Title
WPP, No. 36: Studies on Tone

Publication Date
1977-07-01

STUDIES ON TONE

july 1977

UCLA
WORKING PAPERS
IN PHONETICS 36

The papers presented in this volume have been discussed at the UCLA tone seminar during the academic year 1976-1977. The seminar was organized by Jean-Marie Hombert with the participation of Saed Ali, Steve Anderson, Ava Berinstein, Cathe Browman, Sandy Disner, Victoria Fromkin, Louis Goldstein, Steve Greenberg, Peter Ladefoged, Carl La Velle, Ian Maddieson and Eric Zee.

This work was funded by the National Science Foundation (NSF SOC75-07158).

As on previous occasions, the material which is presented here is simply a record for our own use, a report as required by the funding agencies, and a preliminary account of work in progress.

Correspondence concerning Working Papers in Phonetics should be addressed to:

Department of Linguistics
c/o Secretary of the Phonetics Laboratory
UCLA
Los Angeles, California 90024

Table of contents

Abstract

A note on tone and the abstractness controversy

Victoria A. Fromkin

[Published in *Studies in African Linguistics*, Supplement 6, November 1976. 47-62]

Kiparsky's 1968 paper (1973) concerning the nature of lexical representation opened up a debate among generative phonologists which continues to this day. The controversy is not between those who advocate a totally 'abstract' representation and those who accept Postal's (1968) 'naturalness condition' but among those who are united in placing 'psychological reality' as a primary condition for grammars. All those involved are therefore concerned with finding just that set of constraints on grammars which would limit the class of grammars to those which are learnable and which would still reveal 'linguistically significant generalizations.'

The strongest constraints on lexical representation have been posited by adherents of the theory known as Natural Generative Grammar (NGG) (Hooper 1976). In this theory, a Strong Naturalness Condition (SNC) is included which "requires that all lexical forms be entered in systematic phonetic representation as though the phonological rules had already applied."

Tone languages present a problem for NGG and for any theory which suggests that lexical representation is identical to phonetic representation since tone is a phonological phenomenon, phonetically manifested as pitch and even at the phonetic level it is relative rather than

absolute. Hausa, a tone language with 'downdrift' in which there is a downward contour of pitch manifested throughout a sentence illustrates the problem. In a long utterance (and even a long word) a High tone at the end may be lower in pitch than a Low tone at the beginning. Furthermore phonological rules make use of abstract tonal features e.g. [+High], rather than the phonetic correlates. Since identical tones in polysyllabic morphemes (or words) with sequences of high and low tones would be phonetically different in pitch, if the lexical representation is identical with the phonetic representation phonologically 'identical' tones would have to be represented by different phonetic features or pitch values. Akan represents a similar problem. The SNC thus appears to be too strong a condition on lexical representation.

The lexical representation of contour tones also relates to the abstractness controversy. Strong evidence has been provided in support of representing the contour tones of at least some languages as sequences of level tones, or as deriving from level tones. In this paper a number of examples are provided to show that the SNC is too strong a constraint in such cases. Non-alternating contour tones would have to be represented by contour features in the lexical representation, but this would obscure the fact that these tones 'behave' like level tones. It is shown that the lexical representation of contour tones depends not on the phonetic realizations but on the phonological functioning of the tones in the language system. The differences between languages with underlying (lexical) and derived contours would be obscured by a theory which allows only a phonetic level of representation.

Kiparsky, Paul. (1973) Phonological representations. In Osamu Fujimura (ed.) *Three Dimensions of Linguistic Theory*. 1-136. Tokyo: TEC.

Hooper, Joan. (1976) *Introduction to Natural Generative Phonology*. Academic Press.

Abstract

On the interaction between tone and vowel length:
Evidence from Thai dialects

Jack Gandour

[*Phonetica* (1977) V. 24, 54-65]

The loss of a phonological distinction in vowel length historically among certain Thai dialects (spoken in Thailand) may be seen to be principally conditioned by tones. Available auditory, acoustic and other experimental data on the interaction between tone and vowel (syllable or tone bearing unit) duration evidenced in several geographically and typologically distinct tone languages point to universal phonetic tendency having to do with the production of tone. These phonetic tendencies are believed to have caused the changes in vowel duration thus bringing about a change in the phonological status of vowel length - a restructuring of the synchronic phonological system among dialects and languages of the Tai language family. It is further suggested that the use of indivisible unit-contour tone features leads to an economic, explanatory diachronic and synchronic description of these phenomena.

Abstract

A cross-language study of tone perception

Jack Gandour

[Bell Laboratories]

Richard Harshman

[University of Western Ontario, Canada]

The construction of a universal linguistic-phonetic theory raises questions regarding the nature of phonetic properties or features and the extent to which these features are utilized across particular languages that differ phonetically and/or phonologically. This paper reports on an investigation of the perceptual nature of tone features and the extent to which these features are utilized in 2 typologically and phonologically distinct 'tone' languages - Thai and Yoruba as compared to 1 'non-tone' language - English.

The experimental task involved paired-comparison judgments of dissimilarity for 13 synthetic speech-like stimuli that differed with respect to the level,rate-of-change, direction-of-change of fundamental frequency as well as duration. The method of analysis was based on an individual differences model for multidimensional scaling in which individuals are assumed differentially to weight the several dimensions of a common psychological space. The multidimensional scaling analysis for the pooled English-Thai-Yoruba subjects results in a 5-dimensional tone space that has been interpreted as (1) average pitch, (2) direction, (3) length, (4) extreme endpoint and (5) slope.

The differential weighting of these dimensions or features for a

given individual reveals that the perception of tone, to a large extent, depends on the linguistic (both phonetic and phonological) status of tone in his native language. Thus, the 'direction' and 'slope' dimensions are found to be more heavily weighted by the Thai and Yoruba subjects than by the English subjects. The 'slope' dimension, in turn, is most heavily weighted by the Thai subjects - which suggests that the weight or 'saliency' of a given dimension may be influenced more by its role on the underlying phonological level than the surface phonetic level. Thai, unlike Yoruba, exhibits contour tones in its phonemic inventory although Yoruba does exhibit derived surface phonetic contour tones. The results of a 1-way analysis of variance for each of the 5 dimensions and a stepwise discriminant analysis of the language groups further point to the linguistic-phonetic nature of these perceptual dimensions.

Earlier proposed phonological features of tone [$\pm$ CONTOUR], [$\pm$ RISING] and [$\pm$ FALLING] find support in a regression analysis into the 5-dimensional perceptual tone space.

Abstract

Perception of tones of bisyllabic nouns in Yoruba

Jean-Marie Hombert

[Published in Papers in African Linguistics in Honor of William E. Welmers. *Studies in African Linguistics*, supp 6, 109-121]

The purpose of the first experiment was to investigate the perceptual cues used by Yoruba speakers to distinguish among the six tone patterns found in bisyllabic nouns (i.e. [_ _\\], [_ —], [_ _/], [— _\\], [— —] and [— ¯]). Eight Yoruba speakers were asked to evaluate the perceptual distance (on a scale from 0 to 10) between tone patterns of V_1CV_2 nouns presented by pairs. Subject's responses were analysed using multidimensional scaling techniques. Results from these analyses and from regression analysis indicate that Yoruba speakers are mainly using two perceptual factors to distinguish these 6 tone patterns. The first factor is mainly associated with the direction of change of Fo of V2 and also to some extent to the averaged Fo value of V2. The second factor can be analyzed as resulting from the combination of three cues: amount of Fo change of V2, averaged Fo value of V1 and Fo difference between onset of V2 and offset of V1. The second experiment was designed to investigate the relative importance of primary and secondary cues used to distinguish a mid tone from a low tone in word final position. The results of this second experiment show that falling Fo is the main percpetual cue used to identify low tones in word final position. Although production data indicate a strong correlation between Fo and amplitude, amplitude information does not appear to be a major cue used by Yoruba speakers to distinguish between low and mid tones in word final position.

Abstract

Noun classes and tone in Ngie

Jean-Marie Hombert

[Published in *Studies in Bantu Tonology - Southern California Occasional Papers in Linguistics 3*. Edited by L. Hyman, 3-21.]

This paper presents the Ngie (spoken in the Grasslands area of Cameroon) noun classes and tone system. Most Ngie nouns have a V_1-C_1(G)V_2(C_2) shape. The first vowel (i.e. the prefix) can either be i-, a-, or u- and can have either low or mid tone. The second vowel can have one of the following six tone shapes: low with downgliding before pause L [_\ ], low level L° [_], mid level [—], high H [‾], falling F [\ ] and rising R [_/]. Out of the twelve theoretical tonal combinations, only two patterns are lacking: mid followed by level low and mid followed by rising tone. Comparison with Proto-Bantu reconstructions suggest the following correspondences: a) L-L, L-F and M-F Ngie tone classes correspond to an earlier L-LL sequence; b) L-L° and M-M Ngie tone classes correspond to an earlier L-LH sequence; c) L-M and M-H in Ngie correspond to an earlier L-HL sequence. The most likely historical scenario for these developments seems to be the following: a) all prefixes were low; b) in certain contexts, the tone of the prefix was changed from low to mid except in cases in which the prefix was followed by a NC consonant cluster; c) nasals were lost when followed by a voiceless segment and they merged with the following segment when the latter was also a (homorganic) nasal. It is also pointed out that the modern tone patterns correspond more or less to those of the <u>first two</u> syllables (i.e. the prefix and the first syllable of the stem) of the reconstructed patterns of Proto-Bantu forms.

Abstract of M.A. thesis

An instrumental investigation of Yoruba tone

Carl Richard La Velle

[University of California, Los Angeles, 1977]

This thesis reports on an acoustic study of Yoruba tones which is aimed at providing information as to how phonological tone is realized phonetically. Throughout this study it is assumed that the interaction of tones at the phonetic level can be explained, to a large extent, as resulting from the influence of coarticulations between neighboring tones. Both low-raising and downdrift have been argued to result from unavoidable tonal interaction at the phonetic level. Although the raw data reveal some evidence of both these processes, in the case of downdrift I only found evidence of a tendency. I conclude that low-raising reflects a natural phonological strategy that is motivated by structural and physiological constraints on tonetic space. As a result of multiple regression and correlation studies within utterance types, I found a high percentage of significant correlation coefficients across tonetic contexts. From a comparison of correlations of tone pairs within each utterance type, none showed a significant difference. Finally, I found that the correlation of a tone to another tone outside of its immediate context did not vary significantly with alternations in that context. The results of my correlation studies, I maintain argue against the likelihood that proximity even minimally affects the association of one tone to another in the same utterance. Additionally, these findings suggest that for purposes of speech production, the individual tones are organized into a unit at least as large as the phrase.

Abstract

Tone reversal in Ciluba - a new theory

Ian Maddieson

[Published in *Studies in Bantu Tonology* (Southern California Occasional Publications in Linguistics #3) pp.141-166. 1976. (Edited by L.M. Hyman)]

This paper reviews the problem of the historically reversed tones of Ciluba and proposes a new explanation. Coupez, following Hulstaert, observed that high tone in Ciluba corresponds with low tone in Proto-Bantu and low with high. A straightforward reversal is an implausible historical rule. Van Spaandonck sought a more natural explanation, suggesting that tones were reversed in only a few items and these could be explained by a series of rules of tone copying, displacement, etc. By comparing the Ciluba reflexes of Guthrie's Common Bantu forms the present paper shows that the reversal is pervasive. This is first shown for nouns and noun prefixes. If an initial high tone syllable is assumed to have occurred before the noun prefix the reversed tones can be derived by a simple rule of tone-spreading, followed by loss of the initial syllable and reassignment of the tone pattern to the remaining syllables. Recent research shows that the most closely related languages to the Luba cluster are those in the Ila-Bemba group (Heine 1973, Henrici 1973). In languages of this group such as Uwulamba and Cibemba the 'preprefix', originally a demonstrative of rather limited occurrence in Proto-Bantu, occurs very frequently with nouns; in Ciluba no trace of the preprefix is found. It is therefore proposed that the initial high tone syllable assumed for pre-Ciluba nouns, and subsequently lost, was the preprefix.

Verb tones are also reversed from Proto-Bantu. Infinitives display a subset of the patterns found with nouns and the same explanation applies. But to explain reversed tones in finite verbs it is necessary to assume an initial low tone syllable. A preprefix occurs with relative verbal forms in Cibemba and Uwulamba. Ciyao provides evidence that it had an original low tone. It is suggested that this preprefix became widespread with verbs in pre-Ciluba, tone-spreading and contraction also applying to these forms. The result of these processes is a language with reversed tones that shows no overt trace of the preprefix in verbs or nouns.

Abstract

Universals of tone

Ian Maddieson

[to be published in *Universals of Human Language* Vol. 2 edited by J. M. Greenberg et al., Stanford University Press, 1977]

Three areas of phonological universals of tone are considered: inventories of tone, tone rules, and interaction of tone with nontonal features. It is proposed that five is the maximum number of tone levels that may contrast, and these five levels represent a phonetic scale. Languages with less than five levels use a smaller portion of the scale, with their marked tones generally drawn from the upper part. Contour tones do not occur unless at least one level tone also occurs in the inventory, and bidirectional contours only occur if at least one simple contour also occurs. These constraints on contour tones are related to parallel constraints on sequences of level tones. Tonal assimilation and related processes of displacement and contraction are discussed and some asymmetries in their operation are explained by the usually marked status of high tones. The unusually frequent occurrence of rules of tonal polarity is pointed out. The final section comments on phonetic and phonological relationships of tones with adjacent consonants and with the quality, duration and phonation type of the vowels on which tones are manifested.

(This work was also supported by an NSF grant to Stanford University for research on language universals.)

Difficulty of producing different Fo in speech

Jean-Marie Hombert

[Department of Linguistics
University of California, Los Angeles]

1. Introduction.

The notion of articulatory difficulty is often used but rarely quantified in the linguistic literature. Such data are extremely valuable when explaining the naturalness of phonological processes. This study investigates the difficulty associated with the production of different fundamental frequencies in speech.

Fundamental frequency (i.e., the rate of vibration of the vocal cords) is controlled by aerodynamic and muscular forces. Although there is still some disagreement about the relative weight of these two sets of forces (Liebermann 1967, 1970; Ohala 1970, 1973, in press) it seems that muscular tension is the primary factor controlling Fo changes. Electromyographic (EMG) recordings allow us to evaluate the timing and to some extent the relative level of activity of different muscles. But these data are extremely difficult to relate to the notion of articulatory difficulty. The main problem is that, assuming that all the relevant muscles have been sampled, it is not possible to associate an "overall muscular activity" to a given Fo change by adding together the EMG activity detected in each muscle because the amplitude of each recording is a function of electrode placement and also because we have no reason to believe that equal EMG activity in two different muscles correspond to equal physiological effort and consequently to directly summable articulatory difficulty. Furthermore, even if we were able to evaluate the muscular effort involved in a given tone, we would still have to attribute relative weights to the pulmonic effort (i.e., aerodynamic forces) vis a vis the muscular effort in order to get a global measurement of articulatory difficulty.

Assuming there is a direct correlation between the difficulty involved in doing a certain task and the level of accuracy reached when doing this task, it is possible to by-pass the problems mentioned earlier and get an evaluation of the aritculatory difficulty involved in producing certain tones. Let us consider an experiment in which subjects are asked to perform Fo changes in which the beginning and the end point of the Fo contour is given. Subjects' accuracy at imitating the target tone can be considered as being related to the articulatory difficulty involved in producing this target tone. Ohala and Ewan (1973) and Sundberg (1973) ran a somewhat similar experiment although they were mainly interested in maximum speed of Fo change and consequently ask their subjects to perform their Fo changes as fast as they could between the onset and the offset frequencies (two pure tones representing these frequencies were provided to the subject just before his own production). Their results showed that naive subjects (i.e., non-singers) perform

Fo drop faster than Fo rise over the same frequency range. This type of experimental paradigm has two main shortcomings if one wants to relate these results with the articulatory difficulty associated with various tone shapes. Firstly, it is not clear that the results obtained with maximum speed of Fo change would still hold with slower rate of Fo change more commonly found in natural speech. Secondly it would be better to use a paradigm in which the frequency range of the target tone is adapted to the frequency range of the subject in order to avoid creating unwanted articulatory constraints caused by the subject's attempt at speaking outside her/his normal Fo range.

2. Experimental paradigm.

In order to remedy these shortcomings the following experiment was carried out: Five American male speakers were asked to produce the following nine tone patterns:

- -3 level tones: low level, mid level and high level
- -3 rising tones: from low to mid, from mid to high, and from low to high
- -3 falling tones: from high to mid, from mid to low, and from high to low

The stimuli were presented visually in a word list in which each token was represented using Chao's tone numbers (1930). (i.e., 11, 33, 55, 13, 35, 15, 53, 31, 51). There were five randomized repetitions of each tone with the vowel [i] and five repetitions with the vowel [a]. The stimuli were presented visually in order to avoid influencing the subjects' production by providing them with an acoustic signal of given duration and amplitude. Two vowels [i] and [a] were used in order to see whether certain tone patterns were produced more accurately with high vowels than with low vowels or vice-versa. Fundamental frequency measurements were made on a PDP-12 computer using the CEPSTRUM method on the waveform digitized at 10 KHz. In order to get comparable Fo and amplitude values from tokens of different durations, each token was divided into five intervals, the average Fo and the average amplitude were computed for each interval.

3. Results.

The results from the three level tones: low level (11), mid level (33), high level (55) are presented in Figure 1. The three upper curves represent the average fundamental frequency patterns from the five subjects. The numbers indicated at the onset and offset of each tone refer to the subject number. Each curve is composed by five points representing the five intervals mentioned earlier. Each one of these points is the average of 50 measurements (5 subjects x 10 repetitions (5 repetitions of [i] and 5 repetitions of [a])). The three lower curves represent the averaged intensity curves for these three Fo patterns. Three facts should be noticed from Figure 1: (1) the low tone is realized as a low falling tone, (2) the frequency spacing between low and mid is roughly equal to the spacing between mid and high, (3) there is a

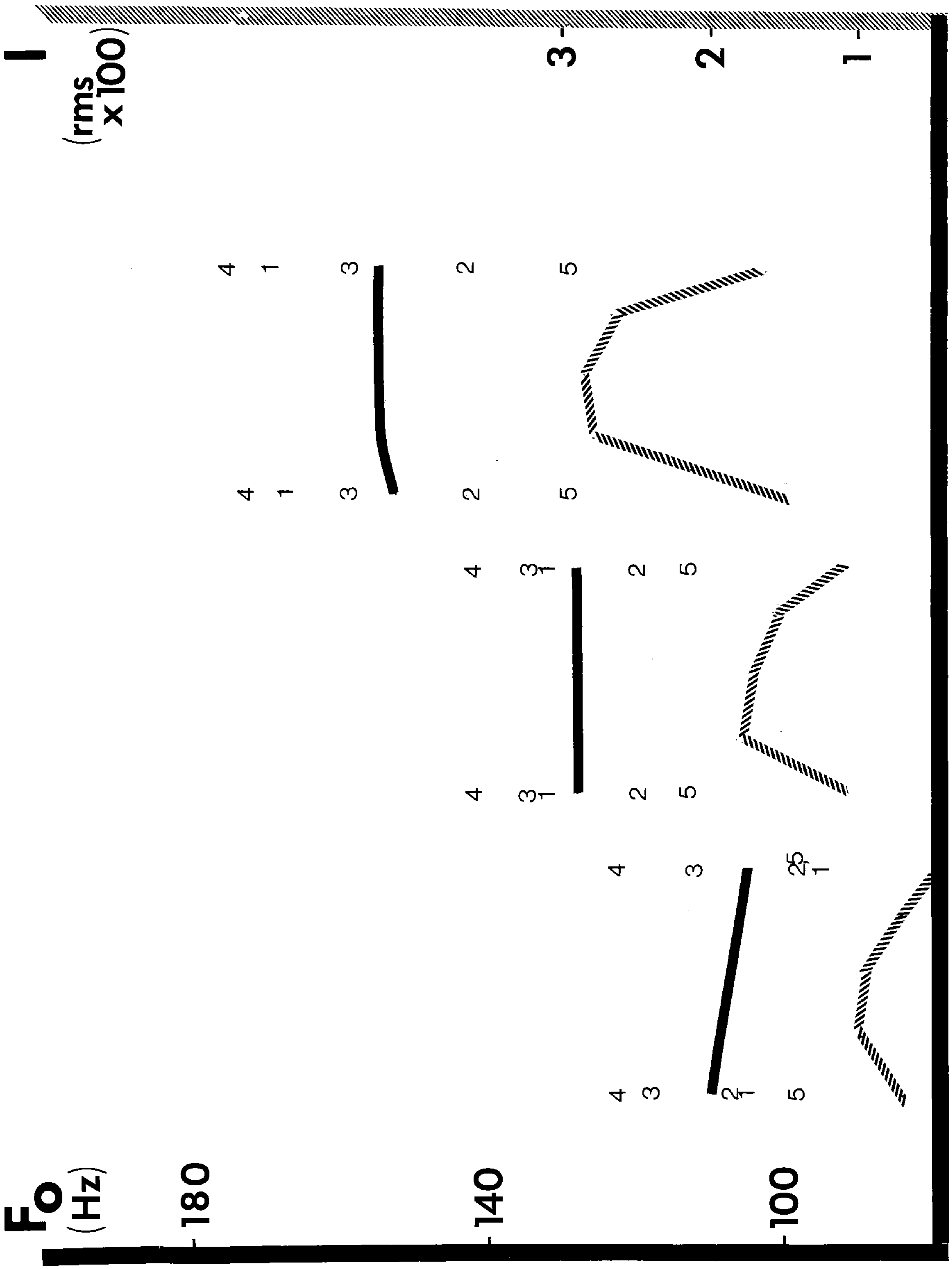


Fig. 1. Fo and amplitude values for the three level tones.

good correlation between amplitude levels and Fo levels.

Figure 2 shows the three rising tones; as can be seen:

(1) The low to mid rising tone (13) covers a smaller frequency range than the mid to high contour (35) (8.8 Hz vs. 33.6 Hz)

(2) the magnitude of the slope of 13 and 15 (i.e., the two rising tones starting at low level) is small during the first half of these tones.

(3) Here also, there is a good correlation between amplitude values and frequency values. (Notice for instance that the third interval has a greater intensity value than the second interval).

As far as the falling tones are concerned (Figure 3), we notice that:

(1) The falling tone from high to mid (53) covers a smaller range than the falling tone from mid to low (31) (20.7 Hz. vs. 29.9 Hz).

(2) The beginning point of the 53 and 51 tones (i.e., the two falling tones which are starting high) have a starting point higher than the high level tone (55).

(3) The magnitude of the slope at the onset of the 53 and 51 tones and at the offset of the 53 tone is small.

(4) There is again a good correlation between amplitude values and frequency values.

A direct correlation was obtained between frequency height and tone duration (i.e., the lower the tone, the shorter the duration). This correlation was found to be significant for three subjects ($P<0.05$); the same tendency was observed for the fourth subject ($P<0.1$) but not for the last subject. For the rising tones, the following durational hierarchy was found for all five subjects: duration of 13 < duration of 35 < duration of 15. Notice that this hierarchy corresponds to the hierarchy based on the frequency range covered by these tones.

31 is the shortest falling tone; no significant differences in duration were found between 53 and 51. It is interesting to point out that falling tones are always shorter than rising tones over comparable frequency range. I mentioned earlier that 2 vowels ([i] and [a]) were used in this experiment. Since no interesting differences were noticed between contour tones produced with these 2 vowels, they were averaged together. However, the intrinsic Fo differences between [i] and [a] exhibit an interesting pattern when produced with the three level tones. These differences are presented in Table 1.

Table 1. Intrinsic Fo differences (in Hz) between [i] and [a]

Low tone	.7	nonsignificant
Mid tone	7.9	$P < .01$
High tone	12.3	$P < .01$

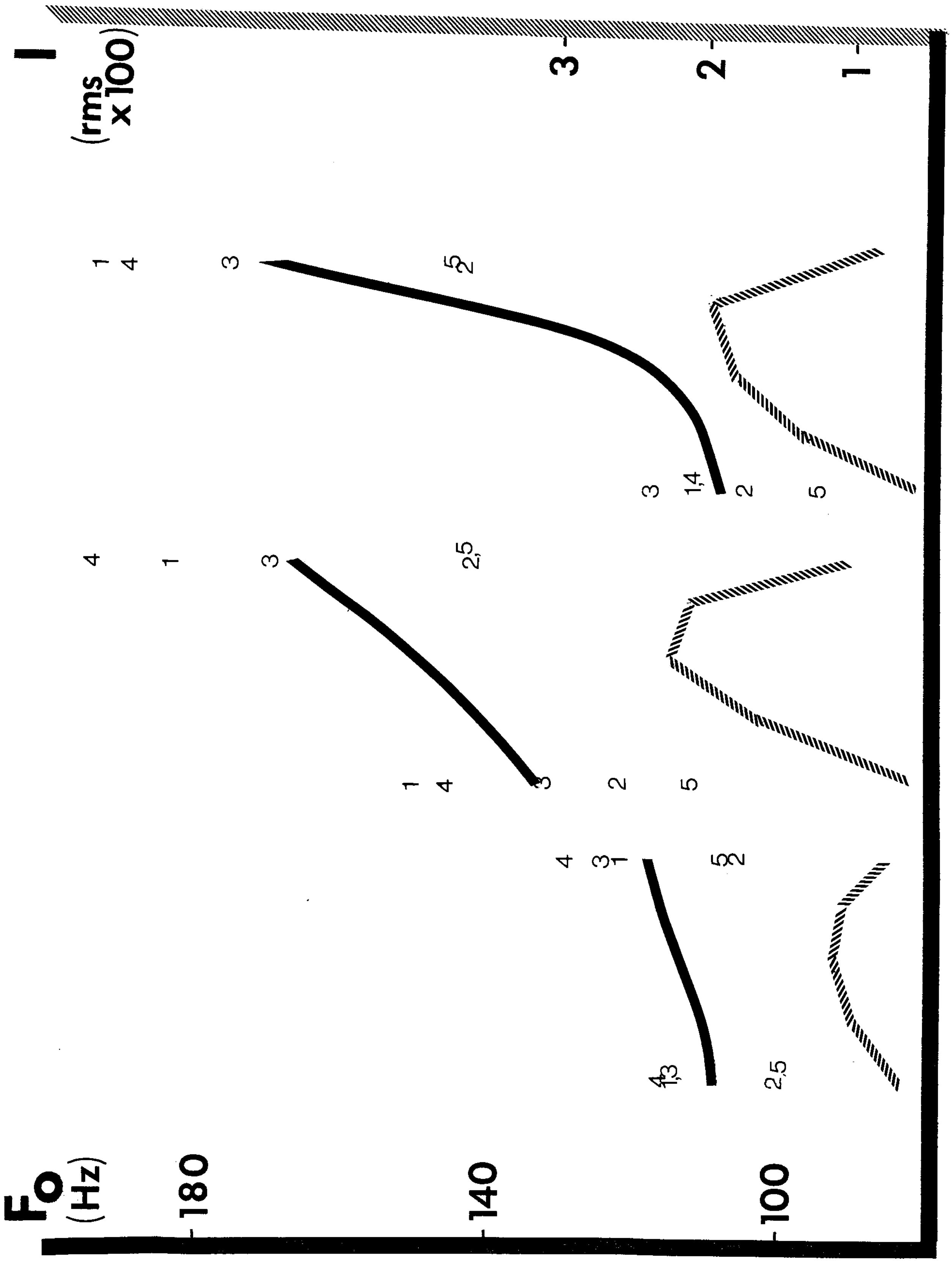


Fig. 2. Fo and amplitude values for the three rising tones.

