral tone (*le*) and before all four of the tones.

The basic curve of tone 3 (like all the average curves of tone 3) constitutes the most notable discrepancy with CHAO's [4] contours. Ending between points 2 and 3, it does not rise as high as CHAO's tone 3 in isolation, nor does it simply fall without rising (the 'half 3rd tone') as does CHAO's tone 3 before neutral tone. This discrepancy may be accounted for by assuming that the basic tone 3 curve represents an additional contextual variant of tone 3 which is different from either of those described by CHAO. The comparable shapes of all the tone 3 average curves, seen in figure 3, as well as the pre-neutral contour presented by DREHER and LEE [8], call into question whether the half 3rd tone occurs before a neutral tone. (On the other hand, WANG and LI [14] present spectrograms which show pre-neutral tone 3 with no rise, in the sequences *yǒu de* and *zǐ de*.)

In comparing their normalized, simplified contours of the tones in pre-neutral position with those in isolation, DREHER and LEE [8, p. 23]

found 'that while the second part of tone 3 is sharply curtailed in duration, it still has its characteristic dip'. DREHER and LEE's tone 3 (measured from the syllable *zǒu*) does have a final rising segment before neutral tone (*le*), but this rising segment is much shorter and rises much less – ending between points 1 and 2 – than that in isolation which rises to point 3, in terms of the same five-point scale of equal intervals of pitch. In view of this work, the traditional description of tone 3 before a neutral tone as a simple fall must be rejected, and the citation tones of the present study will have to be regarded as pre-neutral variants with contours of essentially the same shapes as those of tones in isolation.

### *A Uniform Set of Basic Curves*

The additional duration observed in types 2, 3, and 5 (table II) offers a basis for interpreting that portion of the average curves preceding the turning point of the hump or shoulder as an anticipatory movement of voice pitch occuring during the initial voiced consonant or non-syllabic vowel. Since the fundamental frequency measurements of these syllables were made beginning on the first voiced segment of the syllable, one may expect the first part of these curves to show the course of the fundamental during the segment preceding the syllabic vowel. Indeed, from an inspection of the standard 2,000 Hz/in spectrograms of these syllables, it is very clear that the turning point of the hump or shoulder coincides with the segment boundary between the initial voiced consonant or non-syllabic vowel and the syllabic vowel in all syllables of types 2, 3, and 5.

By disregarding the first part of these average curves, the part identified with the voiced consonant or non-syllabic vowel, one will eliminate roughly 30% of their durations. And, since the mean durations of the entire voiced part of the syllable of types 2, 3, and 5 are found to be about 50% greater than those of the other types, one will see that the remaining portions of these curves correspond very well, tone for tone, with the other curves – both in shape and in mean duration. With only minor exceptions, then, a uniform set of basic curves will emerge for the four tones in all syllable types, allowing the excluded portions in types 2, 3, and 5 to be reasonably viewed as resulting simply from automatic excursions of the fundamental that anticipate the starting points of the basic curves. Such a uniform set

will apparently have much the same shapes as the curves shown in figure 5 for syllables of type 4, which have already been considered 'basic' in an earlier section of this paper. If these basic curves can be accepted as the characteristic shapes of the tonal contours of Mandarin, it must follow that the domain of tone in Mandarin does not include initial voiced consonants or non-syllabic vowels.

### *On the Shape of Pre-Neutral Tone 3*

The results of this study (especially the similarity in shape among all the average curves of tone 3), together with the work of DREHER and LEE [8], support the assumption that Mandarin citation tone 3 has different contours in isolation (214), in pre-neutral position (212), and in other non-final positions (21). If this is true, one may conclude that CHAO's [4] half 3rd tone (21) occurs before tones 1, 2, and 4, but not before a neutral tone, and that the contour of Mandarin pre-neutral tone 3 has a shape that is closer to that of tone 3 in isolation.