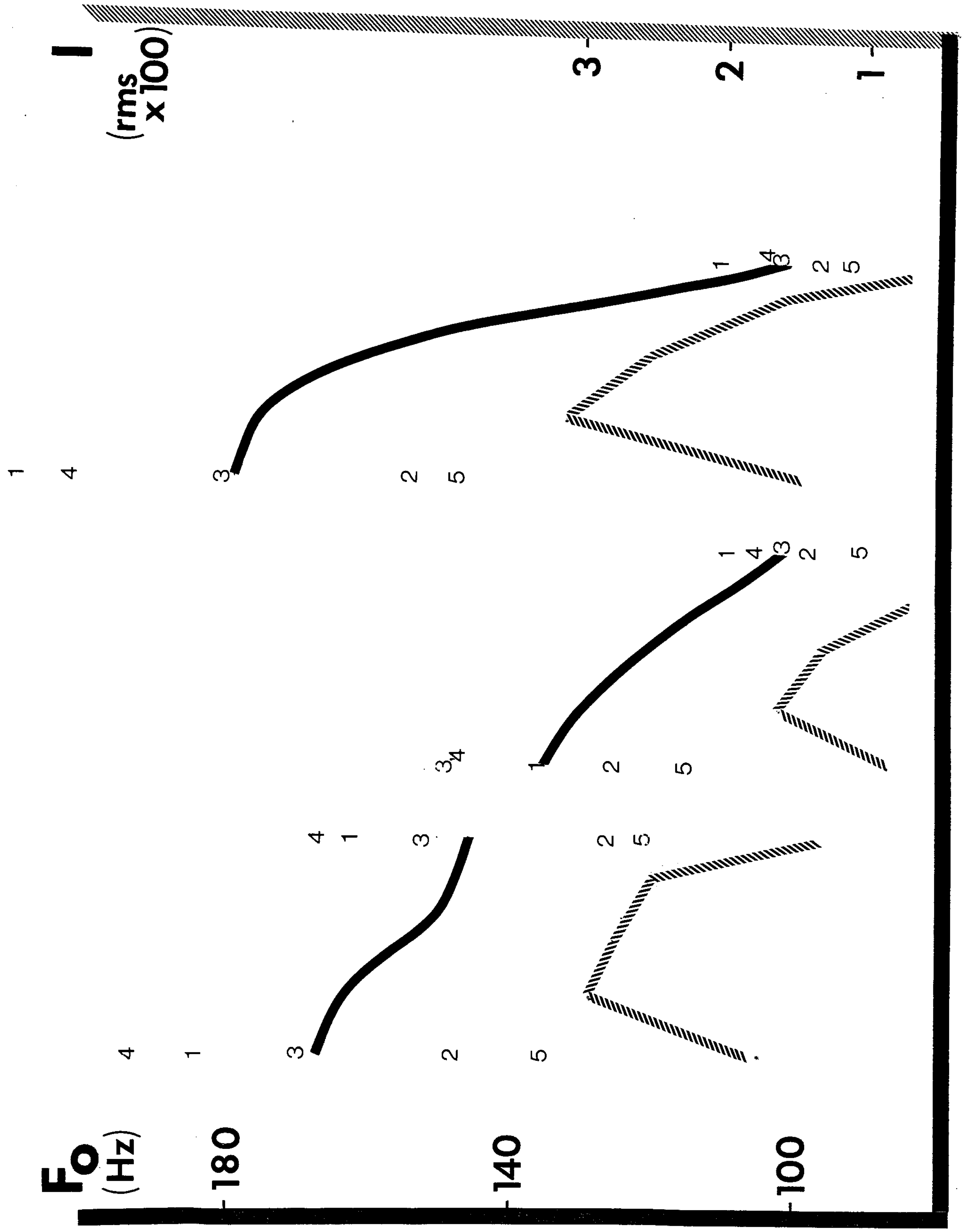


Fig. 3. Fo and amplitude values for the three falling tones.

4. Discussion

The duration data show that even for Fo change realized at a speed used in normal speech, the mechanism controlling Fo lowering is faster than the mechanism controlling Fo rising. As far as level tones are concerned, it is not clear why Fo height should be directly correlated to vowel duration.[1] The Fo data indicate that it is more difficult to control contour tones going from one extremity of the range towards the middle of the range than contour tones going from either the center of the Fo range towards one extremity or from one extremity to the other. The consistently good correlation obtained between amplitude and frequency values suggest that aerodynamic factors play a significant role in controlling Fo in natural speech. The fact that intrinsic Fo differences between high and low vowels are greater at higher Fo confirms the results of previous studies obtained from other languages such as Danish (Petersen 1976), French (Di Cristo and Chafcouloff 1976) and Yoruba (Hombert 1977).[2]

5. Conclusion.

In conclusion, these data are too preliminary to allow an accurate quantification of the articulatory difficulty associated with each tone shape. As a first approximation, it is proposed that contour tones covering a small Fo range (e.g., one stop in Chao's scale) represent a greater articulatory complexity than contour tones covering a wider frequency range. This claim is based on the observation that Fo changes are relatively slow during the initial part of the vowel. However, this constraint should not include contour tones ending at the extremity of the frequency range (i.e., 1 or 5).

Footnotes.

1. Pike (1974) claims that the opposite hierarchy is found in tone languages (i.e., low tones have a longer duration than mid tones which in turn have a longer duration than high tones).

2. Since it was found that in the case of Danish and French, subjects with higher Fo exhibited greater intrinsic differences between high and low vowels, these data suggest an acoustic rather than an articulatory explanation.

References.

Chao, Y.R. 1930. A system of tone letters. *Le Maitre Phonetique* (3), 45:24-27.

Di Cristo, A. and Chafcouloff, M. 1976. An acoustic investigation of microprosodic effects in French vowels. Paper presented at the 14th Conference on Acoustics, High Tatra, Czechoslovakia.

Hombert, J.M. 1977. Consonant types, vowel height and tone in Yoruba *Studies in African Linguistics* 8.2.

Lieberman, P. 1967. *Intonation, perception and language* MIT Press, 38.

Lieberman, P. 1970. A study of prosodic features. *Status Reports on Speech Research.* Haskins Lab., SR 23, 179-208.

Ohala, J.J. 1970. Aspects of the control and production of speech. *Working Papers in Phonetics.* 15.

Ohala, J.J. 1973. The physiology of tone. In L. Hyman (Ed.), Consonant types and tone. *Southern California Occasional Papers in Linguistics* 1. Univ. of So. Calif. 1-14.

Ohala, J.J. (in press) The production of tone. In *Tone: A linguistic survey* V. Fromkin (ed.) Academic Press.

Ohala, J.J. and Ewan, W.G. 1973. Speed of pitch change. *Journal of the Acoustical Society of America* 53, 345.

Petersen, N.R. 1976. Intrinsic fundamental frequency of Danish vowels. *Annual Report of the Institute of Phonetics* University of Copenhagen, 1-27.

Pike, E. 1974. A multiple stress system vs. a tone system. *IJAL* 40, 169-175.

Sundberg, J. 1973. Data on maximum speed of pitch changes. *STL-QPSR* 4. 39-47.

A Model of Tone Systems

Jean-Marie Hombert

[Department of Linguistics
University of California, Los Angeles]

1. Introduction

1.1. The development of tone and tone space

The process by which linguistically contrastive tones can develop from the influence of neighboring segments is well attested and rather well understood (Brown 1965; Chang 1973, 1975; Haudricourt 1954, 1961; Hombert 1975, in press; Hombert, Ohala and Ewan 1976; Matisoff 1973; Mazaudon, 1977; Sarawit 1973). However, one of the questions which has been left unanswered by researchers working on tonogenesis is the following: do the shapes of the newly phonologized tones remain roughly similar to the fundamental frequency (Fo) shapes from which these tones originated (i.e. preserving the Fo perturbations due to the neighboring consonants which caused the tonal development) or, rather, do they evolve according to some more universal principles?

This paper represents an attempt to demonstrate that the tones are going to be reorganized in the tone space according to universal phonetic principles independently of the Fo shapes prior to the tonal development.

1.2. Articulatory difficulty vs. perceptual distinctiveness

The notion of minimum articulatory difficulty and maximum perceptual distance has been used for a long time to account for diachronic linguistic processes (Passy 1890; de Groot 1931; Martinet 1955). However these notions have rarely been used quantitatively. Lindblom and his colleagues (Lindblom and Sundberg 1969, 1971; Lindblom 1972, 1975; Liljencrants and Lindblom 1972) constructed a model predicting the distribution of vowels in the vowel space.[1] The purpose of this study is to present a comparable attempt with tone systems. The model described here predicts the phonetically optimal tone systems from universal phonetic considerations (both articulatory and perceptual) without reference to the origin and the history of each tone system.

This paper is divided into two parts: first, some general tendencies observed in tone systems concerning the number and the shape of these tones are presented. Second, the model is described and its output is compared with actual data from tone languages.

2. Inventory of tone systems

A rapid survey primarily based on languages presented in Ruhlen's *Guide to the Languages of the World* (1975) and also on data from the

Phonology Archive Project (PAP), Stanford University, indicates that about 30% of the languages of the world are tonal. This figure would go up to about 50% if we would consider a sample of languages in which each language family would be represented by a number of languages proportional to the actual number it contains. This discrepancy between the two figures comes from the fact that most of the languages spoken in areas of great linguistic diversity (e.g. Chinese and Southeast Asian languages, Papuan New Guinea languages and Northwestern Bantu languages) are tonal.

An approximate count of the distribution of tone systems as a function of the number of tones in each system is presented in Table 1.

TABLE 1. *Distribution of Tone Systems*

Systems with 2 tones:	30%
" " 3 " :	30%
" " 4 " :	15%
" " 5 " :	10%
" " 6 or more:	15%

In most cases (90%) two tone systems are constituted by high and low tones. High, mid and low tones represent the most common three tone system (65%), although high, low and falling systems are found (20%). Two types of four tone systems are found quite commonly: high, low, falling and rising (40) and high, mid, low and falling (30%). High, mid, low, falling and rising constitute the most common five tone system (60%) and, finally, the most common six tone system[2] (60%) is composed by two level tones (low and mid in most cases), two falling tones (high to mid and mid to low) and two rising tones (low to mid and mid to high). No clear patterns were found for seven and eight tone systems.

3. The model

3.1. Description of the model

The steps which should theoretically be involved are shown in Fig. 1. Considering the relatively small quantity of data available on physiological constraints involved in the production and perception of tones as well as the very limited number of acoustic descriptions of tone languages, it is clear that the model presented here should be regarded as a first attempt at predicting tone shapes from articulatory and perceptual considerations. The input to the model was constituted by the number of tones N in the system and by 25 tones representing all possible combinations of the two tone numbers (from 1 to 5) used to describe the beginning and the end of each tone (Chao 1930).[3] The model generated all possible systems of N tones selected among the 25 possible tones. N was varied from 2 to 8.

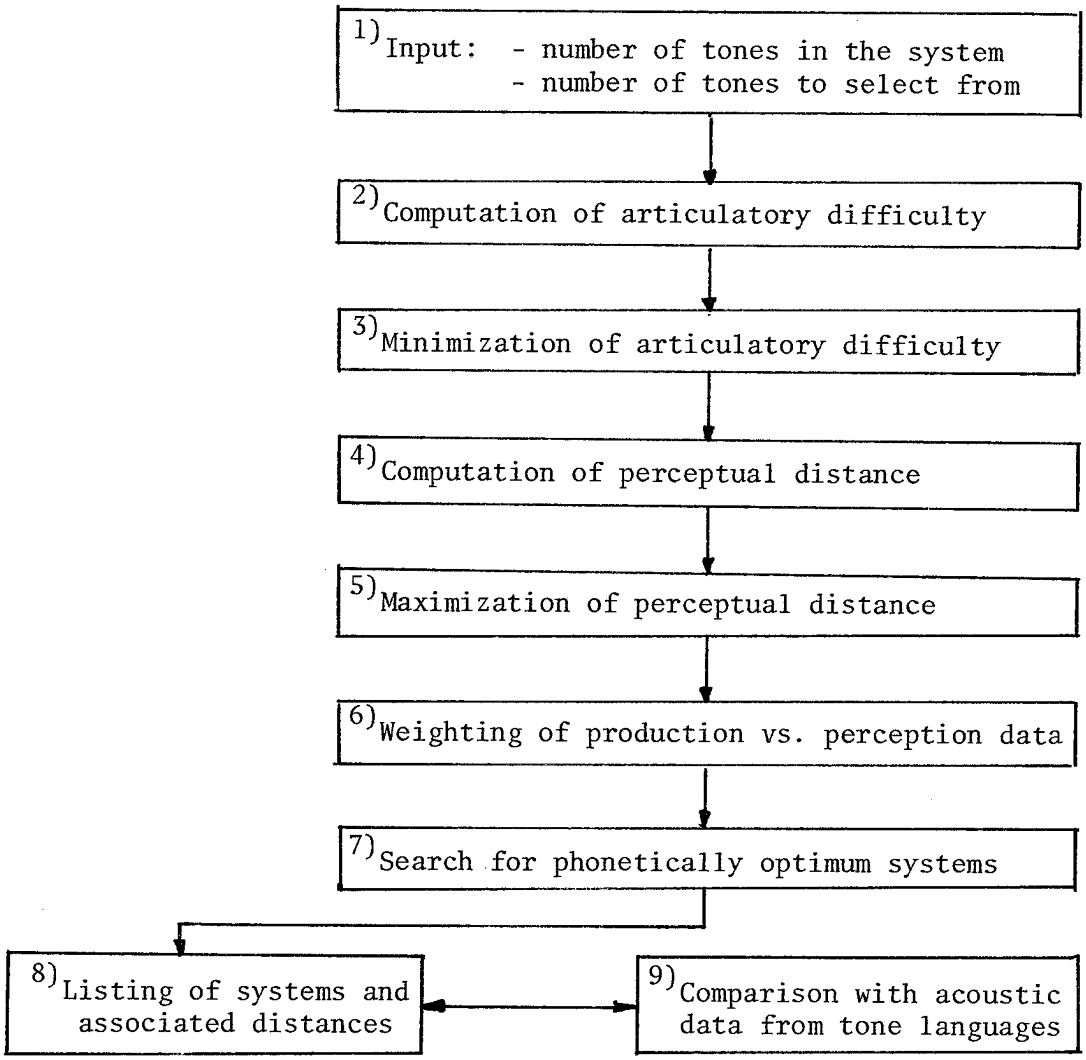


FIGURE 1. Flow-chart of a model predicting tone shapes from articulatory and perceptual considerations

3.2. Articulatory data

In Hombert (1977) a first attempt at quantifying the articulatory difficulty associated with certain tone shapes was made. It was found that contour tones covering a small F_o range (one step according to Chao's tone numbers) were probably articulatorily more complex than level tones[4] or contour tones ending at one extremity of the F_o range (i.e. 1 or 5). Considering the precariousness of these data, it was decided to simply discard the contour tones having a change of only one step (e.g. 12, 54, etc.) from the input set composed of 25 tones. However, contour tones ending at the extremity of the F_o range were retained even when their F_o change was only one step (i.e. 21 or 45). As a result, only 19 tones were considered as input to the perceptual component of the model:

5 level tones: 11, 22, 33, 44, 55
7 rising tones: 13, 14, 15, 24, 25, 35, 45
7 falling tones: 53, 52, 51, 42, 41, 31, 21

3.3. Tone perception

Data from experiments using pure tones (Brady et al 1961, Heinz et al 1968, Pollack 1968, Nabelek and Hirsh 1969, Nabelek et al 1970, Tsumura et al 1973) indicate that the end point of the tone plays a more important perceptual role than the beginning of the tone. In a recent study of tone perception involving English, Thai and Yoruba subjects, Gandour and Harshman (1977) used multidimensional scaling techniques to tease out the perceptual dimensions used by their subjects in making their similarity judgments on synthesized stimuli. Using PARAFAC analysis (Harshman 1970) they extracted five dimensions interpretively labeled (1) Average Pitch, (2) Direction, (3) Length, (4) Extreme endpoint, and (5) Slope. Of these five dimensions, Gandour and Harshman claim that average pitch and extreme endpoint are nonlinguistic dimensions.[5] Hombert (1976) used the same PARAFAC technique to investigate the perception of tone patterns associated with bisyllabic (V_1CV_2) meaningful nouns in Yoruba. Two dimensions were extracted, the first one correlated clearly with the direction of F_o change (falling, level, rising). The interpretation of the second dimension is not clear; the magnitude of the slope of V_2 seems to be the more important factor. However, results of lineal regression analysis indicate that the averaged fundamental frequency of V_2 and the difference between fundamental frequency offset of V_1 and the onset of V_2 are also included in dimension 2.

It is possible that with a larger number of subjects these factor would have been extracted as separate dimensions.

3.4. Computation of perceptual distance

Considering these experimental results, it was decided to use four perceptual factors: the <u>average</u> F_o, the <u>onset</u>, the <u>offset</u>, and the <u>slope</u> of the tone.[6] Since the sign of the slope was considered, this last factor includes direction of change as well. Distinction based on duration was not considered here.

The perceptual distance between any two tones (taken from the set of 19 tones presented in 3.2) was then computed in two ways:

a. city-block distance:

$$D_{cb}\ (T_{ij},\ T_{kl}) = \sum_{m=1}^{4} \alpha m\ \Delta f_m\ (i,j,k,l)$$

b. euclidian distance:

$$D_e\ (T_{ijk}\ T_{kl}) = \sqrt{\sum_{m=1}^{4} \alpha_m\ \Delta f_m^2}\ (i,j,k,l)$$

with Δf_1 = difference in average F_o between T_{ij} and T_{kl} = $\left|\frac{i+j}{2} - \frac{k+l}{2}\right|$

Δf_2 = difference at onset between T_{ij} and T_{kl} = $|i - k|$

Δf_3 = difference at offset between T_{ij} and T_{kl} = $|j - l|$

Δf_4 = difference in slope between T_{ij} and T_{kl} = $|(j-i) - (l-k)|$

Theoretically, weighting factors obtained from multidimensional scaling analyses could have been used instead of arbitrary values. But since the available data are very restricted, it was decided to test several sets of weighting factors. These eight sets are presented in the left hand column of Table 2.

3.5. Optimization

Given the number of tones [N] in the system and the set of weighting factors [α], the distance D (T_{ij}, T_{kl}) was computed for each pair of tones for each possible combination of N tones selected among the 19 tones mentioned earlier. For each tone system considered, the smallest distance (DMIN) between two toens was stored. A system was considered perceptually optimum for a given set of input parameters (i.e. number of tones and set of weighting factors) when it was found to have the greatest DMIN after all possible tone systems with N tones were compared. In other words, the chosen perceptual criteria was to keep the two closest tones of a given system maximally apart. A simple example will help clarify this procedure. Let us consider two three tone systems S_1 (11, 33, 55) and S_2 (11, 51, 55) and the set of weighting factors $\alpha(1,1,1,1)$, i.e. α_1 (average F_o) = 1, α_2 (offset) = 1, α_3 (onset) = 1, α_4 (slope) = 1. The DMIN for each tone system can be computed as follows:

S_1: D (11,33) = $\alpha_1\ \Delta f_1 + \alpha_2\ \Delta f_2 + \alpha_3\ \Delta f_3 + \alpha_4\ \Delta f_4$

= $1 \times |1-3| + 1 \times |1-3| + 1 \times |1-3| + 1 \times |0-0|$

= 2 + 2 + 2

= 6

D (11,55) = 12

D (33,55) = 6

thus DMIN (S_1) = 6

S_2: D (11,51) = 2 + 4 + 0 + 4 = 10
D (11,55) = 12
D (51,55) = 10
thus DMIN (S_2) = 10

Consequently, with this set of weighting factors (i.e. equal weight given to average F_o, onset of the tone, offset of the tone, and slope of the tone) and this type of perceptual distance, S_2 (11,51,55) would be perceptually preferable to S_1 (11,33,55).

2.6. Relative weight of perceptual and articulatory data

The total lack of data in this area forces us to disregard this step. It should be noted however that some implicit weighting decision was made when some of the tones of the complete set of 25 tones were discarded for articulatory reasons. At this stage, articulatory factors were given complete priority enabling these tones to be "saved" on perceptual grounds.

2.7. Output of the model

For each condition (i.e. N and α) the output was constituted by three types of data: (i) the number of systems which were perceptually equivalent under this condition; (ii) the shapes of the tones of these systems; and (iii) the minimum distance between any two tones for these systems (since these systems are perceptually equivalent, they have the same distance between their two closest tones).

The number of systems perceptually equivalent is presented in Table 2 when city-block distance is used, and in Table 3 when Euclidian distance is used.[7] The predicted systems are in very good agreement with the most commonly found tone systems presented in section 2, that is, 55, 11, 15, 51 for the four tone system; 55, 33, 11, 15, 51 for the five tone system, and 2 level, 2 rising and 2 falling for the six tone system. The predictions were generallly bad for the three tone system (i.e. we were not able to predict 55, 33, 11), and for the two tone system the predictions were split between 11, 55 (i.e. the correct prediction) and 15, 51. We will present a possible explanation for this difference in accuracy of prediction for two vs. three tone systems in a moment. It should be emphasized that these predictions are relatively independent of the set of weighting factors used and consequently can be considered as a perceptually stable solution. Let us notice however that the predictions are improved when a relatively heavier weight is attached to the average F_o parameter for systems with small numbers of tones and when a relatively heavier weight is attached to the slope parameter for systems with larger numbers of tones.

Figure 2 shows the normalized distance[8] between the two closest tones of all systems perceptually optimum as a function of the number of tones. The dots represent distances computed with the city-block

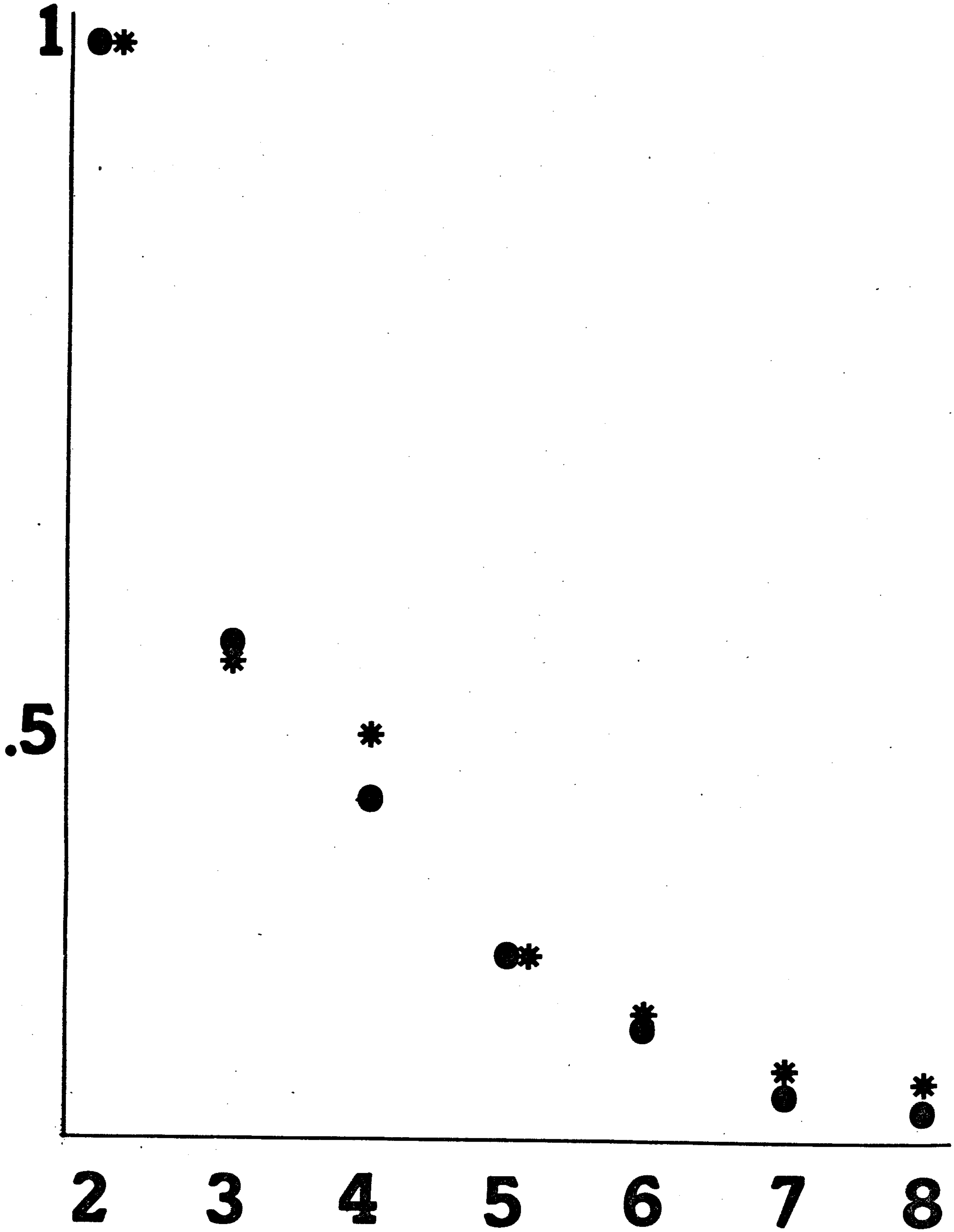
TABLE 2. Number of systems perceptually equivalent - city block distance

average F_o	offset	onset	slope	2	3	4	5	6
1	1	0	0	1	1	4	199	39
2	1	0	0	1	3	48	8	1
0	1	1	0	2	8	13	1	4361
1	1	1	0	1	2	3	6	1065
1	1	0	1	1	6	1	22	1
0	1	0	1	1	1	32	8	1
1	0	0	1	1	2	7	104	11
1	1	1	1	1	6	1	9	342

TABLE 3. Number of systems perceptually equivalent (Euclidian distance)

				2	3	4	5	6
1	1	0	0	1	1	4	13	1
2	1	0	0	1	2	7	6	1
0	1	1	0	2	2	1	1	342
1	1	1	0	1	12	1	1	342
1	1	0	1	1	13	1	48	4
0	1	0	1	1	1	1	12	1
1	0	0	1	2	2	1	104	11
1	1	1	1	2	6	1	9	129

Fig 2. Normalized perceptual distance between the two closest tones of perceptually optimum systems as a function of the number of tones in the system.



distance, the asterisks represent Euclidian distances. Each point is the average of eight distances obtained from the eight different sets of weighting factors. Two important facts should be noted:

1. The two types of distances (city-block and Euclidian) give extremely similar results as far as minimum perceptual distances are concerned.

2. The two extremities of the graph deviate noticeably from the almost linear relationship obtained for 3, 4, 5, and 6 tone systems.

a. The flattening of the curve obtained for 7 and 8 tone systems indicates some sort of saturation in the tone space. A huge number of tone systems are then found to be perceptually equivalent. This implies high instability. These data are in agreement with the fact that it is extremely rare to find tone languages with 7 or 8 tones using only F_o to distinguish these tones.[9]

b. The fact that the minimum distance is much larger in the case of two tone systems than for other systems suggests that we were wrong in assuming that when a system has two level tones, these two tones are 11 and 55. These data seem to indicate that these two tones should be closer to each other.[10] This may explain why we got more accurate predictions for the two tone systems than for the three tone systems. This discrepancy is only an artifact of the conventions we used to code high and low level tones.

4. Conclusion

Although it is clear that articulatory, perceptual and acoustic data are badly needed to improve and validate the model, the results presented here are quite encouraging. They suggest that tone shapes can be predicted with reasonable success knowing the number of tones in the system, only on the basis of articulatory and perceptual data, and independently of the origin and history of the particular systems considered. These results suggest that (i) the perceptual parameter slope acquires more importance when the number of tones increases; (ii) fundamental frequency alone is not sufficient to code 7 or 8 tones (in these cases secondary cues are used, e.g. phonation types, syllable structure); and (iii) the spacing between the top and the bottom of the fundamental frequency range is smaller for two tone systems than for systems with greater numbers of tones.

Footnotes

1. In fact they used a "perceptual" representation of the acoustic space by transforming the linear frequency scale into mel scale .

2. Since languages with more than five tones were poorly represented in Ruhlen's and PAP's surveys, I used data from Tai (Brown 1965), Sarawit (1973) and Miao-Yao languages (Chang 1973).

3. Relatively smaller numbers correspond to relatively lower F_o (e.g. "51" would represent a tone falling from high to low).

4. We obviously do not imply that the voiced portion of the tone bearing unit has a perfect steady-state F_o.

5. It should be noted that if these two dimensions are considered non-linguistic, Gandour and Harshman's dimensions do not allow measuring differences between level tones in a tone language.

6. Obviously these three factors are not independent: the slope can be derived from the onset and the offset of the tone for instance. Choosing slope as a separate factor implies that information can be provided by the slope itself.

7. Since the number of systems for 7 and 8 tone systems was extremely high, only the distance was computed in most of these cases.

8. This normalized distance was obtained by dividing the computed distance by the distance obtained for a two tone system.

9. Systems with more than 6 tones generally use secondary cues such as phonation types of syllable structure to distinguish some of the tones.

10. In a recent paper, Maddieson (1977) suggested that the frequency range used by speakers of tone language was a function of the number of level tones in the system (the greater number of level tones, the greater the frequency range). These data partially support and partially refute Maddieson's claim since they indicate that the distance between high and low is smaller in a two tone system than in a system with a greater number of tones but they do not indicate that the same increase in F_o range is found between 3 and 4 or 4 and 5 tone systems.

References

Brady, P.T., House, A.S. and Stevens, K.N. 1961. Perception of sounds characterized by a rapidly changing resonant frequency. *Journal of the Acoustical Society of America* 33, 1357-1362.

Brown, J.M. 1965. *From ancient Thai to modern dialects.* Bangkok: Social Science Assoc. Press of Thailand.

Chang, K. 1973. The reconstruction of proto-Miao-Yao tones. *Bulletin of the Institute of History and Philology-Academia Sinica* 44, (4), 541-628.

Chang, K. 1975. Tonal developments among Chinese dialects. *Bulletin of the Institute of History and Philology-Academia Sinica* 46, (4) 636-710.

Chao, Y.R. 1930. A system of tone letters. *Le Maitre Phonètique* (3), 45:24-27.

Gandour, J. and R. Harshman. 1977. *Cross language study of tone perception.* Bell Laboratories, Murray Hill, New Jersey.

Groot, A.N. De. 1931. Phonologie und Phonetik als Funktionswissenschaften. *Travaux du Cercle Linguistique de Prague* 4. Prague.

Harshman, R.A. 1970. Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multidimensional factor analysis. *UCLA Working Papers in Phonetics* 16.

Haudricourt, A.G. 1954. De l'origine des tons en vietnamien. *Journal Asiatique* 242, 69-82.

Haudricourt, A.G. 1961. Bipartition et tripartition des systèmes de tons dans quelques langages d'Extreme-Orient. *Bulletin de la Société Linguistique de Paris* 56, 163-180.

Heinz, J.M., Lindblom, B. and Lindqvist, J. 1968. Patterns of residual masking for speech-like characteristics. *IEEE Transactions Audio and Electroacoustics*, AU-16, 107-111.

Hombert, J.M. 1975. Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in language. University of California Berkeley Ph.D. dissertation.

Hombert, J.M. 1976. Perception of bisyllabic nouns in Yoruba, In L.M. Hyman, L.C. Jacobson and R.G. Schuh (Eds.), *Papers in African Linguistics in Honor of W.E. Welmers.* Studies in African Linguistics, 6, 109-122.

Hombert, J.M. 1977. Difficulty of producing different Fo in speech. *Journal of the Acoustical Society of America* 60, 1, 544.

Hombert, J.M. In Press. Consonant types, vowel height and tone. In *Tone: A linguistic survey*. V. Fromkin (ed). Academic Press.

Hombert, J.M., J.J. Ohala, and W.G. Ewan. 1976. Tonogenesis: Theories and queries. *Report of the Phonology Laboratory* 1, Berkeley, 48-77.

Lindblom, B.E.F. 1972. Phonetics and the description of language. *Procedings of the 7th International Congress of Phonetic Sciences*. Montreal. Mouton. The Hague.

Lindblom, B.E.F. 1975. Experiments in sound structure. Paper read at the 8th International Congress of Phonetic Sciences, Leeds.

Lindblom, B.E.F. and J. Sundberg. 1969. A quantitative model of vowel production and the distinctive features of Swedish vowels. *STL-QPSR (KTH) Stockholm* 1, 14-30.

Lindblom, B.E.F. and J. Sundberg. 1971. Acoustical consequences of lip, tongue, jaw and larynx movement. *JASA* 50, 1166-79.

Maddieson, I. 1977. Tone loans, a question concerning tone-spacing and a method of answering it. *UCLA Working Papers in Phonetics* 36.

Martinet, A. 1955. *Economie des changements phonétiques*. Berne: Francke.

Matisoff, J. 1973. Tonogenesis in Southeast Asia. In L.M. Hyman (ed), *Consonant types and tone. Southern Calif. Occasional Papers in Linguistics*. Univ. of So. Calif. pp. 71-96.

Mazaudon, M. 1977. Tibeto-Burman tonogenetics. To appear in: *Linguistics of the Tibeto-Burman Area.*

Nabelek, I.V. and Hirsh, I.J. 1969. On the discrimination of frequency transitions. *JASA* 45, 1510-1519.

Nabelek, I.V., Nabelek, A.K. and Hirsh, I.J. 1970. Pitch of tone bursts of changing frequency. *JASA* 48, 536-533.

Passy, P. 1890. *Étude sur les changements phonétiques et leur caractères généraux*. Paris.

Pollack, I. 1968. Detection of rate of change of auditory frequency. *Journal of Experimental Psychology* 77, 535-541.

Ruhlen, M. 1975. *A guide to the languages of the world.* Stanford.

Sarawit, M. 1973. The Proto-Tai vowel system. Unpublished Ph.D. dissertation, University of Michigan, Ann Arbor.

Tsumura, T., Sone, T. and Nimura, T. 1973. Auditory detection of frequency transition. *Journal of the Acoustical Society of America* 53, (1), 17-25.

The effect of aspiration on the fundamental frequency of the following vowel

Jean-Marie Hombert and Peter Ladefoged

[Paper presented at the 91st meeting of the Acoustical Society of America]

In this paper we will be concerned with the effect of aspiration on the fundamental frequency of the following vowel. More precisely we will be interested in comparing fundamental frequency values after voiceless aspirated vs. voiceless unaspirated stops. This study will bring new insights into issues such as the timing of articulatory gestures involved in the production of these stops as well as the phonetic plausibility of certain sound changes (e.g. the development of contrastive tones from consonant mergers).

1. Existing phonetic data

A number of phonetic studies (Fischer-Jørgensen 1972; Hombert 1975; House and Fairbanks 1953; Lea 1972, 1973; Lehiste and Peterson 1961; Löfqvist 1975 ; Mohr 1971; Slis and Cohen 1969) show that the fundamental frequency at the onset of a vowel is significantly higher after a voiceless stop than after a voiced stop. The process by which these intrinsic fundamental frequency perturbations (i.e. non-intended by the speaker) become linguistically relevant by giving rise to contrastive tones is rather well understood (Haudricourt 1961; Hombert 1975; Hyman forthcoming; Matisoff 1973; Mazaudon forthcoming). The data are not as straightforward when one compares the effect of a voiceless aspirated vs. voiceless unaspirated stop on the Fo of the following vowel.

It is generally assumed (Hombert 1975) that voiceless aspirated stops give rise to a higher Fo at the onset of the following vowel. This claim is supported by data from languages such as Korean (Han 1967; Kim K. 1968), Thai (Ewan 1976) and Danish(Jeel 1975). However, conflicting data are found from the same as well as other languages. They sometimes show no difference in Fo onset depending on the preceding consonant as illustrated by Danish (Fischer-Jørgensen 1968b) and sometimes they exhibit the opposite pattern, that is a higher Fo after voiceless unaspirated than after voiceless stops as in Korean (Kagaya 1974); Thai (Gandour 1974; Erickson 1975) and Hindi (Kagaya and Hirose 1975). In order to bring new data which hopefully will help clarify this issue, we decided to investigate the two series of voiceless stops in English and French (the English series is voiceless aspirated as opposed to the French series which is supposed to be voiceless unaspirated).

2. Experimental procedure

Two American English speakers (1 Female and 1 Male) and two French speakers (1 Female and 1 Male) were used in this experiment. They spoke 6 CV nonsense words where C=[p,t,k,b,d,g] and V=[i]. The word list consisted of 10 tokens of each test word arranged in random order. Each test word was uttered in the frame "Say ___ louder". No special instructions were given with respect to the speed of reading. The recording was done in a sound treated room. The speech waveform was sampled at an effective rate of 20,000 Hz (Sample rate=10,000 and tape played back at half speed). Since most Fo extractor methods perform poorly in determining the FO during the first few cycles and since these measurements were crucial in our study, we made our measurements directly on the digitized waveform. In order to get maximum accuracy in locating similar points on each consecutive period, frequencies above 1000 Hz were attenuated before digitization by a low pass filter with a 24 db per octave slope.

3. Results

The voice onset time (v.o.t.) values for the French [p,t,k] were shorter than those for the corresponding English sounds (approximatively 15, 30 and 50 msec respectively vs. 45, 65 and 85 msec for English). As was already noticed by Fischer-Jørgensen (1972) these v.o.t. values from French are longer than one would expect for so-called voiceless unaspirated sounds. On the four graphs, in Figure 1, fundamental frequency is plotted as a function of successive vocal fold cycles (horizontal axis). Although this axis distorts the data when compared with the usual time axis, it gives more useful information for our purposes. The upper line represents the average Fo after voiceless unaspirated stops in English and after voiceless unaspirated in French; the lower line corresponds to the two other series (voiceless unaspirated or "voiced" in English and voiced in French). Fo values corresponding to each individual stop [p,t,k,b,d,g] are represented by corresponding symbols on the graph. It is clear from these data that there is not a direct correlation between the duration of aspiration after a voiceless consonant and the onset fundamental frequency of the following vowel since English and French voiceless consonants have very similar perturbatory effect on the Fo of the following vowel.

4. Discussion

In order to be able to explain why voiceless aspirated and voiceless unaspirated stops have similar effects on the Fo of the following vowel, let us examine how the various parameters controlling Fo are affected by the aspirated vs. unaspirated distinction.

State of the glottis:

Glottal opening: The fact that the glottis is widely open upon release of a voiceless aspirated stop as opposed to a more adducted position for a voiceless unaspirated stop is irrelevant here. We are concerned only with the state of the glottis when it is vibrating. The important question has to do with the relative timing of onset of voicing and glottal adduction. There is no reason to suppose that the onset of voicing occurs at different degrees of glottal adduction in these sounds. There is a difference in the muscular activity, in that PCA activity is more important 50 to 100 msec before voice onset (Hirose 1975). But there is not evidence that this will affect the rate of vibration of the vocal folds.

Muscular tension: Horizontal tension. The vocalis (VOC), the lateral cricoaretenoid (LCA), and the cricothyroid (CT) are associated with Fo increase (Hirose and Gay 1972; Ohala 1970). The sternohyoid has been reported to cause (Ohala 1972) or to be associated with Fo lowering (Erickson 1975b). None of these muscles seem to be associated with the voiceless aspirated vs. voiceless unaspirated distinction, and in most cases these muscles are not involved in the voiced/voiceless distinction either (Hirose and Gay 1972). However, it has been reported that CT activity is used in Hindi to differentiate between voiced and voiceless stops (Dixit 1975) ; this difference in CT activity is not obvious from Kagaya and Hirose's data (1975). In any case, even if the CT is associated with the voiced/voiceless distinction in Hindi, it does not affect our point since we are concerned with aspirated/unaspirated differences, and there is no significant difference between CT activity for voiceless aspirated vs. voiceless unaspirated stops in Hindi.

Vertical tension: It has been proposed (Ohala 1972) that vertical movements of the larynx change the rate of vibration of the vocal folds by changing primarly the vertical rather than the anterior-posterior tension of the cords. Experimental data show that in Thai (Ewan 1973, 1976; Ewan and Krones 1974; Gandour and Maddieson 1976) and Hindi (Ewan and Krones 1974) voiceless aspirated and unaspirated stops have a higher larynx position than their voiced counterparts; but there is no consistent difference between aspirated and unaspirated stops in Thai and Hindi. Here again we have a parameter (larynx height) which accounts for Fo differences after voiced vs. voicless stops but does not predict any difference in Fo after voiceless aspirated vs. voiceless unaspirated stops.

Aerodynamic Forces: Transglottal pressure:

Subglottal Pressure. It has been found in a number of languages such as English (Ladefoged 1967, 1968, 1974;Lieberman 1967; McGlone and Shipp 1972; Murry and Brown 1975; Netsell 1969), Swedish (Löfqvist 1975), Hindi (Ohala and Ohala 1972; Dixit 1975) and Dutch (Slis 1970) that the respiratory system generates a constant subglottal pressure irrespective of the stop categories (voiced vs. voiceless, aspirated vs. unaspirated). It seems that the variations in subglottal pressure can be predicted from glottal and supraglottal adjustments (e.g. glottal opening, mouth opening) even in the case of languages such as Jinghpo (Ladefoged, 1971) and Korean (Kim 1965; Lee and Smith

1972) in which different subglottal pressures have been associated with different stop categories. The subglottal pressure is generally lower at the onset of a vowel following an aspirated stop as a result of the high rate of airflow through the glottis during the aspiration period. Since other things being equal a lower subglottal pressure will lead to a lower Fo, we would therefore expect a slightly lower Fo after the aspirated series (which is the opposite of what is generally assumed).

Oral Pressure. Our measurements were made after the release of the stop closure. At that time the oral pressure is very low and probably does not differ in the different series of stop consonants. If there is any difference it would be that the oral pressure was higher after the aspirated stops, since the higher flow rate would cause a higher back pressure. If the oral pressure is higher after aspirated stops, then the transglottal pressure drop would be lower. This, again, would lead to the unexpected result of a lowering of Fo after aspirated stops.

Airflow rate. When the rate of airflow increases, the Bernoulli force increases and consequently Fo increases. It has been suggested that the reason why Fo was higher after voiceless aspirated consonants than after unaspirated ones was that the rate of airflow was higher in the first case (Ladefoged, 1967, 1974; Ohala, 1973). However, experimental data do not show a higher rate of airflow *at vowel onset* depending on the preceding consonant (Isshiki and Ringel, 1964; Van Hattum and Worth, 1967). The data of Klatt, Stevens and Mead (1968) show a slightly higher airflow after the aspirated stop for only a short period of time (about 50 msec). Although it is true that some data (Subtelny et al, 1969) show a higher airflow rate after the aspirated vs. the unaspirated series (1070 vs. 650 ml/s) for a significantly longer period of time (over 200 msec after vowel onset) it seems likely that these results reflect the poor frequency response of the flowmeter (warm-wire method) used in this experiment (Kozhevnikov and Chistovitch, 1966; Hardy, 1967; Van den Berg, 1962) rather than an actual long-term airflow rate difference caused by the preceding consonant. In fact, the model of speech aerodynamics proposed by Ohala (1976) gives a smaller value for airflow at the onset of a vowel following an aspirated stop as opposed to a higher value after a voiceless unaspirated stop. It does not seem that available airflow data allow us to predict significant differences in Fo after voiceless aspirated vs. unaspirated stops.

5. Conclusion

In conclusion we have shown that the parameters controlling the rate of vibration of the vocal folds are not significantly different at the onset of a vowel following a voiceless aspirated stop vs. a voiceless unaspirated stop. This explains the data presented at the beginning of this paper where it was shown that English voiceless aspirated stops and French voiceless unaspirated stops had very similar effects on the fundamental frequency of the following vowel. There is neither evidence nor reason to expect that the increase in Fo after a voiceless stop will be positively correlated with the duration of the aspiration.

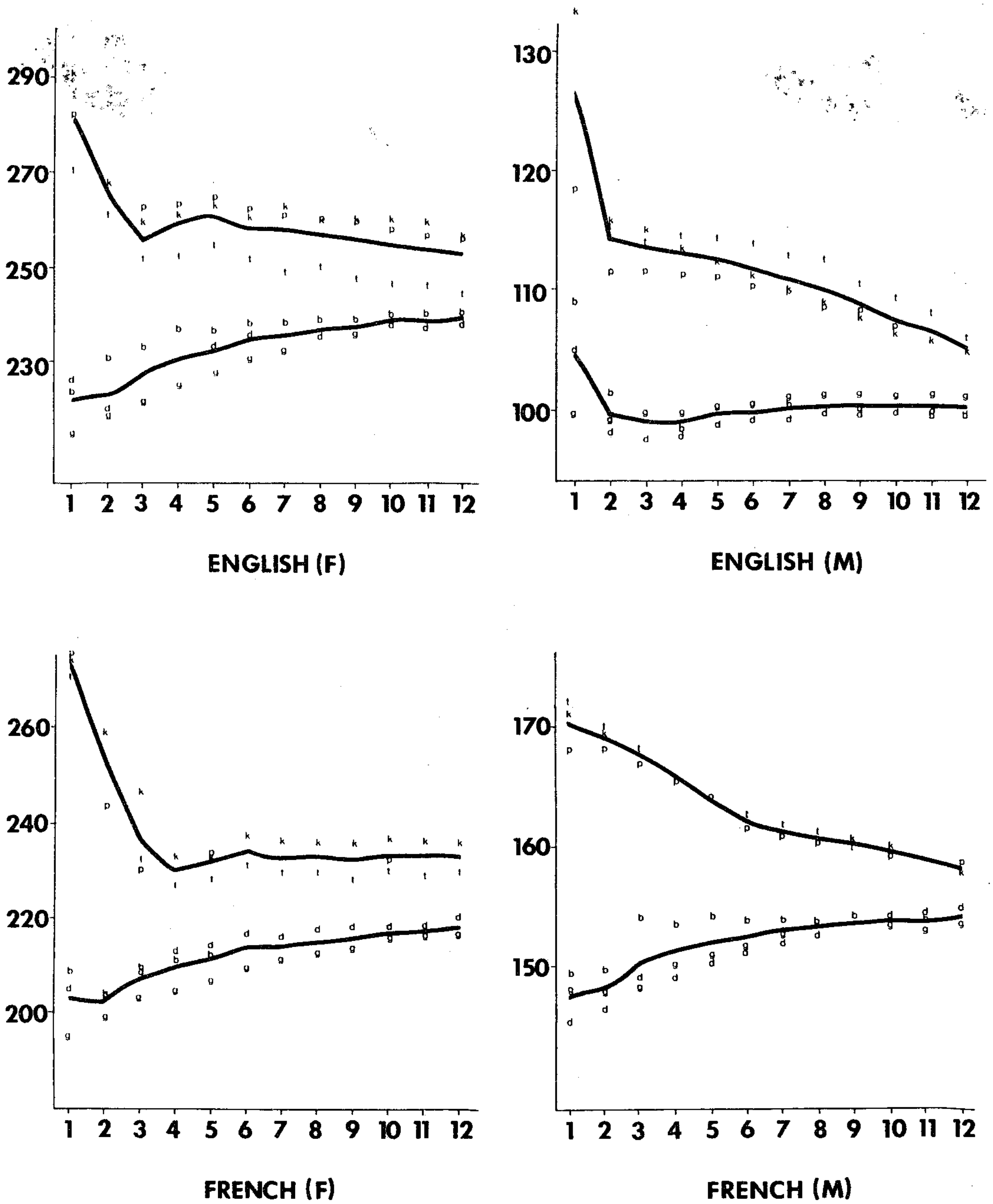


Fig. 1

Fundamental frequency measurements (vertical axis) of vowels following [p, t, k] and [b, d,g] as a function of glottal period (horizontal axis) for two American English speakers (1 female and 1 male) and two French speakers (1 female and 1 male). The upper curve represents average Fo measurements after ptk, the lower curve represents average Fo measurements after bdg.

References

Berg, van den, J. (1962) Modern Research in experimental phoniatrics. *Folia Phoniatrica.* 14, 81-149.

Dixit, R.P. (1975) Neuromuscular aspects of laryngeal control: with special reference to Hindi. Univ. of Texas, Austin, Ph.D. dissertation.

Erickson, D. (1975) Phonetic implications for an historical account of tonogenesis in Thai. *Studies in Tai Linguistics in honor of W.J. Gedney.* Edited by J.G. Harris and J.R. Chamberlain. Bangkok Central Inst. of English Lang. Office of State Universities.

Erickson, D. and J.E. Atkinson (1975) The function of strap muscles in speech. *Journ. Acoust. Soc. of Amer.* 58,1, F4.

Ewan, W.G. (1976) *Laryngeal behavior in speech.* UC Berkeley. Ph.D. dissertation.

Ewan, W.G. and R. Krones. (1974) Measuring larynx movement using the thyroumbrometer. *Journ. of Phonetics*, 2, 327-335.

Fischer-Jørgensen, E. (1968) Les occlusives francaises et danoises d'un sujet bilingue. *Word.* 24, 112-153.

Fischer-Jørgensen, E. (1972) PTK et BDG Français en position vocalique accentuée in A. Valdman (ed.) *Papers in Linguistics and Phonetics to the memory of Pierre Delattre.* Mouton-The Hague. 143-200.

Gandour, J.T. (1974) Consonant types and tone in Siamese. *Journ. of Phonetics*, 2, 337-50.

Han, M. (1967) Studies in the phonology of Asian languages. Vol. V. Acoustic features in the manner differentiation of Korean stop consonants. Acoustic Phonetics Res. Lab. Univ. of Southern Calif.

Hardy, J.C. (1967) Techniques of measuring intraoral air pressure and rate of airflow. *Journ. of Speech and Hearing Res.*10, 650-654.

Hattum van, R.J. and Worth, J.H. (1967) Airflow rates in normal speakers. *The Cleft Palate Journ.* 4.2. 137-147.

Haudricourt, A.G. (1954) De l'origine des tons en vietnamien. *Journ. Asiatique.* 242, 69-82.

Haudricourt, A.G. Bipartition et tripartition des systèmes de tons dans quelques langages d'Extreme-Orient. *Bulletin de la Societe Linguistique de Paris.* 56, 163-180.

Hirano, M. and Ohala, J. (1969) Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. *Journ. of Speech & Hearing Res.*12, 362-373.

Hirose, H. (1975) The posterior cricoarytenoid as a speech muscle. *Annual Bullet. Res. Inst of Logopedics and Phoniatrics*, 9, 47-66.

Hirose, H. and Gay, T. (1972) The activity of the intrinsic laryngeal muscles in voicing control; an electromyographic study. *Phonetica* 25, 140-164.

Hirose, H. Lisker, L. and Abramson, A. (1973) Physiological aspects of certain laryngeal features in stop production. (Abstract). *Journ. Acoust. Soc. of Amer.* 53, 294-295.

Hombert, J.M. (1975) *Towards a theory of tonogenesis: an empirical, physiologically and perceptually-based account of the development of tonal contrasts in language.* Ph.D. dissertation, UC Berkeley.

House, A.S. and G. Fairbanks (1953) The influence of consonant environment upon the secondary acoustical characteristics of vowels. *Journ. of the Acoust.Soc. of Amer.* 25, 105-113.

Isshiki, N. and Ringel, R. (1964) Air flow during the production of selected consonants. *Journ. of Speech and Hearing Res.* 7, 233-244.

Jeel, V. (1975) An investigation of the fundamental frequency of vowels after various Danish consonants, in particular stop consonants. *Annual Report of the Inst.of Phonetics.* Univ. of Copenhagen 9, 191-211.

Kagaya, R. (1974) A fiberoptic and acoustic study of the Korean stops, affricates and fricatives. *Journ. of Phonetics* 2, (2) 161-180.

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

Kim, C.W. (1965) On the autonomy of the tensity feature in stop classification with special reference to Korean stops. *Word* 21, 339-359.

Klatt, D.H., Stevens, K.N. & Meade, J. (1968) Studies of articulatory activity and airflow during speech in sound production in man. In A. Bouhuys (Ed.) *Annals of the New York Academy of Sciences.* New York, 42-55.

Kozhevnikov, V.A. and Chistovich, L.A. (1966) *Speech: articulation and perception.* Distributed by U.S. Dept. Commerce, Washington D.C. Moscow-Leningrad.

Ladefoged, P. (1967) *Three areas of experimental phonetics.* Oxford Univ. Press.

Ladefoged, P. (1968) Linguistic aspects of respiratory phenomena. In A. Bouhuys (Ed.s) Sound production in man. *Annals of the New York Academy of Sciences* 155, 141-151.

Ladefoged, P. (1971) *Preliminaries to linguistic phonetics.* Chicago Univ. Press.

Ladefoged, P. (1974) Respiration, laryngeal activity and linguistics In B.D. Wyke (Ed.) *Proceedings of the International Symposium on Ventilatory and Phonatory Control Systems.* Oxford Univ.Press.

Lea, W.A. (1972) *Intonational cues to the constituent structure and phonemics of spoken English.* Purdue Univ. Ph.D. dissertation.

Lea, W.A. (1973) Segmental and suprasegmental influences on fundamental frequency contours. In L.M. Hyman (Ed.) Consonant Types and Tone. *Southern Calif. Occasional Papers in Linguistics.* Univ. of So. Calif. 1, 15-70.

Lee, C.Y. and Smith, T.S. A study of subglottal air pressure in Korean stop consonants. *Journal of Phonetics.*

Lehiste, I. and Peterson, G.E. (1961) Some basic considerations in the analysis of intonation. *Journ. of the Acoust. Soc. Amer.* 33, 419-425.

Lieberman, P. (1967) *Intonation, perception and language.* MIT Res. Monograph. 38.

Löfqvist, A. (1975) Intrinsic and extrinisc Fo variations in Swedish tonal accents. *Phonetica,* 31, 228-247.

McGlone, R.E. and Shipp. T. (1972) Comparison of subglottal airpressures associated with /p/ and /b/. *Journ. Acoust. Soc. Amer.* 51, 664-665.

Mohr. B. (1971) Intrinsic variations in the speech signal. *Phonetica,* 23, 65-93.

Murry, T. and Brown, W.S. (1975) Aerodynamic interactions associated with voiced-voiceless stop cognates. To appear in *Proceedings of 8th Intern.Congress of Phonetic Sciences,* Leeds, August.

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

Ohala, J.J. How is pitch lowered? *Journ. Acoust. of Amer.* 52. 124.

Ohala, J.J. (1973) The physiology of tone. In L. Hyman (Ed.) Consonant types and tone. *Southern Calif. Occasional Papers in Linguistics* 1. Univ. of Southern Calif.

Ohala, J.J. (1976) A model of speech aerodynamics. *Report of the Phonology Lab.* Berkeley, 1, 93-107. (To appear in the Proceedings of the 8th Int'l Congress of phonetic sciences, Leeds, August 1975.)

Ohala, M. and Ohala, J.J. (1972) The problem of aspiration in Hindi phonetics. *Project on Linguistic Analysis Reports.* Berkeley, 16, 63-70. *Annual Bulletin.* Research Institute of Logopedics and Phoniatrics, 6, 39-46.

Slis, I.H. (1970) Articulatory measurements on voiced, voiceless and nasal consonants. *Phonetica* 21, 193-200.

Cerebral dominance for pitch contrasts in tone language speakers and in musically untrained and trained English speakers

Diana Van Lancker and Victoria A. Fromkin

[Department of Linguistics,

University of California, Los Angeles]

Abstract.

Speakers of a tone language, Thai, recognize pitch contrasts which are linguistically significant in their language better at the right ear in a dichotic listening task, but show no ear advantage for the same pitch contrasts occurring in a nonlinguistic context. American English speakers, divided into musically untrained and trained groups, show no ear advantage for those same pitch contrasts occurring in a non-linguistic context. American English speakers, divided into musically untrained and trained groups, show no ear advantage for those same pitch contrasts. The only effect of musical training is an enhancement of left ear accuracy for pitch contrast recognition.

A large number of studies on cerebral function have investigated left and right hemispheric specialization (1). For auditory stimuli, some researchers have asked whether lateralization to one cerebral hemisphere is based on acoustic features or on the functional context of the stimulus. Normal humans can be tested by dichotic listening (2), in which a subject hears two different, simultaneous sounds (one at each ear. The subject typically makes errors in reporting what he or she hears. These errors are greater at one or the other ear, depending on the type of stimulus. Because of stronger contralateral connections (than ipsilateral) between ear and hemisphere, a "right ear advantage" (higher accuracy at the right ear than the left) is assumed to indicate preferential processing by the left hemisphere, and vice versa. Such findings for many diverse stimuli have contributed to hypotheses on hemispheric specialization.