### *On the Domain of Tone in Mandarin*

The interpretation of the first part of the average curves in types 2, 3, and 5 as merely an anticipatory adjustment of the voice suggests that the domain of tone in Mandarin is not the entire voiced part of the syllable or the entire vocalic part of the syllable, as it has been described [3, 6, 7, 9, 10, 12, 13], but rather is confined to the syllabic vowel and any voiced segment that may follow it in the syllable, that is the *rhyming part* of the syllable.


### *Zusammenfassung*

**Über den Tonbereich im Mandarinischen. Akustische Beobachtungen**

Die Messungen der Grundfrequenz von 136 ad hoc gesprochenen Einsilbern ergaben Durchschnittskurven für die vier distinktiven Töne in neun Silbentypen. Es zeigt sich, daß die Grundkonturen der mandarinischen Töne nur vom silbischen Vokal bis zum folgenden stimmhaften Segment reichen. Diejenigen Teile der Tonhöhe, die einen stimmhaften Anfangskonsonanten oder einen unsilbischen Vokal begleiten, nehmen lediglich die folgende Stimmbewegung vorweg.

## *Résumé*

**Le domaine du ton en mandarin. Quelques observations acoustiques**

Les mesures de la fréquence fondamentale de 136 syllabes du mandarin ont donné des courbes moyennes pour les quatre tons distinctifs pour neuf types de syllabes. Il semble que les contours de base des tons du mandarin ne coïncident qu'avec la voyelle syllabique et n'importe quel segment sonore qui la suit, tandis que les portions de hauteur mélodique apparaissant avec une consonne initiale sonore ou avec une voyelle non syllabique ne sont que des ajustements anticipés de la voix.

## *References*

1 ABRAMSON, A.: The vowels and tones of standard Thai: acoustical measurements and experiments (Indiana University Research Center in Anthropology, Folklore, and Linguistics, Bloomington 1962).
2 BROTZMAN, R. L.: Progress report on Mandarin tone study. Project on linguistic analysis report No. 8, pp. 1–35 (Ohio State University Research Foundation, Columbus 1964).
3 CHAO, Y. R.: A grammar of spoken Chinese (University of California Press, Berkeley 1968).
4 CHAO, Y. R.: Mandarin primer (Harvard University Press, Cambridge 1948).
5 CHAO, Y. R. and YANG, L. S.: Concise dictionary of spoken Chinese (Harvard University Press, Cambridge 1947).
6 CHENG, R. L.: Mandarin phonological structure. J. Ling. *2:* 135–158 (1966).
7 DOW, F. D. M.: An outline of Mandarin phonetics (Australian National University Press, Canberra 1972).
8 DREHER, J. J. and LEE, P. C.: Instrumental investigation of single and paired Mandarin tonemes (Douglas Advanced Research Laboratory, Huntington Beach 1966).
9 KRATOCHVIL, P.: The Chinese language today: features of an emerging standard (Hutchinson, London 1968).
10 KRATOCHVIL, P.: Review of a grammar of spoken Chinese by Y. R. CHAO. Language *46:* 513–524 (1970).
11 REEDS, J. A. and WANG, W. S.-Y.: The perception of stops after *s*. Phonetica *6:* 78–81 (1961).
12 SHEN, Y.; CHAO, J. C. Y., and PETERSON, G.: Some spectrographic light on Mandarin tone-2 and tone-3. Study of sounds (Phon. Soc. Japan) *9:* 265–314 (1961).
13 WANG, W. S.-Y.: Phonological features of tone. Int. J. Amer. Ling. *33:* 93–105 (1967).
14 WANG, W. S.-Y. and LI, K.-P.: Tone 3 in Pekinese. J. Speech and Hearing Res. *10:* 629–636 (1967).
15 WANG, W. S.-Y.; LI, K.-P., and BROTZMAN, R. L.: Research on Mandarin phonology. Project on linguistic analysis report No. 6r (Ohio State University Research Foundation, Columbus 1963).

Request reprints from: Dr. JOHN M. HOWIE, Assistant Professor of French and Linguistics, 27 Arts and Science Bldg., University of Missouri, *Columbia, MO 65201* (USA)