Most dichotic listening results have suggested that the functional context of a given stimulus predicts ear advantage. For example, if a stimulus is perceived as part of the linguistic system native to the subject, then those stimuli are more accurately heard at the right ear. In one study, vowels in syllables yielded a right ear advantage, whereas the same sounds embedded in a nonlinguistic context yielded a left ear advantage (3). Environmental sounds (4) and some musical stimuli (5) have shown a left ear advantage. These data support the "functional" theory of hemispheric specialization for processing auditory (and visual and other)stimuli.

Our experiments on Thai tones strongly confirm the functional theory of hemispheric specialization. In the first experiment (6), we found that pitch contrasts were better identified at the right ear when constituting the tone contrasts of a tone language (Thai), but not when presented (to the same subjects) as hums. In a tone language, pitch alone may contrast the meanings of words which are otherwise identical in sounds. In a non-tone language, such as English, pitch is used to convey information (e.g., the intonation of an interrogative sentence will differ from that of a declarative) but the meaning of 'cat' will remain the same whether it is produced with a high, medium, low rising, or falling pitch in a sentence such as 'The cat is on the mat'. In Thai, however, (or in languages such as Chinese, Twi, Hausa, etc.) a word like 'naa' will mean different things depending on the pitch. (See Table I)

For the dichotic listening study,we prepared three audio tapes, using stimuli produced by a female native Thai speaker. The stimuli were three kinds: 1) five Thai words differing only in tone (the tone-word stimuli; 2) five Thai words contrasting only in initial consonant (the consonant-word stimuli), all occurring on mid-tone; 3) and five "hums" (to produce these the five Thai tone contrasts were hummed). (Table I).

For each of the three sets, every stimulus was paired with every other stimulus, producing 20 pairs. These were presented twice, with channels reversed the second time; the earphones were then shifted and the forty pairs presented again. Thus there were 80 stimulus pairs for each set. The pairs were presented with a six-second interval, with the exception of a block of 40 pairs in the tone word set, which were presented as doubled pairs. This was done to increase errors.

The subjects were provided with eight answer sheets with five columns, ten rows per page, eight pages for each stimulus set. Each column was headed by the appropriate 'tone-word' or 'consonant-word', designated in both Thai and English orthography, and by iconic diacritics to represent the tones. These diacritics plus the Thai word for the tones headed the 'hum' column. For each stimulus pair, the subjects were directed to respond by marking an L under the appropriate column to indicate the stimulus heard in the left ear, and an R under the column that identified the stimulus heard in the right ear. The order of reporting left-ear-right-ear, or right-ear-left-ear was reversed every ten pairs.

Before each of the three stimulus sets, practice sessions were conducted for all subjects using binaurally presented stimuli.

For the first study using 23 right-handed native Thai speakers, we found a significant right ear effect for tone-words and consonant-words, but no ear effects for hums. We inferred that the Thai speakers were processing the feature contrasts in the pitches (tone-words) and consonants as language, and were therefore engaging the left hemisphere for these two tasks. The hums yielded no significant ear effects because they were not linguistically significant (although containing the same contrasts as the tones).

Our second experiment was conducted to determine whether a similar result would be obtained using subjects who were not native speakers of a tone language but who were musically trained. Our interest was to see if the right ear effect found in the Thai subjects was primarily due to the linguistic system of these speakers or whether it could have been due to greater familiarity with pitch contrasts. Our interest in this question was enhanced by a recent finding of Bever and Chiarello , who reported that musically trained subjects recognized melodies better at their right ear, whereas untrained subjects demonstrated the usual left ear advantage for musical sounds (in a monaural listening task). These authors suggested that musical training may have "real neurological concomitants" in requiring different strategies which engage the left hemisphere for otherwise right hemisphere tasks.

For this experiment, therefore, we selected 40 right-handed male native English speakers as subjects, divided into two groups with respect to musical training. The musically untrained group had had no formal musical training in the past ten years and had four years or less total exposure. Nearly all had had zero to one year in elementary school and no musical activity since. The musically trained subjects had had at least eight years formal training, some as many as 40 years active participation. Most were currently playing an instrument or singing in a chorus.

These subjects received the binaural training session and heard the same dichotic stimuli as the Thai population.

For both untrained and trained subjects, a Wilcoxon signed ranks test on total errors (counting wrong answers and ear location errors, which we called intrusions) yielded a significant difference between ears for the consonant-word recognition. (This was also true of the Thai subjects). No significant difference by this scoring was found for either group (untrained or trained)on either the tones or the hums. This is in contrast with the Thai subjects, who showed a significant right ear effect on the tone-words (but not on the hums).

Figure 1 summarizes the results of the three groups, musically untrained (u), trained (t), and Thais (T). The percentage of error scores for the left ear ("right ear advantage") are nearly identical in the untrained and trained English speakers in each task. In the tone-word task, both groups of English speakers perform differently from Thai speakers. In the consonant-words task, all three groups show a right ear advantage, and no significant difference between the groups is found. In the hums task, all three groups show no significant ear advantage.

When wrong-answer errors only (not intrusions) were subjected to statistical analysis, the trained (but not the untrained) subjects were found to have a left ear advantage for tone-words ($p < .05$) and for hums ($p < .01$).

Our experimental results do not support the hypothesis that musical training leads to the engagement of the left hemisphere in tasks involving non-linguistic pitch contours.

We did find support for Gordon's claim that musical sophistication correlates with degree of advantage, because (when wrong-answer errors only were analyzed in our data) a significant left ear superiority was found for tones and hums for the trained subjects only. However, the direction of the ear advantage was in contradiction to what might be predicted from either the Bever and Chiarello or the Gordon studies. This may be because our stimuli were pitch contours presented in single or doubled pairs, while the other two studies used melodies.

A summary of our findings can be seen in Figure 2, which presents the significance levels of right versus left ear percentage of error scores for the three subject groups on the three tasks.

Standard interpretation of dichotic listening results leads us to conclude that Thai speakers processed pitch contrasts which have a linguistic function more efficiently in their left (language) hemisphere. English speakers, regardless of musical training, did not show a left hemisphere preference for these same contrasts, but did show a significant ear difference for consonant-word contrasts as did the Thai subjects. The only effect of musical training in English speakers was to increase the degree of *right* hemisphere processing of tones and hums. Left hemisphere specialization was found for sounds only when they were part of a linguistic system in the perceiver.

Table I. 3 sets of stimuli used

(1) Tone-words

Stimulus	Tone	Length (ms)	English gloss
nāa	mid tone	625	"field"
nàa	low tone	650	(a nickname)
nâa	falling	575	"face"
náa	high tone	625	"aunt"
nǎa	rising	650	"thick"

(2) Consonant-words

Stimulus	Tone	Length (ms)	English gloss
dāa	mid tone	700	(a nickname)
nāa	mid tone	650	"field"
sāa	mid tone	700	"diminish"
cāa	mid tone	650	"tea"
lāa	mid tone	600	"goodbye"

(3) Hums

Stimulus	Tone	Length (ms)
—	mid tone	650
`	low tone	550
^	falling	550
´	high tone	575
ˇ	rising	525

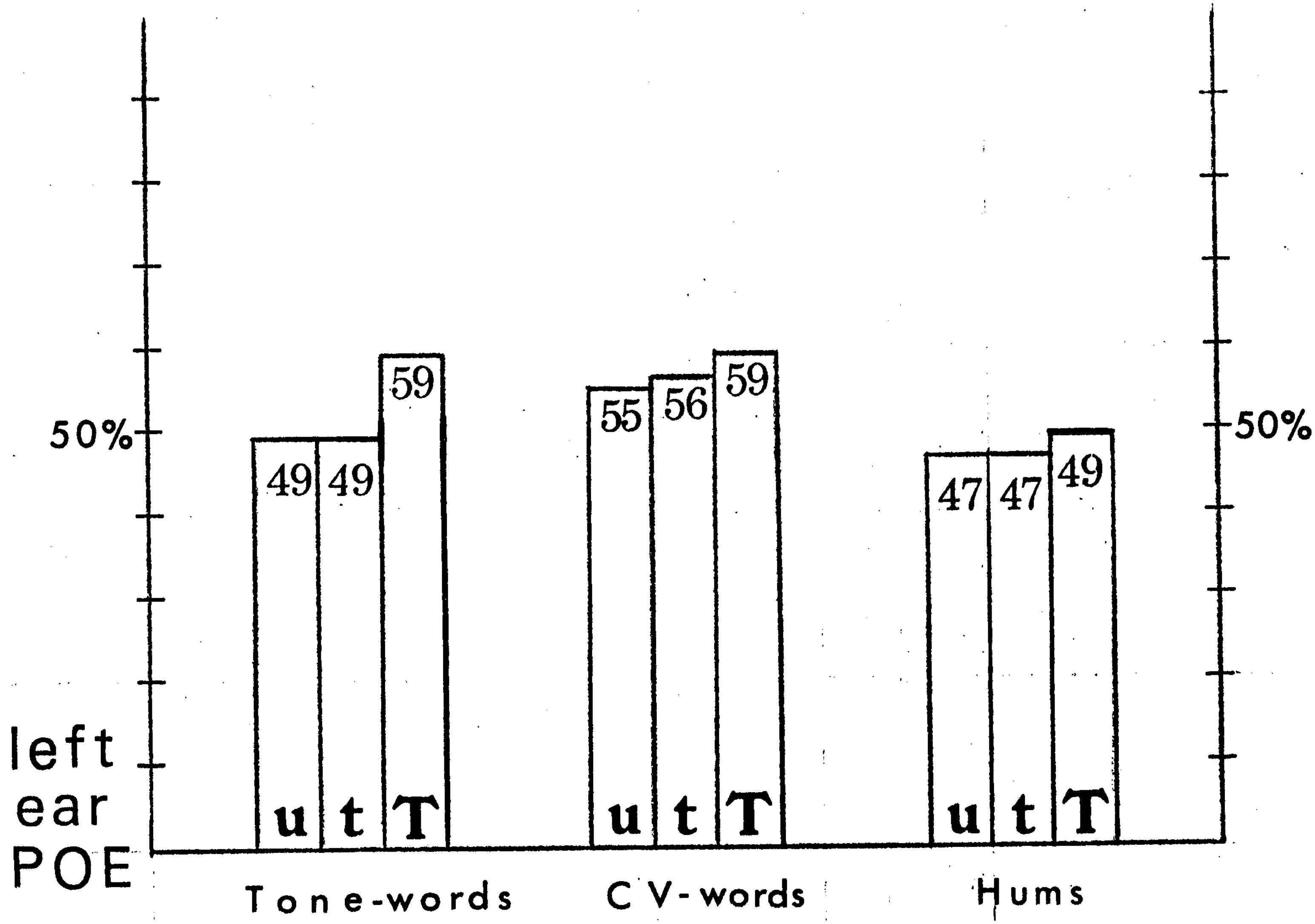


Figure 1. Percentage of errors (made at the left) for musically untrained (u) musically trained (t) and Thai groups (T) on the three dichotic listening tasks.

	tones	CVs	hums
u	NS	p<.05	NS
t	NS	p<.02	NS
T	p≪.01	p<.01	NS

Figure 2. Significance levels for ear difference in musically untrained (u), musically trained (t) and Thai groups (T) on the three dichotic listening tasks.

Rererences and Notes

1. Some of the literature on hemispheric function is reviewed, for example, in V. Mountcastle (ed.) *Interhemispheric Relations and Cerebral Dominance* (Johns Hopkins Univ. Press, Baltimore, 1962); J. Bogen, *Bulletin of the Los Angeles Neurological Societies 34,* 135 (1969; S. Dimond, *The Double Brain,* (Williams and Wilkins Co, Baltimore, 1972); S. Dimond and J.G. Beaumont (eds.), *Hemisphere Function in the Human Brain* (Paul Elek, Ltd, London, 1974); F. O. Schmitt and F. C. Worden, *The Neurosciences: Third Study Program* (MIT Press, Cambridge, Mass. 1974).

2. For reviews of dichotic listening research see C. I. Berlin, S. S. Lowe-Bell, J. K. Cullen and C. L. Thompson, "Dichotic Speech Perception: An Interpretation of Right-ear Advantage and Temporal Offset Effects" *Journal of the Acoustical Society of America 53.3,* 699 (1973); *Brain and Language* (Dichotic Studies I) 1.4 (October, 1974); J. Cutting, "Perception of Speech and Non-Speech, with Speech Relevant and Irrelevant Transitions" *Haskins Labs Status Report on Speech Research SR-33,* 37 (1973); D. Van Lancker, "Heterogeneity in Language and Speech: Neuro-linguistic Studies", *UCLA Working Papers in Phonetics 29,*(1975).

3. Spellacy and S. Blumstein, "The Influence of Language set in Ear Preference in Phoneme Recognition" *Cortex VI,* 430 (1970).

4. R. B. Chaney and J. C. Webster, "Information in Certain Multidimensional Sounds" *Journal of the Acoustical Society of America 40.2,* 477 (1966) F. Curry, "A Comparison of Left-handed and Right-handed Subjects on Verbal and Non-verbal Recognition Tasks: *Cortex III,* 343, (1967).

5. H. Gordon, "Hemispheric Asymmetries in the Perception of Musical Chords" *Cortex VI,* 387 (1970).

6. D. Van Lancker and V. A. Fromkin, "Hemispheric Specialization for Pitch and Tone: Evidence from Thai" *Journal of Phonetics* 1, 101 (1973).

7. T. G. Bever and R. J. Chiarello "Cerebral Dominance in Musicians and Nonmusicians" *Science* 185, 537 (9 August 1974).

8. H. Gordon, "Hemispheric Asymmetry and Musical Performance" *Science* 189, 68 (4 July 1975).

Tone loans: a question concerning tone spacing

and a method of answering it

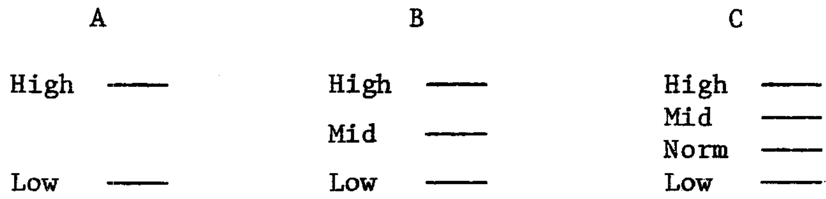
Ian Maddieson

1. Tone-spacing

An unanswered question relating to the nature of possible tone systems concerns the spacing of level tones. Is there a relationship between the number of contrastive level tones in a language and the amount of a speaker's pitch range which is employed to distinguish them? Two straightforward types of relationship between the number of tones and their spacing may be conceived and presented as extreme possibilities. One of these types of relationship would posit a more-or-less fixed interval between tones. In a language with only two level tones these tones would be separated by this interval. In a language with three level tones the highest and lowest would be separated by twice this interval, and so on. In other words, the pitch range used for paradigmatically contrasting tones would be determined by the number of tones in the language. The other type of relationship would posit that level tones are maximally separated in the speaker's available pitch range. In this case, the two tones in a language with only two levels would be produced towards the extremes of the pitch range. In a language with three level tones the mid tone would be produced in the center of the pitch range. Additional tones would subdivide the pitch range further. If this relationship holds, the pitch range used is assumed to be known, and it is the size of the interval between tones that is determined by the number of tones in the language. The 'spacing' being discussed is within an abstract phonetic space which represents a point of contact between physical and psychological reality. Syntagmatic variability of tones is not concerned here.

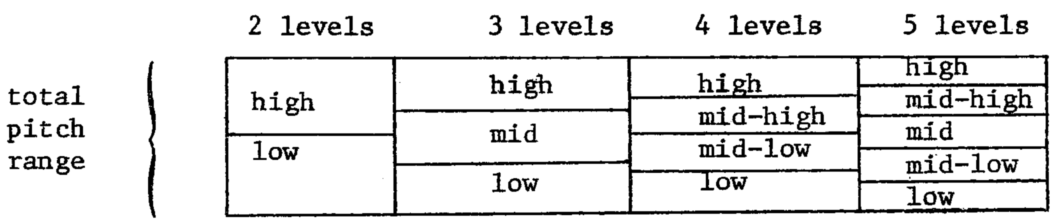
Where the issue has been discussed in the literature various linguists have offered views which seem to support both of the above positions, as well as some which are intermediate. In his book Tone Languages, Pike (1948) expressed his opinion that "a language with two registers tends to have the contrastive levels further apart than are the contrastive levels of four-register systems". He illustrates his view of this tendency with the diagram in (1).

(1)



The diagram indicates that, in Pike's view, the high and low tones in languages with two (A), three (B) or four (C) tones can be on the same level. In other words, the range is the same and the interval between tones is determined by the number of tones. Westerman and Ward (1933) seem to agree with this observation, to judge from the advice for learning tones which they offer to students of African languages, and a similar view seems to underlie the tone letter notation and numerical transcriptions for tones proposed by Chao (1930). Wang (1967) also proposed that larger numbers of level tones can be more narrowly separated than smaller numbers of tones. As he put it "the greater the number of distinct tones in the paradigm the narrower the phonetic range of each tone would be". The diagram he provides to illustrate his view can be redrawn as in (2).

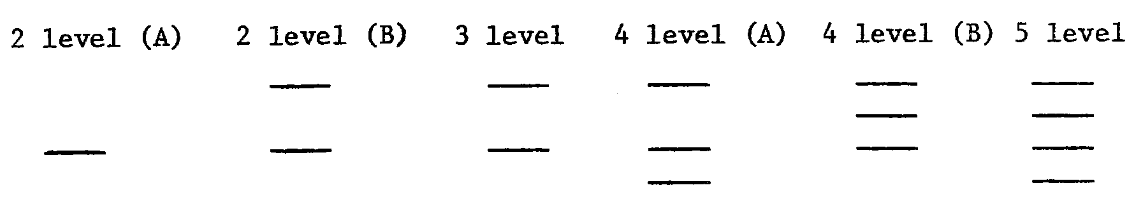
(2)



It is not clear if Wang is suggesting that the tones vary freely over the ranges indicated, or if he considers that fluctuation within the range is conditioned variation predictable from context, intonation, etc.

Woo (1969) suggests that in systems with two levels. However, in her view, systems with four or five level tones divide the upper and/or lower portion of the speaker's pitch range into smaller intervals. She actually suggests that an additional feature, such as a modified mode of glottal vibration, is required to distinguish the fourth and fifth tones in a system and that pitch alone is not sufficient, but if we construct a diagram just based on her view of the division of the pitch range it would have the form in (3).

(3)



In a discussion of universals of tone (Maddieson 1977) the present author proposed essentially that the first of the relationships outlined above is true. That is to say, there is a fixed interval between paradigmatically neighbouring tones and the range is determined by the number of tones (at least in the majority of languages). Possible tone systems would include those represented by the diagram (4) besides others not represented there.

(4) 2 level 3 level 4 level 5 level

The situation is summarized by (5) (from Maddieson 1977).

(5) A larger number of tone levels occupy a larger pitch range than a smaller number.

It is worth emphasizing again that what is under discussion here is the pitch range which is used for distinguishing tones that are in like contexts, i.e. abstracted from any syntagmatic variability. In connection with (5) it is assumed that languages with fewer tones will permit greater variability in their realization whether through superimposition of intonational contours, changes in pitch due to position in a sequence of tones or interaction with segmental structure.

The question is, do the languages of the world lean toward the Pikean or the Maddiesonian extreme in regard to the spacing of tones? (Or do they conform to an intermediate position, or simply vary one from another in regard to spacing of tones?).

If the languages of the world conform to Pike's diagram in (1) in the spacing of their tones, then it may be said that the tones are maximally separated in the available space. On the other hand if the tones conform to the spacing indicated by the diagram in (4) then it may be said that tones need only to be adequately separated by a standard interval. The principles of maximal separation or merely adequate separation may be posited as competing theories of the basic nature of phonological inventories, which also require positing different strategies for production and perception of phonological units.

The claim that phonological entities are maximally separated has been proposed, for example, to account for the spacing of vowels of different qualities (Liljencrants and Lindblom 1972). For a language with only three vowels in its inventory, this principle predicts that they will be at the apexes of the basic vowel triangle, i.e. /i, a, u/. In general this seems to be true (Sedlak 1969). The principle also seems to successfully predict tone languages requires special interpretation because of the usual presence of contour tones in the tonal inventories, and also because of the frequent cooccurrence restrictions between tones and consonants. Consequently the data presented here is largely from African languages. Nonetheless, an adequate number of languages to draw some conclusions is probably represented.

3. Loanwords taken from the same donor language

We will first examine loanwords from Hausa, a Chadic language spoken principally in northern areas of Nigeria. It was the language of the most influential political units in these areas in the nineteenth century. Subsequently, it became an official language of the Northern Region of Nigeria whose boundaries included speakers of very many other languages. It has played an important role in trade, religion, education and administration in the region. Consequently, many languages have loanwords from Hausa. We will describe the pattern of loaning in three of these - Gwari, Nupe and Kpan.

Hausa contrasts two level tones, High (H) and Low (L) (Abraham 1959). The sequences of these tones that are permitted in a word are relatively unrestricted, but words with some patterns (such as LL) are rare and others may be typical only of certain grammatical categories, e.g. plural nouns. There are no significant cooccurrence restrictions on the tones with segments of particular types. A third tone has often been recognized: this is a falling tone that usually appears with a short vowel (though often the syllable concerned has a geminate consonant or a consonant cluster). On various grounds the falling tone may be analysed as a sequence of high and low tones on the same syllable (HL). For example, the falling tone behaves like a sequence of H + L tones in relation to the processes of downdrift and high-raising. Meyers (1976) has shown that downdrifting applies to a sequence of like tones (all H or all L) and calls this a rule of like-tone lowering. However, when H and L tones alternate in sequence a steeper lowering pattern occurs. The process of high-tone raising increases the pitch of the last of a series of 2 or more H tones before a following L. When a falling tone follows H tone(s) the same raising of the beginning of the fall can be observed. This is illustrated in (7).

(7)	<u>H H L H</u>		<u>H HL H</u>	
	sá:bón gàrí	[¯ ¯ _ ¯]	fítôwá	[¯ ╲ ¯]
	'new town'		'coming out'	

The phonetic output of downdrift and like-tone lowering will be shown to have some importance in the loaning of Hausa words.

In each case the tones of the Hausa words which have been borrowed are those given by Abraham (1962). The first borrowing language examined is Gwari, classified by Greenberg (1963) as a langauge in the Kwa branch of Niger-Congo. The Hausa loanwords in Gwari have been extracted from Hyman and Magaji (1970). The dialect represented is that of Kuta, a city-state which maintained political independence until the end of the nineteenth century, despite owing some allegiance to the Hausa Sarkin Zazzau the location in this two dimensional acoustic space of the most commonly found five and seven vowel inventories /i, e, a, o, u/ and /i, e, ɛ, a, ɔ, o, u/. It is less easy to see if successful predictions are made about consonant inventories, perhaps because the dimensions involved are less well understood. Considering particular subsystems of consonants simplifies the question somewhat. For example, for nasals one might posit a 'place of articulation' dimension. In this case the maximally distinct pair of nasals would be the bilabial and uvular pair /m, N/. However, a more usual system of two nasals contains a bilabial and an alveolar or dental member /m, n/. For reasons of this sort it is difficult to maintain that consonants are constrained to be maximally separate (cf. the discussion of place of articulation in Ladefoged 1971). There is, however, evidence that when two consonants are in a similar area of the articulatory space they tend to make maximal use of the part that is available to them (Ladefoged (1975). For example, the labiodental fricatives in Ewe are usually made with the upper lip raised so as to make these sounds more distinct from the bilabial fricatives.

As far as tones are concerned, maximal separation would imply that articulation is the major factor in determining tone-spacing, since the size of the total pitch range is constrained much more by physiological limits on pitch production than by limits on pitch perception. On the other hand, adequate separation would imply a much greater role for perception, since discriminability must be a major factor in determining the size of an adequate interval.

2. Tone-loans as evidence for tone-spacing

Because of the relative nature of pitch contrasts there are difficulties in determining whether tone levels are maximally separated or not. Some idea of the spacing of tones can be obtained from comparing the intervals between the fundamental frequencies of tones measured in various instrumental studies of tone languages (cf. Maddieson 1977). However, this technique is unsatisfactory as there is no control for individual characteristics of speakers, and the contexts of the tones measured differ from study to study. Also differences of experimental technique and measurement procedures introduce other variables which can affect the actual numbers. Consequently, an interval of, say, 15 Hz in one study may not be equivalent to an interval of 15 Hz in another case. A series of instrumental studies could be made which would standardize the form of the utterances as far as possible and standardize all the technical aspects of the procedures used, but the individual variable could only be eliminated from playing a distorting role if a large number of speakers from each language could be studied. Other approaches to the question could involve measuring the contrast between tone intervals in the speech of those who are able to speak two or more tone languages fluently, or creating 'instant bilinguals' by teaching words from one tone language to speakers of another with a typologically different tone system.

This paper reports on another technique. It examines loan words which have been taken from one tone language into another. When words from a language with fewer level tones are taken into a language with more tones the loan correspondences between tones show which intervals are treated as equivalent. As an illustration, consider a disyllabic word with a High-Low tone sequence in a language with only two tone levels which is taken into a language that has three level tones. In this three level system this word might be treated as a High-Mid or Mid-Low word, as in (6a or b), or as a High-Low word, as in (6c).

(6)

	2 level language		3 level language
a)	H L	⟶	H M
b)	H L	⟶	M L
c)	H L	⟶	H L

If the result of such borrowing is as in (6a) or (6b) then the interval between H and L in the two tone language is treated as a smaller interval than the high to low interval in the three level language. In other words, something like a standard interval between adjacent tones describes the relationship of tones in the example. On the other hand, if the result is as in (6c), then the evidence shows that the two levels in the two level system are separated as widely as the most widely spaced tones in the 3 level language. In other words, this pattern of borrowing suggests that the tones are maximally distinct. From examining a number of words from a number of pairings of languages it should be possible to detect if there is a consistent tendency for one or the other of these two situations to be the case and hence to answer the question about the spacing of tones that was asked at the outset of this paper. Note that since this technique involves looking at the phonological tones of the loan-words, not at actual pitches, the normalization of individual differences is implicit in the data. An agreed perception of the tones of the donor language has been reached in the process of assigning a phonological representation to the words which have been adopted in the recipient language.

In order to be reasonably sure that a general principle is being discovered it is desirable to have the same language represented as the donor in more than one pairing, the same language represented as the recipient in more than one pairing, and some pairings which are quite independent of each other (i.e. have no language in common). There are also important considerations of adequacy of data involved. A sufficient number of words to suggest a reliable equation is desirable. The phonologies of the languages involved must be known. It must be ascertainable that the words involved are in fact loanwords and not historically inherited cognates and it must be clear what the direction of loaning is.

Relatively few cases satisfy these requirements of adequacy of data. No data is available from American tone languages, although it is surely possible that there are, say, Apachean loanwords in Tanoan languages or Yucatec loanwords in Otomanguean tone languages. The data from Asian at Zaria and later at Abuja. Undoubtedly a closer contact with Hausa existed at an earlier date in more northerly Gwari-speaking areas and it is possible that Hausa words have entered Kuta speech from these other dialects (cf. Temple 1919).

There are three basic level tones in Gwari, high (H), mid (M) and low (L). All possible sequences of tones are found, although the vocalic prefixes of nouns may only bear L or M (as most Hausa words are consonant-initial this cannot affect the pattern of borrowing in most cases). A mid tone following a low tone is heard as a lowered mid tone. A high tone following L is heard as a low to mid rise. An examination of the vocabulary in Hyman and Magaji shows that the lowered mid is also heard as the usual variant of M in a word-initial CV syllable when H follows. There are some exceptional cases where a lowered mid occurs on the second syllable of a word even when no L precedes, or where a phonetic level high tone is heard even when L does precede, but these exceptional tone patterns do not generally occur in the words borrowed from Hausa. The loanwords recorded for Gwari are listed in detail in Appendix I. A summary of the tone correspondences which emerge from these loanwords is given in Table 1. In this table, as in the others that follow, the tone patterns of words in the donor language (in this case Hausa) are listed in the first column. In the second column the tone patterns in the borrowed forms are listed. In the third column two numbers usually appear, separated by a slash. The number of words with the particular tone pattern in the donor language that have been borrowed is given after the slash. Before the slash, the number of those words that have the corresponding tone pattern in the recipient language is given. In many instances these numbers are equal. Where they are unequal this is because either some of the borrowed words have entered the recipient language with tone patterns that are difficult to equate with the tone pattern in the donor language, or, in cases where a relatively large number of words with the same tone pattern in the donor language have been borrowed, it is possible to select one pattern in the recipient language as typical and disregard the others as exceptional. In a few cases, several tone patterns in the recipient language are listed as corresponding to one pattern in the donor language. In these cases a brace replaces the slash before the second number. Further details may be found in the appendices.

It may be seen from Table 1 that, with few exceptions, the interval between H and L in Hausa is treated as the equivalent of the interval between H and M or between M and L in Gwari. H in Hausa words containing no L tones is borrowed as Gwari M. Gwari therefore provides evidence against a theory of maximal spacing of tones. The only apparent exception comes in the final two syllables of Hausa HHL words, borrowed as Gwari MHM. This exception is, however, explicable as a result of the syntagmatic effects of the high-tone raising rule in Hausa described above. This rule raises the second H in the HHL sequence so that an interval is created between the first and the second H. Gwari represents this as a step up from M to H and then represents the following HL interval as a step down from H to M, just as the interval in disyllabic Hausa

Table 1

Tone correspondences of Hausa loanwords in Gwari

	Hausa tone pattern	Gwari tone pattern	number	/	out of
level	H H	M M	7	/	7
patterns	H H H	M M M	2	/	2
	H H H H	M M M M	2	/	2
	L L L	L L L	2	/	2
descending	H L	H M M L	8 3	}	15
patterns	H H L	M H M	5	/	6
	H L L	M L L	1	/	1
	H H H L	H H H M	1	/	1
ascending	L H	L M	11	/	16
patterns	L L H	L L M	6	/	6
	L H H	L M M	6	/	8
ascending and descending	L H L	L M M L M L L M H M	3 2 3	}	8
	L H H L	M H H M	1	/	1
	L H L H	L M L M	1	/	1
	(H L H	M L M	2	/	6)
	H H L H	L M L M	1	/	2

words with HL is usually represented as HM in Gwari.

Nupe is a fairly closely related language to Gwari, spoken in an adjoining area to the west. Although trade with Hausaland dates from earlier times, extensive contact with the Hausa language probably did not occur until after the reformist Islamic Sokoto Caliphate had consolidated its influence over Nupeland in the mid nineteenth century (Nadel 1942, Mason 1970). Consequently Hausa loanwords in Nupe and Gwari are independently borrowed well after the separation of the two languages.

The Nupe tone system is described by Smith (1967) and George (1970) as containing three level tones (H, M, L). Contour tones arising from contractions of level tones are also possible. An important rule produces a rising variant of H after L and a voiced consonant. Tone combinations are unrestricted except that the noun prefix /e-/ may only have L or M. The loanwords from Hausa in Nupe have been extracted from Hyman (1970) and are recorded in detail in Appendix II. A summary of the tone correspondences is given in Table 2.

The Nupe data is not quite so easy to interpret as the Gwari. It is obvious that the well-represented Hausa HL and LH patterns are normally represented as Nupe HM and LM, i.e. by a small interval, and H tones in words with no L are represented by Nupe M. However, in words with tone patterns which contain both an ascending and a descending interval it is not uncommon for one of these intervals to be represented by a 'single step' in Nupe tones and the other as a 'double step'. Thus the usual borrowings of Hausa LHL and HLH are Nupe LHM and HLM. The downdrift rule in Hausa probably explains why Hausa HLH is borrowed as Nupe HLM. This rule ensures that a markedly smaller interval separates the medial L from the final H than separates the initial H from the medial L, i.e. Hausa HLH = [¯ _ -] (in fact the final H can even be lower than the medial L (Meyers 1976)). Nupe has no comparable downdrift rule and the HLM pattern probably represents an attempt to capture the difference between the HL and LH intervals in Hausa *when they appear in the same word.* In the two words where the Nupe pattern representing Hausa LHL is MHL a similar explanation could apply (Hausa LHL = [- ¯ _]). But it is unclear why Hausa LHL should more often be represented by Nupe LHM.

Neither of the competing theories of tone spacing can explain why the initial but not the final tone in a Hausa LHL word is represented as Nupe L, particularly as the LH interval in the LHL sequence is phonetically smaller than the HL interval and the Hausa HL in disyllabic words is Nupe HM. But overall, albeit with this rather puzzling example, Nupe also provides evidence against the hypothesis of maximal spacing of tones.

A third language into which Hausa borrowings have been taken is the Jukunoid language Kpan, spoken well to the east of the Hausa heartland. The Kpan were included in the Jukun 'empire' of Kororofa, subsequently conquered from the east by the Chamba and partially under the control of the Fulani emirate of Adamawa until colonial times (Meek 1931a, 1931b).

Table 2

Tone correspondences of Hausa loanwords in Nupe

	Hausa tone pattern	Nupe tone pattern	number	/	out of
level patterns	H H	M M	5	/	5
	L L	L L	1	/	1
descending patterns	H L	H M	7	/	7
	H H L	H H M	2	/	3
	H L L	HL L	1	/	1
	H H H L	H H H M	1	/	1
ascending	L H	L M	5	/	7
	L L H	L L M	2	/	3
	L H H	L M M	1	/	1
ascending and descending	L H L	L H M M H L L H L	7 2 2	}	13
	L H H L	L M M L L H H L	1 1	}	2
	L H L H	L H L M	1	/	1
	H L H	H L M	8	/	9
	H L H L	M L M L	1	/	1
	(H H L H	H H L L	1	/	1)

Hausa loans therefore are of recent origin, probably having largely entered the language in recent decades because of the status of Hausa as an officially recognised language. The loanwords from Hausa into Kpan have been extracted from Shimizu (1972). Shimizu (1971) describes the tone system as one with three level tones (H, M, L). The three tones may freely combine in all possible sequences, although the prefix vowels of nouns are generally L before L or M and M before H. The list of loanwords is given in detail in Appendix III. A summary of the tone correspondences between Hausa and Kpan borrowings is given in Table 3.

In Kpan there is some hesitancy between representing the Hausa interval between tones as one step or two steps. Also, the two HH words come in as HH but the two HHH words come in as MMM. We may count the number of times the Hausa interval is represented in the two different ways in Kpan, disregarding doubtful instances of borrowing. Of the cases where a descending interval is included in the Hausa tone pattern, ten move down one step in Kpan (from H to M or from M to L) and eight move two steps down (from H to L). Of the Hausa ascending patterns there are thirteen which move up one step and five which move two steps up. Thus the cases in which the Hausa tone interval is represented as one step in Kpan tones do outnumber the cases in which it is represented as two steps. Furthermore a number of these two step intervals are readily understandable. Thus, as in Gwari, the loaning of Hausa HHL and HHHL patterns as MHL and MMHL in Kpan is a consequence of the high-tone raising rule of Hausa; the interval from the raised H to L in Hausa is represented by a 'double step' in Kpan but between the initial tone and the final tone of these patterns in Kpan there is only one step, from M to L. If we consider these two patterns as representing essentially a one-step interval we have the numbers shown in (8).

(8)

	interval = one step	interval = two steps
descending	12	6
ascending	13	5

These numbers make it clear that it is prevalent for the Hausa interval between tones to be treated as equivalent to the interval between H or M or M and L in Kpan, although there are some exceptions, particularly in the case of Hausa HL words. A theory of maximal separation of tones is less successful in predicting this pattern than a theory of a standard interval between tones.

4. Loanwords received from different languages

The loans from Hausa into Gwari, Nupe and Kpan fulfil the requirement that loans from the same source into a number of different languages should be examined. This is one way to provide a check against correspondences that result from a fortuitous coincidence of subsidiary phonetic properties of tones in a given pairing of languages. Another way of providing the same check is to find instances where a given language has borrowed from more than one language. Southern Africa pro-

Table 3

Tone correspondences of Hausa loanwords in Kpan

	Hausa tone pattern	Kpan tone pattern	number	/	out of
level	H H	H H	2	/	3
patterns	H H H	M M M	2	/	2
	L L	L L	1	/	1
	L L L	L L L	3	/	3
descending	H L	H L	3	}	9
		H M	1		
patterns		M L	2		
	H H L	M H L	1	/	1
	H L L	H M M	1	}	2
		H L L	1		
	H H HL	M M H L	1	/	1
ascending	L H	L M	3	}	6
		M H	1		
patterns		L H	1		
	L L H	L L H	2	}	3
		L L M	1		
	L H H	L M M	2	/	2
ascending and	L H L	L M L	2	}	5
		L H L	1		
descending		L M M	1		
	H L H	M L M	1	}	3
		H M H	1		
		H L H	1		
	H L H L	M L M L	1	/	1

vides an example of this type, involving borrowing from two different Bantu languages into the Žu|'hõasi dialect of !xũ (Khung), a Khoisan language spoken in South West-Africa (Namibia). Unfortunately the number of words involved is not large, but the case is interesting because the !xũ tone system as described by Snyman (1975) has four level tones, here called Extra High (EH), High, Mid and Low. There are minimal contrasts between all four tones and most sequences are found to occur. The M tone is the most frequent in the vocabulary that Snyman has examined, followed in descending order of frequency by H, L and EH.

The first source of borrowed words in !xũ is the Sotho-Tswana group of dialects. Contact has been principally with Tswana members of this group, speakers of !xũ and Tswana having been involved in recent times in a kind of master-serf relationsnip (Silberbauer and Kuper 1966). In addition workers from both groups have often been employed together on farms in South-West Africa. The loanwords have been extracted from Snyman's !xũ-Afrikaans dictionary, which identifies the Tswana words concerned but does not mark their tones. Tone-patterns for the Tswana words have therefor been sought in a variety of sources (Cole and Mokaila 1962, Cole 1969, Ziervogel 1961, and even Jones 1927, 1928 and Tucker 1929) and checked with Thula Mndaweni, a speaker of a Southern Sotho dialect. The Sotho-Tswana dialects have two level tones. An important rule raises L to H after H in many environments (Cole 1969). Consequently almost none of the loanwords contain a descending tone pattern in !xũ, even when the Sotho-Tswana word from which it comes contains a HL sequence in its underlying tones. The words involved are listed in Appendix IV and the tone correspondences summarized in Table 4. The table shows that none of the borrowed words contains EH in !xũ. The sequence LH in Sotho-Tswana is stable and this sequence is represented as an interval of only one step in !xũ (L to M or M to H). The fact that three of the correspondences to Sotho-Tswana HL contain an ascending step in !xũ may be due to the penultimate stress rule in Sotho-Tswana which results in a lengthened vowel, possibly with a rising onglide, on the H syllable of a disyllabic HL word (cf. Jones 1927, 1928). In particular, the two words that represent this stressed vowel as a geminate with contrasting tones in !xũ seem to indicate that the lengthening played a role in determining the shape of the borrowed word. These words are given in (9).

(9)	gloss	Tswana	!xu
	'hill'	thótà	tùūtā
	'pound, beat'	bátà	bāáthá

The two Sotho-Tswana HL words which come into !xũ as HH amd MM are fairly straightforwardly derived from forms to which the rule HL → HH has applied, while MH might derive from patterns like those in (9) with contraction of the geminate vowel.

Insofar as these borrowings provide evidence, they provide support for the view that a standard not a maximized interval separates paradigmatically neighbouring tones. Maximal separation would predict the

Table 4

Tone correspondences of Sotho-Tswana loanwords in !xũ

	Sotho-Tswana tone pattern	!xũ tone pattern	number	/	out of
level patterns	H H	H H	3	/	3
	H H H	M M M	2	/	2
	L L	L L	2	/	3
	L L L	L L L	1	/	1
descending patterns	H L	H M	1	}	6
		M M	1		
		H H	1		
		M H	1		
		L M M	1		
		M H H	1		
	H H L	L M	1	/	1
ascending patterns	L H	L M	5	}	7
		M H	1		
	L H H	M H H	1	/	2

occurrence of EH in the loans and a larger interval in the words borrowed from Sotho-Tswana LH.

The second language from which borrowings into !xũ are noted by Snyman is Herero. Herero is from a quite different subgroup of Bantu from Sotho-Tswana but it also has a tone system with two level tones (Köhler 1958, von Essen 1971). Unfortunately Herero dictionaries (Irle 1917, Meinhof 1937) do not mark tones and the specific tone-patterns on the Herero words involved cannot be identified at present. However, again none of the words contain EH tones in !xũ, which would be expected to appear if tones were maximally separated, and in general there is no more than one step between their tones. Thus the Herero borrowings appear consistent with the observations on the other cases discussed above.

In connection with the discussion of !xũ it seems appropriate to quote a remark by Lanham and Hallowes (1956) on a related language spoken in South Africa. On the basis of minimal contrast they recognise three level tones in Eastern Bushman, then additionally argue that:

> "another factor lending support to the recognition of three level tonemes is the unusually big fall from the tones marked as high tones .... to those marked as low tones. This fall is noticeably greater than that heard between the high and low tonemes in Nguni and Sotho (with their ditonemic structure) ..."

5. Other data

In addition to the cases reported above there are also a couple of rather fragmentary instances of loanwords reported from other African languages. M. Trifković has kindly provided me with some data from the Lendu dialects she has studied. Lendu is in the Central Sudanic branch of Nilo-Saharan. A principle dialect called Baleða has four tone levels, H, M, L and Extra Low (EL). There are lexical and syntactic restrictions that are not entirely clear on the two lowest of these (Trifković 1977). Another dialect, ʒiða, spoken by the Bahema perople (a group of Bantu origin who have adopted Lendu as their language) has only three level tones, not distinguishing the two lowest levels. Some ʒiða words have been borrowed into Baleða. These words can be shown to be loanwords because the original Baleða lexical items can still be traced. The two upper levels of ʒiða and Baleða correspond straightforwardly (ʒiða M becomes Baleða M and ʒiða H becomes Baleða H), but ʒiða L becomes Baleða L, not EL, except where the morphemic category of the word requires otherwise. Thus, the verb /ra/ 'go' and the plural pronoun /ma/ 'we, us' must take EL tones, but nouns and adjectives which may bear either L or EL take Baleða L to correspond to ʒiða L, as in /hwàhwà/ 'slender', /lè/ 'woman', /yàyá/ 'mother'. In sum, the lowest tone in ʒiða is not represented by the lowest tone in Baleða in cases where there is any freedom to choose.

Canu (1968) reports on loanwords from a variety of sources in Mò:re, a Voltaic language with three level tones spoken in Upper Volta. A few of these words have come from Bambara (or one of the other dialects of the Mandekan language), which has two tone levels (Courtenay 1974), or from Songhai, which, although described by Williamson (1967) as not a tone language "in the usual sense", can nonetheless be described in terms of two pitch levels. Among the loans are the words in (10).

(10)

donor language	original word	Mò:re form	gloss
Bambara	bárká	bārkà	'blessing'
Bambara	kàrȁ	kàrè-m (-m = class suffix)	'reading'
Malinke (cf. Susu xìbárù)	kìbàró	kībá-rē (-re = class suffix)	'story'
Songhai	wato, wakato (with high level pitch)	wākātò	'hour, time'

These few items from Lendu and Mò:re would not be persuasive on their own, but they are consistent with the data from other languages that show that a theory that tones are maximally separated is unable to explain the patterns generally found in borrowed words.

6. Apparent counterexamples to a standard interval

No persuasive cases have been found which argue for the theory of maximal separation of tones. The accumulated data is generally compatible with the theory that tones are separated by a standardized interval. There are, however, two cases that seem to be inconsistent on the surface with the data so far presented. These concern loans from Chinese into other Asian languages.

One of these cases involves recent loans from Swatow (Chaochow, Chaozhou) Chinese into Standard Thai (Siamese). Egerod (1959) lists close to two hundred loanwords belonging to this layer of borrowing. The tone correspondences are also tabulated and discussed in his article, so the wordlists and correspondences are not included here. Egerod describes a Swatow tone system with two level tones, H and L, and four contour tones, high falling, low falling, high rising and low rising. Thai has three level tones, H, M and L and one rising and one falling tone. Usually the Swatow underlying H comes in as Thai H and the Swatow underlying L comes in as Thai L. Because neither of these Swatow level tones comes in as Thai M, it seems that they are treated as being as far apart as the H and L tones in Thai. However, a number of things suggest that the Swatow tone system described by Egerod is not the correct one for the source of the loans in Thai. In the first place, as Egerod points out, the Thai mid tone appears in many of the borrowed words that have an underlying high rising tone in the dialect of Swatow he describes, whereas in Fielde's dictionary of Swatow (1883) this tone (shang ping) is described as level. Thus, the relevant Swatow subdialect appears to be one which had three level tones, the same number as Thai. Further evidence that Egerod's tonal system is not the right one for the source of the loans comes from the fact that the short syllables with final stops in Swatow come into Thai with H tone regardless of whether the tone in Egerod's dialect was H or L. Since Thai permits both H and L on syllables of this shape it seems as if the relevant Swatow dialect must be one that had lost the tonal contrast in these syllables. It may also be noted that the high rising tone derived from other tones in sandhi positions in Egerod's Swatow does not correspond with Thai M, as the underlying high rising tone does. Instead, it corresponds to Thai falling tone. This suggests that this sandhi tone must have had a different shape in the relevant dialect, perhaps being a low falling tone, as the underlying low falling tones in Swatow come into Thai as falling tones. It thus seems fairly certain that Thai borrowed from a phonetically different Swatow dialect, unlike the one described by Egerod in certain tonal characteristics. As this dialect had three level tones, like Thai, no evidence on the question at issue in this paper can be obtained from this case.

A second possible case not consistent with the foregoing data would seem to be the Yanghao dialect of Miao, spoken in K'ai li *hsien* of Kweichow, China. Yang hao (Ying 1972) has four level tones, as well as two rising and two falling tones. It has borrowed extensively in recent times from Chinese and it is presumed that these loans come from the local variety of Yunnanese (Southwestern Mandarin) which has two level tones, as well as a low rising and a low falling tone. Using the usual five-digit scale for transcribing tones (Chao 1930) the Yanghao level tones are, in descending pitch, 55, 44, 33 and 11. The lower of the two level tones in Chinese is borrowed as Yang hao 33, but the higher level tone is not represented by the next highest tone (44) but by Yang hao 55. In this case it would seem that a wrong prediction is made by a theory of standard intervals between tones. However, there are restrictions on co-occurrence of initial consonants and tones in Yanghao syllables. The 55 tone may appear only with voiceless initial consonants (including voiceless aspirated). In the local Kweichow Mandarin dialect the higher level tone may appear with voiced initials (mainly liquids and nasals) as well as with voiceless initials. It seems that rather than resolve the question by dividing the Chinese loanwords between the 44 and 55 tones according to the voicing of the initial consonant, Miao speakers (many bilingual, according to Ying) were aware that these words bore the same tone in Chinese. The tonal unity of this group of words was thus preserved by assigning them all to the same category in Yang hao. It is not clear what factor(s) may have led to the chosen category being the higher of the two tones available. All that may be said is that this case contains a complicating factor which makes it difficult to use it as evidence to decide the question at issue in this paper.

7. Conclusions

This paper shows that in general a theory of standard intervals between tones correctly accounts for most of the data on tone loans presented. On the basis of this evidence we may posit that it is generally true for the tone languages of the world that level tones are paradigmatically separated by a standard interval, although it would not be inconsistent with the evidence to posit that a small decrement in the size of this interval occurs as the number of level tones increases. This might explain the hesitancy sometimes observed between different tone patterns in a group of loanwords that had the same pattern in the donor language. A theory of maximal separation of level tones does not account for the data presented and consequently may be viewed as false.

This conclusion, in the first place, makes predictions about the kind of linguistic interference and language-learning difficulties that might arise for speakers of tone languages and about other cases of loaning. But, more significantly, it takes an important step towards defining the constraints on overall tone inventories and provides significant input to the process of building models of the production and perception of tones. Selection between competing proposals for phonological features of tone must take account of it. For example, a feature set containing [Extreme], as proposed in Maddieson (1970), is better able to explain why the highest tone in a system of two levels is not equated with the highest tone in a system with four or five levels than are the proposals by Wang (1967), Woo (1969) and others.

The paper also shows that tone rules, especially those which affect the syntagmatic relationships of tones, importantly affect the way in which tones are perceived and reproduced in the process of borrowing. Because other factors besides the number of paradigmatically contrastive tone levels are involved in determining the way in which a tone system with fewer levels is represented in loanwords in a language with more levels, indirect confirmation is provided of another basic condition on tone languages. Namely, languages with less elaborate tone inventories permit more syntagmatic variability in the realisation of their tones.

References

Abraham, R.C. (1959) *The Language of the Hausa People* University of London Press, London.

Abraham, R.C. (1962) *Dictionary of the Hausa Language* (2nd edition) University of London Press, London.

Canu, Gaston (1968) "Remarques sur quelques emprunts lexicaux en Mò:re" *Journal of West African Languages* 5:25-34.

Chao, Yuan-ren (1930) "A system of tone letters" *Maître Phonétique* 30:24-27.

Cole, D.T. (1969) "Tonal morphology in Tswana" *Ethnological and Linguistic Studies in Honour of N.J. van Warmelo* Department of Bantu Administration and Development, Pretoria:: 39-45.

Cole, D.T. and D.M. Mokaila (1962) *A Course in Tswana* Georgetown University, Washington, D.C.

Courtenay, Karen (1974) "On the nature of the Bambara tone system" *Studies in African Linguistics* 5:303-323.

Egerod, Søren (1959) "Swatow loanwords in Siamese" *Acta Orientalia* 23:137-156.

Fielde, Adele Marion (1883) *Pronouncing and Defining Dictionary of the Swatow Dialect According to Syllables and Tones* American Presbyterian Mission Press, Shanghai.

George, Isaac (1970) "Nupe tonology" *Studies in African Linguistics* 1:100-122.

Greenberg, Joseph H. (1963) *The Languages of Africa* University of Indiana, Bloomington.

Hyman, Larry M. (1970) "The role of borrowing in the justification of phonological grammars" *Studies in African Linguistics* 1:1-48.

Hyman, Larry M. and Daniel J. Magaji (1970) *Essentials of Gwari Grammar* (Occasional Publication 27) Institute of African Studies, University of Ibadan, Ibadan.

Irle, J. (1917) *Deutsch-Herero Wörterbuch (Abhandlungen des Hamburgischen Kolonialinstituts* 32, Reihe B 18) Friederichsen, Hamburg.

Jones, Daniel (1927) "Words distinguished by tone in Sechuana" *Festschrift Meinhof* Friederichsen, Hamburg: 88-98.

Jones, Daniel (1928) *The Tones of Sechuana Nouns* (Memorandum 6) International Institute of African Languages and Cultures, London.

Köhler, Oswin (1958) "Tonggestalt und Tonmuster in Infinitiv des Verbum im Herero" *Afrika und Übersee* 42:97-110, 159-172.

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

Ladefoged, Peter (1975) *A Course in Phonetics* Harcourt Brace Jovanovitch, New York.

Lanham, L.W. and D.P. Hallowes (1956) "An outline of the structure of Eastern Bushman" *African Studies* 15:97-118.

Liljencrants, Johan and Björn Lindblom (1972) "Numerical simulation of vowel quality systems: the role of perceptual contrast" *Language* 48:839-862.

Maddieson, Ian (1977) "Universals of tone" in *Universals of Human Language* (edd. J.H. Greenberg and C.A. Ferguson) Stanford University Press.

Maddieson, Ian (1970) "The inventory of features required for handling tone....." *Tone in Generative Phonology* (*Research Notes* Department of Linguistics and Nigerian Languages, University of Ibadan 3/2-3) 3-18.

Mason, M.D. (1970) *The Nupe Kingdoms in the Nineteenth Century* Ph.D. thesis, University of Birmingham.

Meek, C.K. (1931a) *A Sudanese Kingdom; an Ethnographical Study of the Jukun-speaking Peoples of Nigeria* Kegan Paul, Trench, Trübner, London.

Meek, C.K. (1931b) *Tribal Studies in Northern Nigeria* (Vol II) Kegan Paul, Trench, Trübner, London.

Meinhof, Carl (1937) *Die Sprache der Herero in Deutsche Südwestafrika* (2nd ed.) (Deubche Kolonialsprachen 1) Reimer, Berlin.

Meyers, Laura (1976) *Aspects of Hausa Tone* (*UCLA Working Papers in Phonetics* 32).

Nadel, S.F. (1942) *A Black Byzantium* Oxford University Press, Oxford.

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

Sedlak, Philip (1969) "Typological considerations of vowel quality systems" *Working Papers in Language Universals* (Stanford) 1:1-40.

Shimizu, Kiyoshi (1971) "The Kente dialect of Kpan: I Phonology" *Research Notes* (Department of Linguistics and Nigerian Languages, University of Ibadan) 4.2/3:11-36.

Shimizu, Kiyoshi (1972) "The Kente dialect of Kpan: III Lexicon" *Research Notes* (Department of Linguistics and Nigerian Languages, University of Ibadan) 5.1 pp 67.

Silberbauer, G.B. and A.J. Kuper (1966) "Kgalagari masters and Bushman serfs: some observations" *African Studies* 25:171-179.

Smith, N.V. (1967) "The phonology of Nupe" *Journal of African Languages* 6:163-169.

Snyman, J.W. (1975) *Žu|'hõasi Fonologie en Woordeboek* Balkema, Cape Town and Rotterdam.

Temple, O. (1919) *Notes on the Tribes, Provinces, Emirates and States of the Northern Provinces of Nigeria* Church Missionary Society, Cape Town (reprinted by Cass, London. 1965).

Trifković, Mirjana (1977) "Tone-splitting in Lendu" Paper presented at the 8th African Linguistics Conference, UCLA, April 1977.

Tucker, A.N. (1929) *The Comparative Phonetics of the Suto-Chuana Group of Bantu Languages* Longmans, London.

Von Essen, Otto (1971) "Bedeutungsbestimmende Silbentonhohen in der Sprache der Herero" *Afrikanische Sprachen und Kulturen - Ein Querschnitt* (ed. V. Six et al.) Deutsches Institut für Afrika-Forschung, Hamburg: 88-94.

Wang, William S-Y (1967) "The phonological features of tone" *International Journal of American Linguistics* 33:93-105.

Westermann, Dietrich and Ida C. Ward (1933) *Practical Phonetics for Students of African Languages* Oxford University Press for the International African Institute, London.

Williamson, Kay (1967) "Songhai word list (Gao dialect)" *Research Notes* (Department of Linguistics and Nigerian Languages, University of Ibadan) 3. pp 34.

Woo, Nancy (1969) *Prosody and Phonology* Ph.D. thesis, M.I.T.

Ying, Lin (1972) "Chinese loanwords in Miao" translated in *Miao and Yao Linguistic Studies* (ed. H.C. Purnell) Department of Asian Studies Cornell, Ithaca : 55-82.

Ziervogel, D. (1961) *Klein Noord-Sotho-Woordeboek* Van Schaik, Pretoria.

Appendix I

Hausa Loanwords in Gwari

Words are listed in groups according to the tone pattern in Hausa. The Hausa form is given first, followed by an English gloss, then the Gwari form, followed by a further gloss if the meaning is sufficiently different from the meaning in Hausa to require it. Comments on particular items follow after the group to which they belong.

H H

dúbú:	'thousand'	dūbū	
gádó:	'bed'	gōdō	
há:lí:	'character'	ālī	
hú:tú:	'resting'	fūtū	'holiday'
ká:yá:	'load' (noun)	kāyā	
núfé:	'Nupeland'	nūfēí	'Nupe person'
sá:tí:	'Saturday'	sātī	

The final -í in /nūfēí/ is assumed to be a suffix in Gwari.

H H H

máƙúllí:	'key'	mākūlī
núfá:wá:	'Nupe people'	ànūfāwāí

The initial à- in /ànūfāwāí/ is a plural noun prefix in Gwari; the final /-í/ in this word is taken to be the suffix noted above.

H H H H

mákárántá	'school'	mākārāntā
lá:rábá:wá:	'Arabs'	àlārābāwāí

Prefix and suffix in /àlārābāwāí/ as in /ànūfāwāí/ above.

H L

gó:dè	'give thanks'	gwódē
í:kò	'power'	yíkwō
kwá:nò	'bowl, basin'	kwánū
ƙárfè	'clock, hour'	kárfē, kólúfyē
láimà	'umbrella'	lémā
mú:dù	(a measure)	múdū
rí:bà	'profit'	gyíbā
ró:gò	'cassava'	rógō
gó:rò	'kolanut'	gōrò
súlài	'shilling'	sūlè, sīlè
té:bùr	'table'	tēbùl

One Hausa word /gárkè/ 'cattle-pen, herd' enters Gwari in the form /gārícē/ with MHM normally representative of Hausa HHL. Another HL word, /ʃáidà/ 'witness' is related to the Gwari /sèīdādā/, but both the LMMM pattern and the syllabic structure are puzzling.

H H L

bíndígà	'gun'	bīdígā
gá:fárà	'pardon, forgiveness'	gāfálā
júmmáʔà	'Friday'	ʒīúmā, àʒīúmā
lá:fí:yà	'health'	lāáfíyā
dábí:nò	'date (fruit)'	dàbīínō

The HL word /gárkè/ has been counted among the examples of this pattern, see above.

H L L

láhàdì	'Sunday'	lāàdì

H H H L

músúlúncì	'Islam'	músúlúncī	

L L L

àkwà:tì	'box'	àkwàtì	
àyàbà	'banana(s)'	àyàɓa	'bananas'

The initial /à-/ in Gwari /àyàɓà/ is treated as a plural prefix and a singular /yàɓà/ is formed on the model of /(ā)kútá/ 'rock(s)'.

L H

bàutá:	'slavery'	bàutā	
ɗàrí:	'hundred'	tàgyī	
fùré:	'blossom'	fùrē	'flower'
gwàfá:	'prop'	gwòfwā	'pillar'
jàká:	'bag'	ʒìkā	
jà:kí:	'donkey'	jàakī	
kò:kó:	'cocoa'	kwòkwō	
kyàutá	'generosity, gift'	kyàutā, càutā	
sò:só	'loofah, sponge'	sòsō	
yàngá:	'boastfulness'	yàngā	'bluff'
wàndó:	'trousers'	wàndō, wàidō	
kè:ké:	'bicycle'	kyēkyé, cēcé	
sànnú:	'greetings!'	hànú, sànú	
àmmá:	'but'	àmá	
gù:gá:	'bucket'	gwògá	
sò:sái	'well, correctly'	sòséi	'very'

Hausa /lè:mó:/ 'lime, lemon, etc' may be the source of Gwari /lòmwî/ but the vowels and tones display unusual correspondences.

The last three words in the list above have the irregular level variant of H after L in Gwari.

L L H

àmfà:ní:	'advantage'	àmpànī	'use' (noun)
kàrà:tú:	'reading'	kàràtū	'education'
mùgùntá:	'wickedness, evil'	mùgùntā	
mùsùlmí:	'Moslem' (noun)	mùsùlmī	
tù:rà:wá:	'white men'	tùràwā	
ùngùlú:	'vulture'	àgùlū, gùlū	

The form /gùlū/ for 'vulture' in Gwari probably arises from the perception of the initial vowel /à-/ as a detachable plural prefix.

L H H

àlká:lí:	'judge'	àlìkālī	
àsí:rí:	'secret' (noun)	àsīrī, àsīgyī	
gàskí:yá:	'truth'	gàicīyā	
ƙò:ƙárí:	'effort'	kùkōrī, kùkōgyī	'attempt'
nàsá:rá:	'Christians'	nàsārā	'Christian'
rà:ƙúmí:	'camel'	làkūmī	
gò:dí:yá:	'gratitude'	gwōdíyá	
tùrá:ré:	'perfume'	tùràrē	

L H L

àní:yà	'intention'	ànīyā	'determination'
fìtílà	'lamp'	fìtīlā	
là:bá:rì	'story'	làbārī, nàbāgyī	
àddúʔà	'invocation'	àdūwà	'prayer'
àsábàr	'Saturday'	àsībì	

àlbásà	'onion'	àlùbāásā	
bàtú:rè	'white man'	bàtūúrē	
tàlá:tà	'Tuesday'	àtàlāátā	

Beside the form given above for 'onion', Gwari also has /lúpūsā/.

L H H L

bàlá:rábè	'Arab person'	bālárēbē	

L H L H

àlká:wàrí:	'promise' (noun)	àlùkāwòlī	

H L H

ázùrfá:	'silver'	āzùrùfwā	
có:kàlí:	'spoon'	cōkàlī, ʃēkàlī	
jálàbí:	'self-serving act'	ʒālábā	'temptation'
ázùmí:	'fast' (noun)	ázúmǐ	
hánkàlí:	'intelligence'	ánkálī	'good sense'
tákàrdá:	'paper'	tákáàdā	'book'

Only in the first two of these words are the Gwari tone patterns a fairly straightforward reflection of the Hausa HLH sequence. The remainder are surprising, not least for their inconsistency.

H H L H

áláyyàfó:	'spinach'	àlēèfō	
áláyyàdí:	'palm oil'	àlēèdí	

Words with falling tone

âllí:	'chalk'	ālíyī	
lâifí:	'crime, fault'	lāìpí	'sin'
àlhàmîs	'Thursday'	àlàmî	
lìttìnîn	'Monday'	àtènî	

ƙwârƙwátà 'louse' kwōròkwōta

In two of these cases Gwari seems to represent Hausa falling tone as ML ('sin' and 'louse'). In final position a falling tone is apparently retained ('Thursday', 'Monday') but there is no native parallel to such a pattern and these forms may be recognized as not yet assimilated loans.

Appendix II

Hausa loanwords in Nupe

Words are listed according to the format described in Appendix I.

H H

gáɗó:	'bed'	gã̄dō
fú:ʃí	'anger'	fūʃī
káskó:	'bowl'	kã̄ʃīkō
ƙárƙó:	'last' (verb)	kã̄līkō
wúrí:	'open space'	wūrī

L L

dàgà	'from'	dã̄gã̄

H L

fúskà	'face'	fúʃíkā
hújjà	'excuse' (noun)	wújā
í:kò	'power'	yíkō
írì	'kind, species'	yírī
múlkì	'authority'	múlíkī
kâi	(exclamation)	káyī
râi	'life'	ráyī

Two Hausa monosyllables with falling tone are included here.

H H L

ígí:yà	'rope'	yígíyā
gáfákà	'school bag'	gã́fã́kā
júmmáʔà	'Friday'	jímâ

The final falling tone seen in the Nupe form for 'Friday' is an unusual occurrence very rare in the native vocabulary.

H L L

láhádì	'Sunday'	lâdè

H H H L

múná:fúncì	'hypocrisy'	mã́nã́fĩ́cī

L H

yàbó:	'thankfulness'	yèbō
bàutá:	'slavery'	bàwùtā
ƙàfó:	'horn'	kã̀fō
wàrkí:	'loincloth'	wòrìkī
àlló:	'board'	èló
wùyá:	'difficulty'	wūyá

L L H

ʃù:gàbá	'leader'	ʃìgàbā
kwàɗàyí:	'covet'	kòdã̀rī
màrhàbá:	'welcome!'	màràhábì

L H H

gàskí:yá:	'truth'	gã̀ʃìkīyā

L H L

àlbárkà	'blessing'	àlùbárīkā
àlbásà	'onion'	àlùbásā
àljánnà	'paradise'	àlìjénã̄
àní:yà	'determination'	àníyã̄
dàbá:rà	'skill'	dàbárā
mùrá:dì	'desire'	mùrádī
wàsíƙà	'letter'	wòʃíkā
àlhájì	'pilgrim'	ālāhájì

Appendix III

Hausa loans in Kpan

Words are listed according to the format described in Appendix I.

H H

dáidái	'correct'	dédé	
kwálbá:	'bottle'	kwábá	
gárwá:	'petrol can'	gārūwà	'oil drum'

The tone pattern of Kpan /gārūwà/ is surprising. Perhaps this word is not directly loaned from Hausa.

H H H

ká:tá:kó:	'plank'	kātākō	
dánkwárá:	'small cap'	dānkwārī	'headtie'

Shimizu suggests Kpan /dānkwārī/ is from Hausa /dán kwálí/ 'velvet' (literally 'son of bobbin'). The etymology given here is preferred but either yields the same tone correspondence.

L L

wà:tò:	'that is, i.e.'	wàtò

L L L

àyàbà	'banana'	àyàbà
àkwà:tì	'box'	àkwàtì
àdùdù	'lidded box'	àdùdù

H L

kwá:kwà	'oil palm'	kwákwà	'palmtree'
má:tò	'motor car'	mátò	
tá:bà	'tobacco'	tábà	
tébùr	'table'	tébù	

wáyà	'wire'	wáyà	
kwá:nò	'bowl, plate'	kwánū	'plate, zinc'
tá:sà	'plate, bowl'	tāsà	
kúdù	'south'	kūdù	
dánƙò	'rubber'	dōnkō	'catapult'

Because they are little changed in Hausa it is difficult to be certain that 'table' and 'wire' did not enter Kpan directly from English. Shimizu connects Hausa /túfà/ 'garment' with /ātūkpō/ 'cloth' in Kpan; this looks doubtful and it is not counted among the loanwords discussed in this paper.

<u>H H L</u>

fátányà	'small hoe'	pītínyà

<u>H L L</u>

kó: wànè	'whichever, every one'	kó wēnē	'every'
mángwàrò	'mango'	àmángòrò	

<u>H H ĤL</u>

kánánzîr	'kerosene'	kānāzíǹ

<u>L H</u>

àkú:	'parrot'	àkū
cìttá:	'pepper'	àʃìtā
ràké:	'sugar-cane'	àrìkē
lè:mó:	'lime, etc'	ārēmó
kè:ké:	'bicycle'	kyèkyé
gàngá:	'drum'	gàngán

The final /-n/ in Kpan /gàngán/ is of unexplained origin. Shimizu suggests that Hausa /kò:kó:/ 'calabash' used in the expression

/kòkwán kái/ 'cranium' (literally 'calabash of head') has been borrowed into Kpan in the expression /kòkò ìkyì/ 'skull'. This would be the only case where Hausa LH became levelled to LL.

<u>L L H</u>

tù:rà:ré:	'perfume'	tùràré	
àkà:wú	'clerk'	àkàwú	(personal name)
àgwà:gwá:	'duck'	àgwàgwā	

<u>L H H</u>

àní:ní:	'button'	ànīnī	
ʃùwá:ká:	(a shrub) vernonia amygdalina	ʃùwākā	'bitter-leaf'

Hausa /mùzú:rú:/ 'male cat' is connected with Kpan /dāmúsūú/ in some way but the first syllable in Kpan is of obscure origin.

<u>L H L</u>

sà:búlù	'soap'	sàpūrù
tùmá:tùr	'tomato'	tùmātù
àddú?à	'prayer'	àdúà
hàrʃá:ʃì	'bullet, shell'	àrìʃāʃī
àlbásà	'onion'	áràbàsà

The final vowel and tone pattern in Kpan suggest that 'tomato' was not borrowed direct from English as Shimizu suggests. The word for 'onion' in Kpan may not be from Hausa; this root is from Arabic and is widespread in West African languages.

<u>H L H</u>

árè:wá:	'north'	ārèwā
có:kàlí:	'spoon'	cókōrí

íyà:yé 'parents' íyàyé

H L H L

síngìláti̇̀ 'singlet' sīngùrītì

This word is considered a direct loan from English by Shimizu, but the specific shape of the Hausa form is much closer to the Kpan word.

L H L H

There is one Kpan form derived from a Hausa word with LHLH tone pattern. This is /tátábā/ from /tàntábàrá:/ 'pigeon', but the syllabic structure and tone pattern have been altered so much that this form is not useful for the purposes of this paper.

Appendix IV

<u>Sotho-Tswana loanwords in !xũ</u>

Words are listed according to the format described in Appendix I. In some cases it has been necessry to select between different forms in the Sotho and Tswana sources consulted to find the most probable origin of the borrowed form in !xũ.

<u>H H</u>

kókó	'chicken'	hōrōkókó	
káné	'can'	kʰani	'milk can'
kxáré	'headpad'	xárí	

'Chicken' also appears with HL pattern in Sotho /kxógò/.

<u>H H H</u>

gátísá	'stamp, beat'	kātīsā	'tame'
kxóbógá	'gather'	kxōbāhā	'scour off'

<u>L L</u>

ɬʰòkà	'lack'	tʰòkà	
ɬʰàgà	'appear'	tʰàhà	'arrive, jump'
sèmè	'whip' (noun)	sāmí	

<u>L L L</u>

bèlèʃà	'carry on back'	bàrìsà	'load'

<u>H L</u>

púdì	'goat'	párí, bārī	
tʰótà	'hill'	tùūtā	
bátà	'pound, beat'	baátʰá	
bò-lódì	'inner bark'	dūrí	
dì-bólò	'ball'	búrū	

The Sotho-Tswana word for 'ball', borrowed from Afrikaans or English, seems to have a more common variant with LL on the stem. The !xũ loanword looks more like a borrowing from a HL stem, but as this group of words is very inconsistent this remains uncertain.

H H L

gádíkà	'roast, fry'	xàrī

The loss of the final syllable in !xũ is unexplained.

L H

kxòlé	'thong, rope'	xòrē	
màdí	'blood'	màdī	'money'
mòá	'breath, spirit'	mòā	
nyàtsí	'lover'	nyàtʃī	
pʰòló	'ox, bull'	pʰòrō	
tàú	'lion'	tàū	
tswàyá	'mark, brand'	tʃwà	

L H H

sèkóló	'school'	sīkóré
bòrókyó	'trousers'	búrúgō
bòrókywé	'trousers'	búrúkʰōe

The two forms of the word for 'trousers' in both languages is an oddity. The original word is Afrikaans.

L H L

One Sotho-Tswana word with LHL pattern is borrowed into !xũ with MM. This is /ɬʰáólà/ 'share out, divide' which enters !xũ as /tʰārō/ 'separate'. This word is not useful for the purposes of this paper.

Tone-spreading and perception[1]

Ian Maddieson

[Paper presented at the 92nd Meeting of the Acoustical

Society of America, San Diego, California,

November, 1976]

Among the various phonological universals of language that have been proposed are two that concern the phenomemon of 'tone-spreading' in tone languages. Examples of tone-spreading rules are illustrated in Figure 1, where H represents a high tone syllable and L represents a low tone syllable. Schematic representations of the pitch contours in these rules are given by continuous lines on the right-hand side of Figure 1.

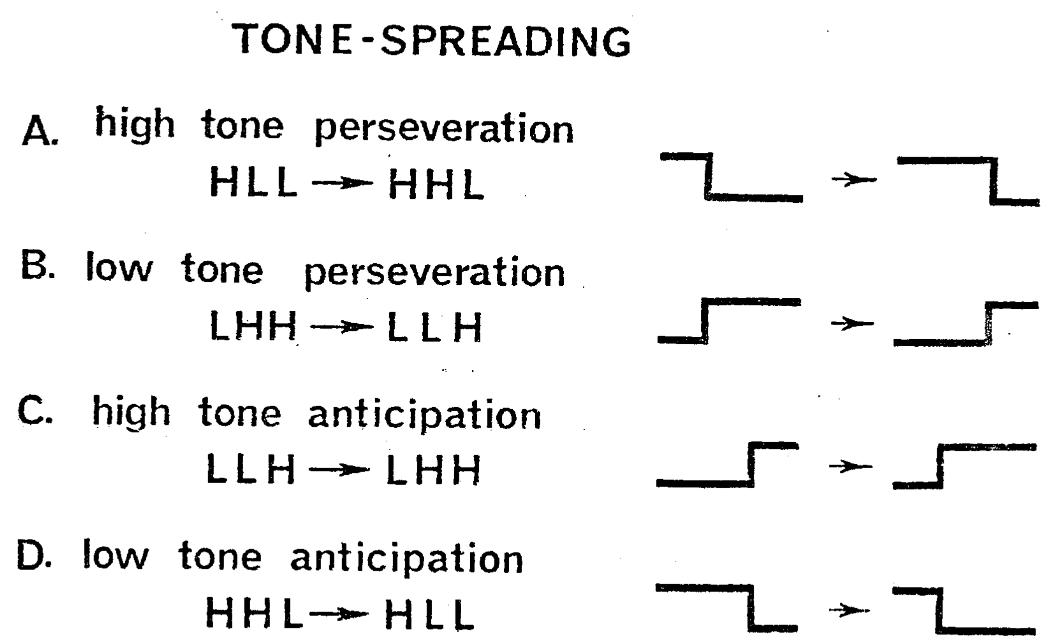


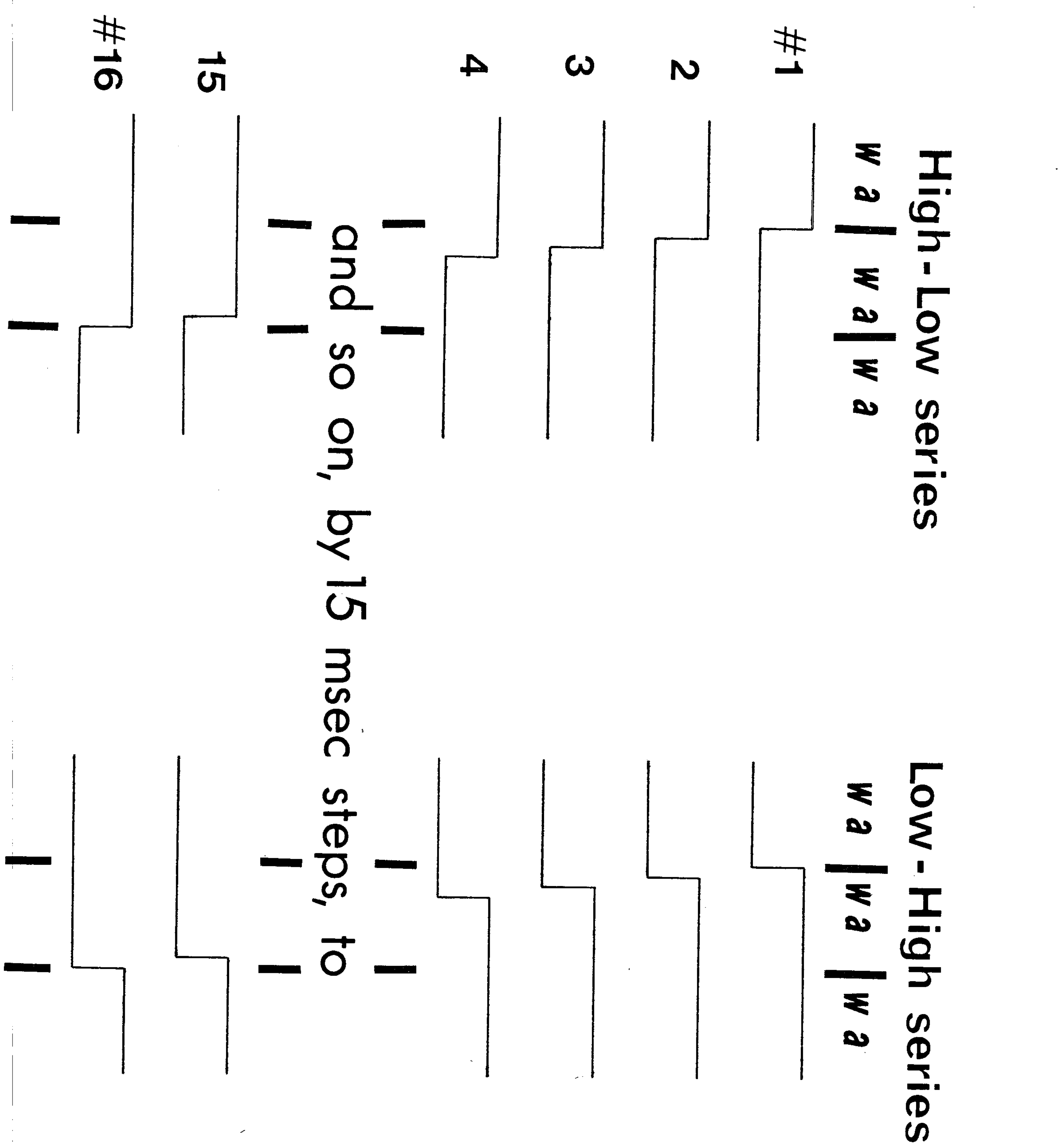
Figure 1.

The proposed universals state that there are asymmetries in the way in which rules of this type are distributed in languages of the world; specifically that tone-spreading is more often perseverative than anticipatory and that spreading of high tones is more likely to occur than spreading of low tones.

An experiment was conducted to test if an asymmetry in the perception of tone sequences can contribute to explaining the asymmetrical distribution of these rules. A synthetic speech parameter file representing a trisyllabic word /wawawa/ was created, using programs implemented on the PDP-12 computer of the UCLA Phonetics Laboratory. Syllable divisions were determined to coincide with the lowest values of Formants 1-3 in the /w/ segments. Two series of stimuli, each containing 16 items, were produced by modifying the F_0 contour of this word. The composition of these series is represented in Figure 2. In the High-Low series the entire first syllable had a higher F_0 (125 Hz) and the entire third syllable had a lower F_0 (112.5 Hz).[2] In Stimulus 1 the whole of the middle syllable was low like the first syllable; in Stimulus 16 the whole of the middle syllable was high like the first syllable. In the other stimuli the position of the change in F_0 was progressively delayed by 15 msec steps. A similarly constructed series of Low-High stimuli was also created in which the entire first syllable had the lower F_0 and the entire third syllable had the higher F_0. In each case the change in F_0 was abrupt. It may be seen from Figure 2 that the stimuli numbered 1 and 16 in each series correspond to the patterns in the graphic representation of tone-spreading rules in Figure 1. The other stimuli are, to a greater or lesser extent, ambiguous between HLL and HHL or LHH and LLH.

Two randomly ordered sequences of the 32 stimuli were generated using an Ove III speech synthesizer and recorded on audio tape. Each sequence contained five tokens of each stimulus type, making a total of 320 trials. The inter-stimulus interval was 1 second. The two sequences were presented on consecutive days to a group of subjects who were students at UCLA taking an introductory phonetics course. Presentation was over a loudspeaker in the classroom against a background of ordinary environmental noise. The subjects were asked to judge if the middle syllable of each token sounded more like the first or more like the last in pitch. Subjects indicated their response by marking the word FIRST or the word LAST on a response sheet. Every tenth token was numbered on the response sheet and the number was read on the tape to ensure that subjects did not lose their place.

During the instructions prior to doing the task on each of the two days the subjects heard the stimuli numbered 1 and 16 from each series of stimuli. They were then played a practise sequence in which each of these four stimuli occurred four times in random order. The practise sequence served to familiarize subjects with the nature of the task and with the tempo at which they were required to do it. It also provided a basis for distinguishing those who could do the task from those who found this particular task difficult or confusing to perform. Because the middle syllable is in fact the same as either the first or the last there is a 'correct' response to



High-Low series
w a | w a | w a
Low-High series
w a | w a | w a
#1
2
3
4
and so on, by 15 msec steps, to
15
#16

Figure 2.

the stimuli numbered 1 and 16. A pilot experiment had shown that failure to give this correct response to the practise stimuli predicted a performance that was close to chance on the whole task. Therefore, to eliminate those for whom the task presented irrelevant difficulties, results from only those subjects who gave no more than one aberrant response in the 32 practise trials are reported. There were 16 such subjects whose results are discussed below. No significant learning pattern emerged in comparing the responses from the two days, so the data from both days is pooled.

The subjects' responses amount to deciding which of the number 1 or 16 stimuli a more or less ambiguous stimulus is equivalent to, and sorting the F_0 patterns into the four categories HLL, HHL, LHH and LLH. If there is no perceptual asymmetry influencing the responses, the stimuli 1 through 8 should be predominantly judged to have the middle syllable like the last syllable (i.e. HLL and LHH) and the stimuli 9 through 16 to have the middle syllable like the first syllable (i.e. HHL and LLH). Figure 3 shows the responses to the stimuli in the High-Low series. For each stimulus number the percentage of HLL responses is shown by a solid bar and the percentage of HHL responses is shown by a striped bar. For this series no perceptual asymmetry is apparent. The cross-over from a majority of HLL responses to a majority of HHL responses falls between stimuli 8 and 9 - stimuli in which more than half the middle syllable is low are overwhelmingly judged to be HLL, stimuli in which more than half the middle syllable is high are overwhelmingly judged to be HHL.

The responses to the stimuli in the Low-High series are shown in Figure 4. The percentage of LHH responses for each stimulus number is shown by a solid bar and the percentage of LLH responses by a striped bar. The cross-over point from a majority of LHH responses to a majority of LLH responses occurs around stimulus 10. In this series more of the ambiguous stimuli were heard as having the middle syllable like the final high toned syllable than like the initial low toned syllable. A definite asymmetry is apparent. This bias may be interpreted as favoring high tone spreading over low tone spreading, as one of the proposed universals indicates is the case. But the asymmetry results from a preference for <u>anticipating</u> the following high tone, which is the antithesis of the other proposed universal.

The results may be tabulated and the conclusions drawn by totalling the reponses in various categories. With the High-Low series of stimuli 50.8% of total responses were equivalent to HLL (and 49.2% to HHL). This is not significantly different from 50% by a χ^2 test for goodness of fit. With the Low-High series of stimuli 59.4% of total responses were equivalent to LLH (and 40.6% to LHH). These figures are significant at better than a .01 level. In the two series together, those responses which treated the middle syllable as anticipating the final tone (i.e. HLL + LHH) totalled 55.1% and those responses which treated the middle syllable as high toned (i.e. HHL + LHH) totalled 54.3%. These figures are also significant at better than .01, but their significance results from the asymmetrical responses to the Low-High series.

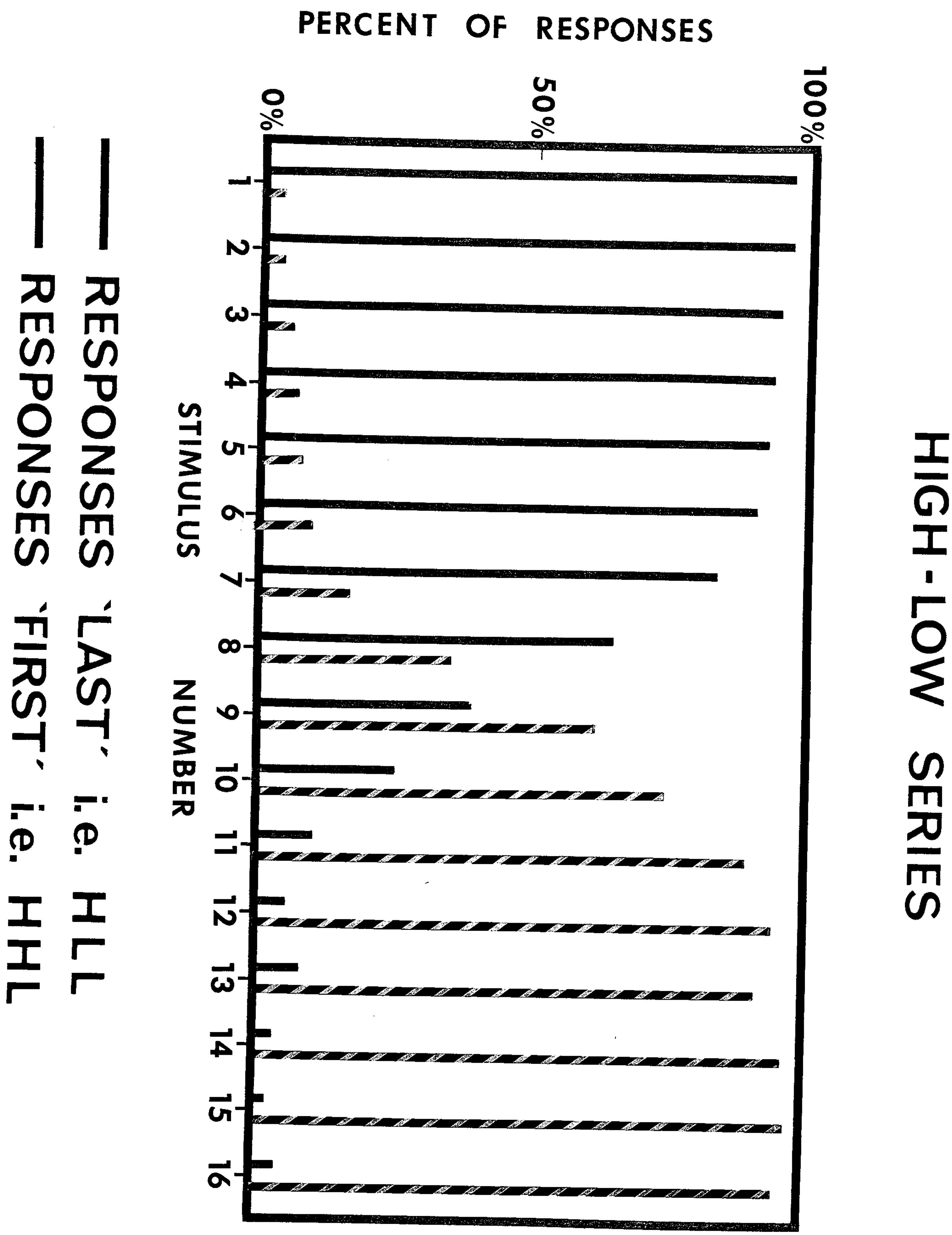

HIGH-LOW SERIES
PERCENT OF RESPONSES
0%
50%
100%
STIMULUS NUMBER
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
RESPONSES 'LAST' i.e. HLL
RESPONSES 'FIRST' i.e. HHL

Figure 3.

The unexpected finding of a preference for anticipation[3] is not sufficient to lead to a denial of the proposed generalization, but it may suggest that the data on which it is based deserve re-examination. A different choice of underlying forms can convert a perseverative rule into an anticipatory rule. For example, consider a case where HLL and HHL alternate. If HLL is the underlying pattern, then HHL is derived by a perseverative rule (Figure 1 A); but if the HHL pattern is chosen as underlying, then HLL is derived by an anticipatory rule (Figure 1 D). It is necessary to ensure that a choice between competing analyses is soundly motivated in order to either confirm or refute the generalization. If it is confirmed then a problem remains in reconciling these experimental data with the linguistic facts.[4]

Footnotes

1. In the design, execution and writing up of the experiment reported in this paper I have benefitted greatly from consultation with others, principally my colleagues in the Phonetics Laboratory at UCLA. I am especially grateful for encouragement and advice from Peter Ladefoged and Louis Goldstein who also wrote a program to tabulate the scores. I should like to record my gratitude to those who endured unpaid the rigors of being subjects both in the pilot and full-scale studies. This work was supported by the NSF.

2. These F_0 levels are within the range typically employed by a male speaker in distinguishing tones in a language with two tone levels.

3. It is possible that this finding would not be repeated if subjects were speakers of a tone language. The experiment will be re-run with such speakers to test this possibility. It has also been suggested to me that the finding results from there being perceptual syllable boundaries located differently from the acoustic boundaries defined for the experiment. If the perceptual boundaries were later than the acoustic boundaries then the acoustic center of the syllable would be in the earlier half of the perceptual syllable. A cross-over point between stimuli 8 and 9 would then indicate that there was a preference for perceiving perseveration in relation to these perceptual boundaries. A listening test conducted with stimuli prepared by removing all or most of either the first or the last syllable of /wawawa/ appeared to refute this possibility. Moreover, the original experiment was designed to discover if there was an asymmetrical response to a lack of correlation of a pitch change with a 'segmental' change, simulating an articulatory slop in production. If indeed there was a delayed perceptual boundary to syllables this would be part of the subject-matter of the investigation.

4. Cf. H. Javkin "Auditory basis of progressive tone spreading" [Abstract] Journal of the Acoustical Society of America 60 Supp 1: S45.

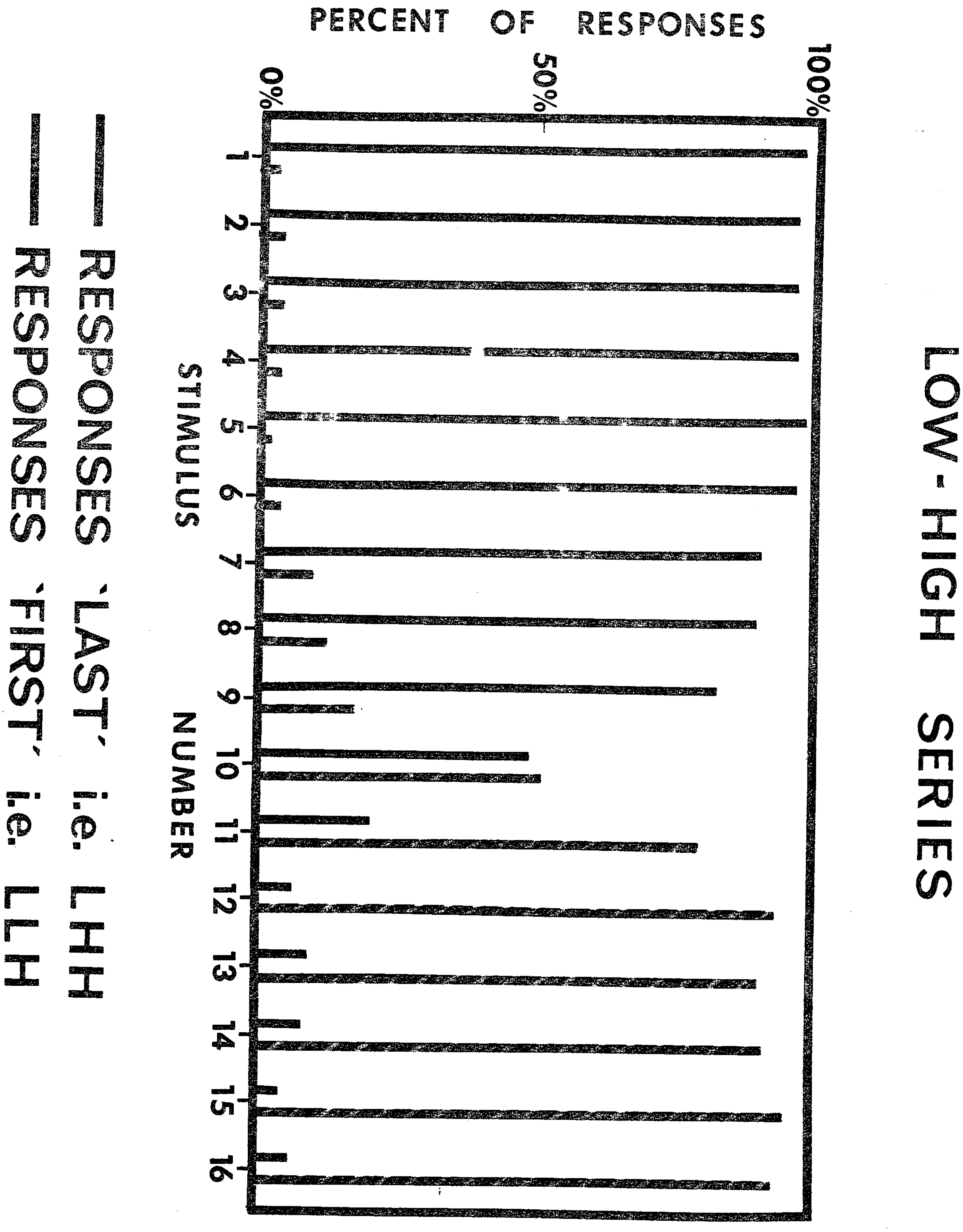

LOW - HIGH SERIES
PERCENT OF RESPONSES
0%
50%
100%
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
STIMULUS NUMBER
RESPONSES 'LAST' i.e. LHH
RESPONSES 'FIRST' i.e. LLH

Figure 4.

Tone Effects on Consonants

Ian Maddieson

Introduction

It is well-known that consonants induce pitch transitions in following vowels that differ according to the type of consonants concerned. There are also pitch differences in vowels which precede consonants. The synchronic and diachronic effects of such differences on the tone systems of tone languages and the possibility that nontonal languages develop tonal contrasts from them has been the subject of much discussion in the linguistic literature (e.g., Hyman 1973a, Gandour 1974, Hombert, Ohala and Ewan 1976).

Less attention has been paid to the fact that differences of tone may affect the adjacent consonants. In fact it has even been denied that such an effect is possible (Hyman 1973b, Hyman and Schuh 1974, Pike 1974). It may well be true that there are fewer phonologically significant effects of consonants on tones. This is not surprising, since there are more different consonants than different tones in any language and the consonants differ along many more dimensions than the tones. However there is a sufficient number of well-attested cases of tone effects on consonants that the possibility must be recognised as a fact. Additional cases, for which the evidence may not be complete, appear more convincing in view of similarities to the better attested cases.

This paper will review in three sections the available evidence showing that tones affect consonants. The first describes diachronic changes of consonants in which tone affects the results. The second describes those cases in which a synchronic consonantal difference has been impressionistically observed to depend on tone. These vary from cases where small phonetic differences are reported to those where phonemic boundaries are crossed. Comments on the diachronic sources are added where appropriate. The third describes phonetic effects of tones on consonants which have been instrumentally determined. In each of these sections similar effects are grouped together as a preliminary step towards establishing a typology. The fourth section of the paper examines the significance of the fact that tones can affect consonants for those cases in which a correlation of consonant types and tone is open to interpretation as either an effect of consonants on tones or of tones on consonants. The final section discusses further the relationship of the diachronic and synchronic effects.

1. Diachronic effects of tones on consonants.

(i) Changes in phonation type.

Beach (1938) reports that, in the Southern African Khoisan language Korana (!ora), */g/ and perhaps other previously voiced segments have become devoiced when one of the higher tones follows. In addition, Beach reconstructs a tone-splitting process in the closely related Nama language by which the two lower tones (described as mid-low and mid-falling) in the parent language developed a rising variant after an originally voiced consonant. Because this split does not affect the higher tones, Beach suggests that the originally voiced consonants must have become devoiced before the high tones prior to the tone-split. After the tone-split, consonants before the lower tones also became devoiced. The comparison with Korana provides both the evidence for the original voiced status of the consonants concerned and the evidence for the early devoicing before high tones.

In Guthrie's 'Group C' of Bantu languages, Ellington (1970-1) shows that /l/ in Ngɔmbɛ corresponds to /t/ before high tone in Lingala, Bobangi, Mongo, Ntomba, and Bolia; but /t/ before low tone is common to the group. A few of Ellington's examples are given in (1).

(1)

	in Ngɔmbɛ	elsewhere in Group C
'thousand'	kɔ́tò	nkɔ́tò
'scorpion'	kɔ̀lɔ́	nkɔ̀tɔ́
'to leak'	tàngà	tàngà
'to count'	lángà	tángà

Ellington formulates a 'sound change rule' (2):

(2) $*t \longrightarrow l \,/\, \_\_ \acute{V}$

and has over 30 examples of corresponding series to illustrate it. He considers the possibility of other causes, but concludes that:

> "since high tone is the only consistently present factor in those cases where /t/ became /l/ then it may be assumed that this is the conditioning factor."

The change (2) is shared by the closely related Bwela (or Dɔkɔ) language (Hulstaert 1961), e.g., Bwela 'to count' is /lángà/. Note that, since /l/ is the normal reflex of Proto-Bantu */d/ in Ngɔmbɛ, the change may have been of /t/ to /d/ in the first instance, i.e., a straightforward voicing change.

This development in Ngɔmbɛ may seem surprising as the association of high tone with voicelessness and low tone with voicing is more prevalent. It should be remembered, however, that where diachronic effects of consonants on tones have been detected the correlation of voicing with higher tone and voicelessness with lower tone is not unusual. It is found, for example, in Sui (Haudricourt 1972), the Central branch of the Tai family of languages and many of the Northern Tai languages (Brown 1965, Sarawit 1973), Nasu and Lü-Ch'üan among the Loloish languages (Matisoff 1972), Chaochow and Shaowu among the Min Chinese dialects (Norman 1973), etc.

The languages of China[1] and South-East Asia provide many examples of changes in phonation type of initial consonatns that depend on the tonal category of the following vowel. For example, in Puyi dialects of the Tai family, original voiced obstruents have modern voiceless aspirated or unaspirated reflexes depending on the tones (Sarawit 1973). A group of Northern Thai dialects is identified as sharing a loss of aspiration in voiceless aspirated stops before reconstructed tones I and IV, whereas voiceless aspirated stops remain unchanged before tones II and III (Jones 1965). In Mandarin and Yüeh dialects of Chinese, the old voiced obstruents became devoiced but became voiceless aspirated if preceding the "even" tone (p'ing-sheng), and voiceless unaspirated elsewhere (Forrest 1965:230, Cheng 1973). Cheng gives the table of correspondences in (3) between voiced initial consonants[2] in Middle Chinese and their reflexes in the Taishan dialect of the Yüeh group. The Middle Chinese initials are based on an interpretation of the Ch'ieh-yun rhyme tables compiled before 601 A.D.

(3) Taishan reflexes of Middle Chinese voiced stops

Middle Chinese initial consonants	b	d	dz	dʒ dz d	g
Taishan reflex before 'even' (p'ing) tone	p^h	h	t^h	$tʃ^h$	k^h
Taishan reflex with other tones	p	ø	t	tʃ	k

In some Yueh Chinese dialects words with the 'lower' (yang) rising tone (shang-sheng) also have voiceless aspirates (Hashimoto 1972:32). In the Hakka group of Chinese dialects the voiced obstruents show voiceless aspirated reflexes in all tones (Yang 1967). Mandarin and Yueh therefore seem to show a selective further development to unaspirated reflexes in some tones.

As for another major Chinese grouping, Ballard (1969) shows that the Wen-Ling dialect of the Wu group has preglottalized variations of nasals and laterals with the originally 'rising' tone. Furthermore, these appear with the higher reflex of this tone, which is normally associated with originally voiceless consonants. Nasals and laterals with one of the other three original tones occur with 'voiced aspiration' (transcribed /ɦm, ɦn, ɦl/) and appear with the lower reflex of the tones. Interestingly some varieties of Wen-Zhou have voiced aspirated nasals and laterals in the rising tone. Two other dialects, Chang-Zhou-Shi and Jin Hua, may have shared this change since they have the unexpected higher reflex of the original rising tone after nasals and laterals. Although nasals and laterals to not currently have different reflexes according to the tones in these two dialects, some explanation is required for the different way the original tones have been split.

Chang (1973, cf. Li, Ch'en and Ch'en 1972) provides extensive evidence of changes similar to the Chinese ones in the Miao languages of the Southern China border area. The series of reconstructed voiced stops have voiceless unaspirated reflexes with 'A2' and 'D2' tones, and voiced aspirated reflexes with 'B2' and 'C2' tones in the Hsien-chin dialect. In Chi-wei the voiced stops become voiceless unaspirated with 'A2' tones and are voiceless stops followed by voiced aspiration with 'B2' 'C2' and 'D2' tones. In Tsung-ti they become voiceless unaspirated with 'A2' and 'D2' tones, but voiceless aspirated with 'B2' and 'C2' tones: the voiced stops remain voiced with 'C2' tone and with the special reflex of 'B2' tone that occurs with nouns, but become voiced aspirated with 'A2' and 'D2' tones and the reflex of 'B2' tone with verbs. Some examples are given in (4) to illustrate these changes.

(4)

gloss	original initial consonant	tonal category	Hsien-chin	Chi-wei	Tsung-ti	Shih-men-k'an
'nine'	*ǰ	A2	čɒ 32	čo 21	ča 42	ǰɦa 35
'to wait'	*d-	B2	dɦau 21	tɦan 22	thon 22	dɦo 13
'to die'	*d-	C2	dɦa 13	tɦa 42	tha 13	da 31
'ten'	*g-	D2	kəu 24	kɦu 22	ku 31	gɦau 31

Chang and Li et al. also show that former prenasalized voiced stops in Tsung-ti become plain prenasalized voiceless stops with A2 and D2 tones but are followed by voiced aspiration with B2 and C2 tones.

(ii) Place or manner of articulation.

Tonal differences have also affected the distinction between fricative and affricate initial consonants in Chinese. Pulleyblank (1970, 1974) shows in detail that the distinction between Middle Chinese voiced palatal and retroflex fricatives and affricates was lost by the end of the Middle Chinese period. A retroflex fricative, perhaps with a breathy offglide, was the merged reflex. Later in Mandarin dialects a split occurred conditioned by the tones. Pekinese has a retroflex voiceless aspirated affricate with the lower even tone (yang p'ing) and a retroflex voiceless fricative with other tones (and also before the vowel /ɨ/ in even tone). The process is illustrated by the diagram in (5).

(5) Early Middle Chinese — Late Middle Chinese — Mandarin

dz, dʒ → $s^ɦ$ → ts^h / even tone

z, ʒ → $s^ɦ$ → s / elsewhere or ____ɨ on any tone.

Examples of modern reflexes are given in (6).

(6)

Early Middle Chinese segment	Mandarin reflex (transliterated): yang p'ing tone	ch'u tone
palatal fricative	ch'eng 'to fill'	sheng 'full'
palatal affricate	ch'eng 'to ride'	sheng 'vehicle'

The words across the rows in (6) rhymed in all but tone in Middle

Chinese, now they differ also in the initial consonant. The words down the columns are now homophonous, having lost a contrast in the initial consonants.

Some tone-dependent changes of final consonants are noted by Jones (1961) in his phonological history of Karen dialects. For example, in the parent of Sgaw and Palaychi dialects, Proto-Karen final */-dh/ became /-h/ after high tone but /-d/ after low tone.

In Kpelle, a Southwestern Mande language, originally voiceless obstruents have three different reflexes (Dwyer 1973, 1974). These may be illustrated, using /p/ as a representative, as in (7).

(7) (i) *p ⟶ p

(ii) *ń-p ⟶ ḿb

(iii) *ǹ-p ⟶ bb

As may be seen, prefixed high and low toned nasals have different effects. The low-toned nasal prefix that results in the voiced geminated obstruents in (7 iii) can be reconstructed from an examination of synchronic alternations (cf. Hyman 1973c). These involve changes in initial consonants when nouns become definite or verbs take a third person pronominal object. Words beginning with a nasal still add a low-toned syllabic nasal prefix in these cases, but those with initial voiceless obstruents replace them with homorganic voiced geminates. Examples are given in (8) (from Welmers 1962, with modified transcription).

(8)

<u>nouns</u>	<u>indefinite</u>	<u>definite</u>
'thing'	sɛ̄ŋ	zzɛ̄ŋ
'vehicle'	kélēŋ	ggélēŋ
'rice'	mɔ̄lɔ̄ŋ	m̀mɔ̄lɔ̄ŋ
<u>verbs</u>	<u>without object</u>	<u>with object</u>
'jump, throw'	pɪ́lɪ̄	bbɪ́lɪ̄
'swell, blow up'	fɛ̄ɛ̄	vvɛ̄ɛ̄
'boil'	nɛ́ŋ	ǹnɛ́ŋ

Dwyer denies that the low tone on the nasal was the cause of the gemination but suggests rather:

> "stress is generally assigned to the first high tone in the nominal phrase. Consequently phrase-initial low-toned nasals are never stressed while phrase-initial high-toned nasals are always stressed. As a result, Gemination applies to low-toned nasals and not to high-toned nasals because of the location of the stress. The reduction of marked features in an unstressed segment is a very common event." (p.69)

Dwyer's explanation is an interesting attempt to find a natural cause of the Kpelle distribution. Other writers, e.g., Welmers (1962) and

Bird (1971) have suggested that the low tone is causally involved in the origin of the geminated voiced ("heavy voiced") segment.

Dwyer objects to the argument that low tone causes the consonant change on the grounds that other low tones do not cause gemination, only low-toned syllabic nasals. He cites such examples as (9) in Central Kpelle. By a tone-lowering and voicing rule (9i) becomes (9ii), not (9iii).

(9) i) kálòŋ 'chief' + pɔ́lɔ́ 'old'

ii) kálòmbɔ̀lɔ̀ 'old chief'

not iii) *kálòmbbɔ̀lɔ̀

However, the non-occurrence of gemination in such an environment is not explained by an appeal to stress, since the second syllable of /kálòŋ/ is an unstressed syllable. Stress is a secondary feature of Kpelle predicted from the tones, and unstressed syllables are not altered by deletion, contraction or neutralization processes. On the other hand, tone is a primary feature in the Southwestern Mande languages. Thus the low tone surely must have played a role in producing the result shown in (7).

(iii) Introduction of consonantal segments due to tones.

In many Scandinavian dialects, two tonal accents distinguish words which are otherwise identical. Originally monosyllabic words have Accent I, originally bisyllabic or polysyllabic words have Accent II. Due to compounding and epenthesis, some Accent I words now have more than one syllable, and due to contraction some Accent II words are now monosyllabic. There are more minimal contrasts in Norwegian dialects than in Swedish. Oslo Norwegian, for example, <u>bunden</u> 'the floor' has Accent I (a rising pitch pattern) but <u>bunden</u> 'bound (participle)' has Accent II (a falling rising pitch pattern) (Kock 1901). In Standard Danish however, the accent distinction is not realised as a tonal difference. Instead, the words with Accent I have a glottal stop (stød) at the end of the vowel in the accented syllable. Larsen (1890) also reports that there is little accentual difference in some of the southernmost Norwegian dialects and a glottal stop characterizes Accent I. Most scholars have been persuaded that the Danish situation represents an innovation caused by tonal differences between accented syllables in the parent language of all the continental Scandinavian dialects. The accented syllables can be distinguished from unaccented syllables, and only stressed syllables with Accent I develop the glottal stop. Kock, following Sievers, suggests that the Danish <u>stød</u>

> "must have arisen in the following way; in the course of the tensing of the vocal cords required to produce the rapidly rising pitch contour of the stem syllable of words with Accent I, the tension was over-articulated so that the vocal cords momentarily closed"

It is not really necessary to maintain that the vocal cords actually closed in producing the contour as a simple cessation of voicing in this position could readily be re-interpreted as a laryngeal segment. A very similar effect seems to have occurred in Latvian (Ekblom 1933), a language in the quasi-tonal Baltic group (cf. Steinbergs 1975).

Another case of high rising tone introducing a glottal stop where none existed before can be detected from comparing dialects of the Mon-Khmer language Jeh (Gradin 1966). In a Southern dialect, there is a 'high tone' which consists of a word-final sharp rise after a level pitch of normal short vowel duration. This sharp rise corresponds to a segment that Gradin calls a 'glottal fricative' in some Northern dialects of Jeh, and to /h/ and /s/ in related adjacent Mon-Khmer languages. If this segment /h/ is a devoiced or aspirated final vowel, the origin of the pitch rise may be connected with the increase in airflow through the glottis caused by widening the aperture between the vocal cords. Gradin notes that the sharp rise in pitch in the Southern dialect can occur with a vowel broken up by a glottal stop. As this involves closing the vocal cords it is not plausible to suggest a direct change h → ʔ . Rather, an originally automatic pitch rise may have been interpreted as phonological, and subsequently the over-articulation of an energetic rise resulted in vocal cord tension reaching the point where vibration was halted.

It is also probable that the glottal stop in the high rising 'broken' tone (ngã) in Northern varieties of Vietnamese originates from a similar process. Mohring (1972) shows that many words in Vietnamese with ngã tone have cognates in Khmuʔ with final /-h/. Again, it seems unlikely that /ʔ/ directly corresponds with /h/. Haudricourt (1954) proposed that both high rising tones in Vietnamese arose from final /-h/ (or /-s/). The second one (sắc) is not accompanied by the intrusive glottalization found in the ngã tone. Mohring is unable to find evidence for final /-h/ in the cognates of words with sắc tones. The absence of cognates with final /-h/ and the absence of the glottal stop in sắc tones emphasize the similarity with the development in Jeh.

One other probable case of the introduction of a glottal stop occurs in an American language. In Kiowa (Sivertsen 1956) the occurrence of a glottal stop syllable-finally before pause can be predicted from the occurrence of low tone in long syllables. This glottal stop varies in other positions with medial constriction of the vowel, or laryngeal (creak) of the vowel, or a plain voiced but 'rearticulated' vowel. Low tone occurs without this glottal feature in short syllables, and long syllables with high tone do not have it. It may therefore be an innovation, rather than a residual phenomenon, introduced by the rule given in (10).

(10) $\emptyset \rightarrow ʔ$ /VV__(C)#

The phonetic process probably involved first, the development of laryngealization simply because vocal cord vibration slows to a very low frequency at the end of a long low-pitched vowel. Subsequently, the complete glottal closure developed in final position.

2. Synchronic rules showing effects of tone on consonants.

In this section phonological rules or allophone distributions which are based on impressionistic description will be covered.

(i) Phonation type.

Two cases in which low tone predisposes voiceless stops to become voiced are reported from Tibeto-Burman languages spoken in North-Eastern India. In Tankhur Naga (Bhat 1969) "the three unaspirated voiceless stops p, t and k become voiced intervocalically especially before a vowel in low tone". In Central Monpa (Das Gupta 1968) the rule is a more strict selection between allophones determined by tone. The stops /p, t, c, k/, which are voiceless unaspirated before high tone, are described as having "softened low tone varieties" with "indistinct voicing" before low tone. Voiceless aspirated stops occur before both tones in these languages without any observed allophonic variation.

In Jingpho (Maran 1971), a more distantly related language spoken in Burma, voiceless final stops after low tone vowels become voiced stops in a gemination process which occurs before certain suffixes (such as an affirmative, possessive or imperative marker). This alternation of voicing does not occur after high tone vowels, as (11) shows.

(11) yàk 'difficult' yàggai 'it is difficult'
cát 'tight' cáttai 'it is tight'

It can be shown that the voicing is an innovation in Jingpho. An archaic ritualized form of the language apparently does not exhibit voicing of final consonants (Maran 1971:194), and Proto-Lolo-Burmese cannot be shown to have had any voiced final stops (Burling 1967, Matisoff 1972).

Among the Kwa languages of West Africa, Dunstan and Igwe (1966) report that the "glottal fricatives" (symbolized /h, ɦ, hj, ɦj/) "appear to be partially voiced before low tones". This remark probably means that a murmured onset to a low tone vowel occurs.

(ii) Place and manner of articulation.

In Tucano (West and Welch 1967), spoken in an area on the border between Colombia and Brazil, there is an alternation between plain and prenasalized stops after nasalized vowels such that if the vowel has a high tone the stop is prenasalized. An additional condition is that the nasalized vowel must be preceded by a voiceless consonant. Examples are given in (12).

(12) <u>with prenasalization</u>

/kṍpè/	[kṍᵐpɛ̀ʰ]	'left'
/wèhẽ́kì/	[wɛ̀hẽ́ᵑkìʰ]	'fish net'
/wã̀kṹtì/	[wã̀ʰkṹⁿtìʰ]	'remember?'

<u>without prenasalization</u>

/sũ̀kṹã̀/,	[sũ̀kṹã̀ʰ]	'small of the back'[3]
/dĩ̀ʔkã̀kã́/	[nĩ̀ʔkã̀kã́ʰ]	'today'

In Picuris, a Tanoan language of New Mexico, Trager (1971) observed an effect of low tone on the first member of a consonant cluster with a glottal stop as the second member. The effect is illustrated with the example in (13).

(13) /êmʔene/ → [ʌ̂ᵐbʔɛnɛ] (^ represents low tone and primary stress)

In the example Trager describes the /m/ as becoming "partially denasalized and devoiced, and more tense". She attributes the effect (of which no further examples are given) principally to anticipation of some of the qualities of the glottal stop in the /mʔ/ cluster. However as similar clusters preceded by high or mid tone do not show the effect, there is a differential effect of tones on these clusters.

An interesting case where tone seems to affect place of articulation in Atsi, a Burmese language, is reported by Burling (1967). Discussing variations in the realisation of /v/, which in most cases is labio-dental, he notes that before /e/ and /a/, in high-toned syllables only, it has bilabial friction. His description is not entirely clear but it does seem certain that he observed some tone-dependent variation in the articulation of this segment.

(iii) Duration

Malmqvist (1962) reports that in Chengtu, a Western Mandarin dialect of Chinese, final nasals have phonetic variants conditioned by tones. In words ending /-en/, the nasal is long (transcribed [n:]) after the rising tones 1 and 4 (which are 45 and 213 on the Chao five-digit scale). After the falling tones 2 and 3 (21 and 53 respectively) the nasal is short. In words ending /-an/, a fully articulated nasal is only heard after tone 4; with other tones the nasal is described as "weakly articulated" and its presence is principally indicated by nasalization of the preceding vowel.

3. Instrumentally-verified phonetic effects of tones on consonants.

Surprisingly few instrumental investigations of tone languages have been carried out, and only in a small proportion of these have properties of consonants been measured to determine if there are differences dependent on tone. This is in contrast to a considerably larger number of instrumental investigations in which pitch differences caused by adjacent consonants have been studied. These latter studies have been greatly stimulated by the interest in seeking the phonetic basis of tone-splits and tonogenesis. As the possibility of tone effects on consonants has been largely neglected there has been no equivalent stimulus to seek a phonetic understanding of the process involved. Yet there is no *a priori* reason to rule out various forms of co-articulation between tones and adjacent segments. Non-pitch effects of tones on consonants are discussed below.

(i) Phonation type.

Gandour (1974) has shown that the aspirated stops in Standard Thai have two allophones; one is voiceless aspirated, the other has 'voiced aspiration' following a voiceless closure phase. The frequency with which the 'voiced aspirated' variant appears is determined by the tone which follows. Approximate figures are given in (14).

(14) Approximate percentages of breathy allophones of aspirated consonants (Thai)

High Tone	Falling Tone	Mid Tone	Low Tone	Rising Tone
100%	90%	50%	28%	20%

The breathy allophone is normal with high and falling tones but rare with low and rising tones. Gandour suggests that this may be due to retention of breathiness from one series of consonants from which the aspirated stops are historically derived. But, as the allophones occur with aspirated stops from several sources apart from the Proto-Thai voiced (or 'breathy-voiced') stops, it is not clear that this is correct. Nevertheless whether this is differential retention of a feature or an innovation of a tone-determined allophone, the distribution of allophones is the result of the tonal differences.

It may also be noted in this section that in Pekinese Karlgren (1960) experimentally observed a devoicing of final nasals after falling tones. The instrumental evidence is not published or described in detail, but he reports that in kǎn 'dare' (rising tone) the /-n/ "was intense and richly voiced to the very end: whereas in k'ân 'look' (falling tone) the /-n/ "had a weak intensity...and lost its voiced nature at the end (being a voiceless n in its final moments)".

(ii) Larynx movement.

The data collected for a study of larynx height in Thai provided the opportunity to measure the differences in larynx position of the three consonants /p, p^h, and b/ which result from differences of tone. The highest point reached by the larynx during the closure of each of the stops is not significantly affected by the following tone (although the height reached differs for the voiceless and voiced consonants). However, the movement from the highest point down to the low position reached at or near the beginning of the vowel is strongly affected by the tone, as may be seen from the table in (15).

(15) Amount of Larynx Movement in Consonants

Consonant	Rising Tone	Low Tone	Falling Tone	High Tone	Mid Tone
p	9.3	7.9	4.2	4.8	4.1
p^h	6.8	5.1	5.4	4.6	3.7
b	8.9	6.2	5.1	3.7	3.4
Average	8.3	6.4	4.9	4.4	3.7

The units in the table are millimeters of displacement of the larynx tracing on the oscillomink records used in the experiment described in Gandour and Maddieson (1976) (these magnify by a small but unknown factor the amount of the actual movement of the thyroid process). The downward movement of the larynx is largest when a rising tone follows the consonant. The low tone causes the next largest movement and the mid tone the least. This shows that the following tone has an important physiological phonetic effect on these preceding consonants.

(iii) Consonant duration.

It seems in general true that higher tones predict longer consonant durations. Zee (1977) has shown that in Taiwanese long syllables with the segmental shape /si:/, the duration of the /s/ segment is positively correlated with the pitch of the beginning of the vowel. There are five possible tones in this environment. The fricative is longer before the high and falling tones and shorter before the low and rising tones than before the mid tone. In Standard Thai, which has a broadly similar pattern of tones on long syllables to Taiwanese, the high and falling tones predict longer consonant durations and the low, mid and rising tones shorter durations. Sample durations of /b/ are given in (16) below. These are measured from oscillograms prepared in connection with writing the paper on larynx height in Thai (Gandour and Maddieson 1976).

(16) Duration in Msec. of Consonant

Segmental Composition	High Tone	Falling Tone	Low Tone	Rising Tone	Mid Tone
bi:	120	113	106	104	101
bu:	107	106	88	88	83

Elert (1965) shows that in the quasi-tonal Swedish of the Stockholm area consonants *following* the stressed vowel are longer in words with Accent II than those with Accent I. In Accent II the stressed syllable has a level high or rising pitch whereas in Accent I the stressed syllable has a falling pitch (i.e., is low at the end).

The durations of the measured consonants in Swedish show an inverse correlation with the duration of the preceding vowel, and the same may well be largely true for Thai (cf. Abramson 1962), but in Taiwanese the vowel durations cannot be used to predict the duration of the consonant. Vowels on rising tones are long, much longer than vowels on low tones, yet consonants before both these tones are short in duration. Therefore the only generalization that will cover these three related cases is that in (17):

(17) Consonants adjacent to hgih-pitched vowels are longer than those adjacent to lower-pitched vowels.

4. Implications of tone effects on consonants for interpretation of ambiguous cases.

As the previous three sections have shown, some correlations between tones and properties of adjacent consonants can arise because of an effect of the tones on the consonants. We also know that certain correlations between tones and consonants can arise because of effects of consonants on tones. Some tone effects and consonant effects are mirror images of each other. For example, the association between voicing and lower tones can be seen to be a mutual relationship. On the other hand, the consonant-lengthening effect of high tones is not known to be matched by a pitch-raising effect of long consonants (e.g., phonologically long or geminate consonants or consonant types with greater inherent length).

Where the effects are mirror images, a correlation of tone and consonant may be open to two different interpretations. Consider for example a set of comparative data from two related languages showing that where one language has words contrasting high and low tones, the other has words contrasting voiced and voiceless consonants. A correspondence of this kind is illustrated in (18).

(18) Language A Language B

tán	tan
tàn	dan

In this situation, most comparative linguists would probably prefer to believe B is the conservative language while A has innovated a tone system. These correspondances would equally well support the reverse interpretation[5], that language B has replaced a tonal distinction by a contrast among initial consonants lacking in A. Without further evidence, this correspondence alone is ambiguous.

Let us look at an example in which the data are decidedly ambiguous. Luce (1969) reported on two dialects of Riang, a language in the Mon-Khmer group. These dialects are distinguished as Black Riang (BR) and White-Striped Riang (WSR). Luce says of them:

> "They appear to have only two tones, level and falling, according as the original initial of the main syllable was surd or sonant."

The 245 items in Luce's list do not include any minimal pairs demonstrating the tone contrast but there are pairs such as those in (19).[6]

(19)

BR	WSR	
kók	kók	'nape of neck'
kâk	kâk	'to bite'
nám	ná:m	'blood'
nûm	nɯm, nûm	'urine'

Luce does not present detailed arguments for his suggestion that the falling tone depends for its origin on an earlier voiced consonant and the high level tone on an earlier voiceless consonant. Note that in order to derive the falling tone in just those syllables initiated by a voiced consonant in Proto-Riang we have to assume that the effect of the initial consonant was not on the beginning but on the *end* of the vowel. This seems bizarre. There is no support for historical scenarios of this kind in the experimental studies of pitch that are often cited to support reconstructions of consonant influence on tone. It is also necessary to make some awkward assumptions about voicing contrasts and their effects in nasals and approximants.

Note also that two changes are assumed to have occurred. The development of tones; and the loss of voicing in initial voiced consonants. This latter change has been held accountable for the rise to phonological significance of 'register' distinctions as well as tone in Mon-Khmer languages. Proto Mon-Khmer is generally assumed to have lacked tones. Shorto (1967) says

> "It can be shown by comparison with languages which preserve the older distinction that register in Khmer and Mon, and tone in Riang-Lang (Black Riang), result from the loss of earlier voiced-voiceless distinctions in various prevocalic elements in the word-structure; chest register and Riang-Lang low tone[7] reflecting earlier voicing."

If the loss of voicing is indeed real, it may be more explanatory to reverse the assumption that tonal contrasts are innovated in Riang and posit both original tones, and original voiced and voiceless stop consonant series. The loss of the voicing contrast can then be motivated by the fact that a falling contour provides a context favouring the interpretation of a preceding stop as voiceless. Consequently voiceless stops might come to outnumber voiced ones and this predominance could complete the progress towards merging the stop series. We could thus provide motivation for the lack of voicing contrasts in stops and explain why tones contrast following nasals and approximants.

However, the evidence for any original voiced stops in Proto-Mon-Khmer does not seem very satisfactory. The major piece of evidence appears to be the choice of symbols in the Devanagari script when it was adapted to write Old Mon and Old Khmer around the 7th and 8th centuries A.D. (Pinnow 1957). Symbols which in Indian usage represented voiced segments were used to write prevocalic consonants in syllables which today have the second ('chest' or 'low') register. Today, along with pitch and vowel quality exponents of the register distinctions there is a difference of consonant voicing. Shorto (1967) says chest register in Mon "is accompanied by complete or partial voicing...of prevocalic plosives and fricatives in most contexts". The orthographic conventions adopted, and still used, for representing these chest register syllables may have been a simple and sensible attempt to use available letters where these could indicate a subsidiary distinction that implied the major one. And the primary distinction could have been one of tone or register for which no orthographic precedent existed in the Indian tradition. The choice of these letters is far from proof of the existence of a primary voicing distinction in the earlier language.

Thus it may even be the case that, in Mon-Khmer languages with a voicing contrast in initial consonants, voicing originates from the loss of tone. The register distinctions in languages such as Mon and Khmer would then represent an intermediate stage in this process.

5. The relationship of phonetic and phological effects.

The importance of the evidence that tones affect consonants reviewed in sections 1-3 would be greatly enhanced if it could be seen that the phonetic (or sub-phonemic) evidence helps to explain the phonologized synchronic and diachronic effects.

There are three phonetic effects that are instrumentally demonstrated. These are summarized in (20).

(20) (a) higher tones condition longer consonants (Taiwanese, Thai, Swedish).

(b) higher tones favor voiced aspiration after aspirated stops (Thai).

(c) lower tones require a greater downward movement of the larynx in the preceding consonant (Thai).

The lengthening of consonants has also been impressionistically observed in Chengtu where the nasals (the only permitted final consonants) are longer following high or rising tones but shorter following falling tones. As reported in section 3(i) above, Karlgren observed a devoicing of final nasals after falling tone in Pekinese. Both Chengtu and Pekinese are

Mandarin Chinese dialects, and the Pekinese situation may be a development from a difference in length rather than a direct effect on voicing.

It is not clear that any of the diachronic effects reported can be attributed directly to lengthening, but (20a) is often consistent with the results in Asian languages where different reflexes emerge from voiced or aspirated proto-consonants. For example, in Mandarin Chinese the voiceless aspirated reflexes of the original voiced obstruents emerge before what was probably originally a high level tone (p'ing-sheng, "even" tone). Elsewhere these segments become voiceless unaspirated. As aspirated consonants are longer than unaspirated consonants, this result conforms to a prediction that might be based on (20a).

The 'voicing' of aspirated stops before high tones in Thai (20b) suggests more strongly an explanation for such cases as the Mandarin one. A possible scenario is given in (21).

(21)

	(a)		(b)
before high tone	$p^{ɦ}$	$\longrightarrow$	p^{h}
elsewhere	p^{h}	$\longrightarrow$	p

The situation in (21a) represents the same allophone distribution as found in Thai, and (21b) represents a plausible development from it that corresponds to the Mandarin situation. Similarly, a phonetic effect such as the one observed in Thai could be responsible for the distribution of reflexes of originally voiced stops in Chi-wei Miao. Voiceless stops with 'voiced aspiration' occur with a high falling or mid level tone and plain voiceless stops with a low falling tone. However, a reliable reconstruction of the earlier shapes of these tones is required before this can be proposed with confidence.

A second group of effects involves the introduction or preservation of voicing in consonants adjacent to low tones. In Tankhur Naga this is a low-level effect which merely increases the probability of intervocalic voicing occurring. In Monpa a significant selection between allophones occurs, while in Jingpho phoneme boundaries are crossed (voiced stops occur initially). In Nama and Korana significant historical changes result. Again, the differences in consonant duration before high and low tones are consistent with the correlation of high pitch with voicelessness and low pitch with voicing, for voiced consonants are in general shorter than their voiceless chounterparts. Furthermore, an earlier onset of voicing may lead to perception of a segment as phonologically voiced. However, the productive difference (20c) observed between consonants before high and low tones may be more important as the cause of these correlations.

It has been suggested that in order to maintain voicing during a stop it is convenient (Hudgins and Stetson 1935) or even necessary (Hombert et al. 1976) to depress the larynx. This gesture enlarges the cavity behind the articulatory closure and enables the flow of air through the glottis to continue. The Thai data show that before low tone or rising tone a sharp downward movement of the larynx occurs during the consonant. Much smaller movements are found with other tones. These differences in the amount of larynx movement are of similar magnitude to the differences in larynx height between voiced and voiceless consonants which have been proposed as the explanation of the pitch differences on the following vowels (Hombert et al. 1976). The difference between the larynx height for /p/ and /b/ is 3.9 units at the highest point reached during the consonant closure, which is the measure that is least affected by the following tones. The difference in larynx movement between consonants before rising tone and falling tone is also 3.9 units and between rising and mid tone is 4.6 units. If the vocal cords are at all approximated, differences of this magnitude could well lead to a perceptually distinct amount of vocal cord vibration in those cases with the greater amount of larynx movement.

Conclusion.

This paper has reviewed the evidence for tones affecting adjacent consonants and shown that such effects are real and that some recurrent patterns emerge. Despite the existence of some idiosyncratic cases, the majority of the effects recorded seem open to plausible phonetic explanation and some suggestions have been offered toward development of such explanations. A rationale has been proposed for the possibly greater frequency of consonant effects on tones.

It is very much hoped that other linguists will be encouraged to consider the possibility that cases of tones affecting consonants may lie hidden by neglect in the data they are handling, and that phoneticians will be stirred to examine and measure aspects of speech that had been thought to conceal nothing of interest.

Footnotes.

1. Transliteration of Chinese words may be inconsistent as the practice of the sources quoted has generally been retained. This is intended to make it easier to refer to the sources for those who chould not make the transliteration from one transliteration to another.

2. Authorities differ over whether the voiced obstruents in Middle (or 'Ancient') Chinese were plain or breathy voiced ('voiced aspirated'). The term used in the old rhyme tables translates as "muddy" and its meaning is open to dispute. Adherents of both views agree, however, that there is only one kind of voiced obstruent as the rhyme-tables never distinguish separate series.

3. Two different tone patterns are given for this word by West and Welch in their article. The pattern quoted here is more probably the correct one.

4. The test items involved in this experiment contained all combinations of three consonants (/p, p^h, b/), three long vowels (/i:, u:, a:/), and the five tones. Consonant durations could only be reliably measured for the voiced consonants /b/ and the number of syllables of the form /ba:/ was too small on some tones to give trustworthy comparisons. Hence only data for /bi:/ and /bu:/ are presented.

5. Apart from the evidence from languages such as Monpa and Jingpho, there is perceptual evidence (Chistovitch 1969; Haggard, Ambler and Callow, 1970) which would provide a starting point for a model of change in which pitch differences could cause initial differences in phonation type to be perceived when none had previously existed.

6. Luce's transcription has been modified to represent high level tone by ´ and falling tone by ^ . Diacritical marks above vowel letters have otherwise been eliminated.

7. Falling tone, according to Luce's first-hand report.

References

Abramson, Arthur S. (1962) *Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments* Indiana University, Bloomington.

Ballard, W. (1969) *Phonological History of Wu* Ph.D. Dissertation. University of California, Berkeley.

Beach, Douglas M. (1938) *The Phonetics of the Hottentot Language* Heffer, Cambridge.

Bhat, D.N.S. (1969) *Tankhur Naga Vocabulary* Deccan College, Poona.

Bird, Charles (1971) "Observations on initial consonant change in Southwestern Mande" in *Papers in African Linguistics* (ed. C.-W. Kim and H. Stahlke) pp. 153-74. Linguistic Research Inc, Edmonton, Alberta and Champaign, Illinois.

Brown, J.M. (1965) *From Ancient Thai to Modern Dialects* Social Science Association Press, Bangkok.

Burling, Robbins (1967) *Proto-Lolo-Burmese* IJAL 33.2 Part II. Issued simultaneously as Publication 43, Indiana University Publications in Anthropology and Linguistics, Mouton, The Hague.

Chang, Kun (1973) "The reconstruction of Proto-Miao-Yao tones" *Bulletin of the Institute of History and Philology, Academia Sinica* 44:541-628.

Cheng, T.M. (1973) "The phonology of Taishan" *Journal of Chinese Linguistics* 1:256-322.

Chistovitch, L.A. (1969) "Variations of the fundamental voice pitch as a discriminatory cue for consonants" *Soviet Physics - Acoustics* 14:372-378.

Das Gupta, K. (1968) *An Introduction to Central Monpa* North-East Frontier Agency, Shillong.

Dunstan, E. and G.E. Igwe (1966) "Two views of the phonology of the Ọ̀hụ̀hụ dialect of Igbo" *Journal of West African Languages* 3:71-75.

Dwyer, J.D. (1973) *The comparative tonology of Southwestern Mande nominals* Ph.D. thesis, Michigan State University.

Dwyer, David (1974) "The historical development of Southwestern Mande consonants" *Studies in African Linguistics* 5:59-94.

Ekblom, R. (1933) *Die Lettischen Akzentarten* (Arbeten utgivna med understod av Vilhelm Ekman's Universitetsfond) Uppsala University.

Elert, C-C (1965) *Phonologic Studies of Quantity in Swedish* Skriptor, Stockholm.

Ellington, John (1970-1) "Tone and segmental sound change in a Bantu language" *African Language Review* 9:148-158.

Forrest, R.A.D. (1965) *The Chinese Language* (2nd edition) Faber and Faber, London.

Gandour, Jack (1974) "Consonant types and tone in Siamese" *Journal of Phonetics* 2:337-350.

Gandour, Jack and Ian Maddieson (1976) "Measuring larynx height in Standard Thai using the cricothyrometer" *Phonetica* 33:241-267.

Gradin, Dwight (1966) "Consonantal tone in Jeh phonemics" *Mon-Khmer Studies* II (Publication No. 3, Linguistic Circle of Saigon).

Haggard, M., S. Ambler and M. Callow (1970) "Pitch as a voicing cue" *Journal of the Acoustical Society of America* 25:105-113.

Hashimoto, Oi-kan Yue (1972) *Studies in Yue dialects 1: Phonology of Cantonese* (Princeton-Cambridge Studies in Chinese Linguistics 3) Cambridge University Press.

Haudricourt, André-Georges (1954) "De l'origine des tons en vietnamien" *Journal Asiatique* 24:69-82.

Haudricourt, A.G. (1972) "Two-way and three-way splitting of tonal systems in some far Eastern languages" (translated by C. Court) in *Tai Phonetics and Phonology*, Central Institute of English Language, Bangkok, 58-86. Originially appeared in *Bulletin de la Société de Linguistique de Paris* 56 : 163-180 (1961).

Hombert, Jean-Marie, J. Ohala and W.G. Ewan (1976) "Tonogenesis: theories and queries" *Report of the Phonology Laboratory* (Berkeley) 1:48-77.

Hudgins C.V. and R.H. Stetson (1935) "Voicing of consonants by depression of the larynx" *Archives néerlandaises de phonetique expérimentale* 11:1-28.

Hulstaert, G. (1961) "Sur le parler Dɔkɔ" *Aequatoria* 24:121-135.

Hyman, Larry M. (1973a) "The role of consonant types in natural tonal assimilations" *Consonant Types and Tone* 153-179 in Hyman 1973b.

Hyman, Larry M. (1973b) *Consonant Types and Tone (Southern California Occasional Papers in Linguistics 1)* University of Southern California, Los Angeles.

Hyman, Larry M. (1973c) "Notes on the history of Southwestern Mande" *Studies in African Linguistics* 4:183-196.

Hyman, Larry M. and Russell Schuh (1974) "Universals of tone rules: evidence from West Africa" *Linguistic Inquiry* 5:81-115.

Jones, Robert B. (1961) *Karen Linguistic Studies: Description, Comparison and Texts* (University of California Publications in Linguistics 25) University of California Press, Berkeley and Los Angeles.

Jones, Robert B. (1965) "On the reconstruction of Proto-Thai" *Lingua* 14:194-229.

Karlgren, Bernhard (1960) "Tones in Archaic Chinese" *Bulletin of the Museum of Far Eastern Antiquities* (Stockholm) 32:113-142.

Kock, Axel (1901) *Die Alt. und Neuschwedische Accentuierung unter Berücksichtigung der anderen nordischen Sprachen* (Quellen und Forschungen zur Sprach- und Culturgeschichte der Germanischen Völker 87) Trübner, Strassburg.

Larsen, Amund B. (1890) [Indberetning fra adjunkt Amund B. Larsen om reiser foretagne med offentligt stipendium i årene 1888-90 for at studere bygdemålene i Kristiansands stift] *Universitets- og skole-annaler* Ny række 5:267-289 (Continued in 6:209-260. 1891).

Li Yung-sui, Ch'en K'o-chung and Ch'en Ch'i-kung (1972) "Some problems concerning initials and tones in the Miao language" in *Miao and Yao Linguistic Studies* translated by Chang Yü-hung and Chu Kwo-ray (Data Paper No. 88 Southeast Asia program) Cornell University, Ithaca New York.

Luce, G.H. (1965) "Danaw: A dying Austroasiatic Language" *Lingua* 14:98-129.

Malmqvist, Göran (1962) "Studies in Western Mandarin phonology" *Bulletin of the Museum of Far Eastern Antiquities* (Stockholm) 34:129-192.

Maran, Laraw (1971) *Burmese and Jingpho: A study of tonal linguistic processes (Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics 4)*, Center for Asian Studies, University of Illinois, Urbana.

Matisoff, J.A. (1972) *The Loloish Tonal Split Revisited* (Research Monograph 7) Center for South and Southeast Asia Studies. University of California, Berkeley.

Mohring, Hans (1972) "Einige Gedanken zum Ursprung der Töne im Vietnamesischen" *Zeitschrift fur Phonetik, Sprachwissenschaft und Kommunikationsforschung* 25:232-242.

Norman, J. (1973) "Tonal development in Min" *Journal of Chinese Linguistics* 1 : 222-238.

Pike, Eunice V. (1974) "A multiple stress system versus a tone system" *IJAL* 40:169-175.

Pinnow, Heinz-Jürgen (1957) "Sprachgeschichtliche Erwägungen zum Phonemsystem des Khmer" *Zeitschrift für Phonetik* 10:378-391.

Pulleyblank, E.G. (1970) "Late Middle Chinese" *Asia Major* 15:197-239.

Pulleyblank, E.G. (1974) "Late Middle Chinese and Southern dialects" Paper presented at the 6th International Conference on Sino-Tibetan Language and Linguistics.

Sarawit, M. (1973) *The Proto-Tai Vowel System* Ph.D. dissertation. University of Michigan, Ann Arbor.

Shorto, H.L. (1967) "The register distinctions in Mon-Khmer languages" *Wissenschaftliche Zeitschrift der Karl-Marx-Universität Leipzig, Gesellschafts- und Sprachwissenschaftliche Reihe* 16/1-2:246-248.

Sivertsen, Eva (1956) "Pitch problems in Kiowa" *IJAL* 22:117-130.

Steinbergs, Aleks (1975) "Some problems concerning the origin of the Latvian broken tone" *Studies in the Linguistic Sciences* (University of Illinois) 5.2:157-185.

Trager, Felicia (1971) "The phonology of Picuris" *IJAL* 37:29-33.

Waterhouse, Viola ed. (1967) *Phonemic Systems of Colombian Languages* (Summer Institute of Linguistics Publications in Linguistics and Related Fields No 14) S.I.L. University of Oklahoma, Norman, Oklahoma.

West, Birdie and Betty Welch (1967) "Phonemic system of Tucano" in Waterhouse (1967).

Welmers, William E. (1962) "The phonology of Kpelle" *Journal of African Languages* 1:69-93.

Yang, Paul (1967) "Elements of Hakka phonology" *Monumenta Serica* 26:305-351.

Zee, Eric (1977) "The effect of F_0 on the duration of [s]" Paper presented to the 93rd Meeting of the Acoustical Society of America, College Park. Pa.

Duration and Intensity as Correlates of F_0

Eric Zee

[Department of Linguistics
University of California, Los Angeles]

1. Introduction

Ferrein (1741) was probably the first to conduct an experiment on the correlation of pitch with air pressure. He connected the trachea of an excised larynx of a dog to a bellows, and used threads and weights to manipulate the positions of the arytenoid cartilages to create different vocal cord tensions. A manometer was used to measure air pressure. With constant air pressure, pitch increased as he increased the tension of the vocal cords. With increases in air pressure further increases in pitch were produced. That increase in subglottal pressure, or increase in vocal cord tension, increases pitch has been confirmed by more recent studies (Van den Ber, 1957; Ladefoged, 1963; Stevens, 1973, 1975). Ohala and Ewan (1972) reported that, given a comparable fundamental frequency interval, there is a marked tendency for fundamental frequency lowering to be faster than fundamental frequency raising. The present study further investigates these correlations by using a tone language. More precisely, it is concerned with the correlation of fundamental frequency with both duration and intensity. The tone language chosen is Taiwanese which is ideal for investigation as it has high, mid, low, high-falling and low-rising tones. Of the two previous instrumental studies on tones in Taiwanese, Chiang's spectrographic analysis (1967) presents the phonetic shapes of the tones in Taiwanese but it does not deal with the correlation of fundamental frequency with either duration or intensity; Weingartner (1970) analyzes the Taiwanese tones in terms of duration and intensity, but unfortunately he does not provide information regarding his experimental procedures, the instruments he used and his statistical method (for a review of Weingartner, 1970, see Cheng, 1972).

2. Procedure

The five Taiwanese tones used for investigation are high, mid, low, high-falling and low-rising. All these tones occur on the syllable [si]:

high	[si˥]	'silk'
mid	[si˧]	'yes'
low	[si˩]	'four'
high-falling	[si˥˩]	'to die'
low-rising	[si˩˧]	'time'

A word list containing 75 tokens of these [si] words (5 words x 15 repetitions) was prepared. Two male native Taiwanese speakers (both university students) were instructed to read the list at a normal rate of speech. The test word was placed in a carrier frame:

[guá kóŋ ___ tài kēi tʰIa]

'I say _____ everyone listens'

The recording was made in a sound treated room at a single session for each speaker. The tapes were analyzed, using a PDP-12 computer at the UCLA Phonetics Laboratory. Fundamental frequency measurement was obtained every 10 msec, using the Cepstrum method. RMS amplitude was also taken every 10 msec, using a square window. Duration measurements were made on the waveform displayed on the computer screen.

Since utterances have different lengths, it is necessary to select comparable points at which the differing fundamental frequencies may be compared. Accordingly, the values of fundamental frequency and RMS amplitude for each token were divived into five parts, each containing as nearly as possible equal numbers of 10 msec interval. The fundamental frequency and RMS values of each part were then taken to be the means of the sets of 10 msec intervals within that part. The 15 repetitions of each syllable were then averaged so as to give mean values for each of the five parts. Thus, the first point on any curve in Figure 1 and Figure 2 represents the mean of the mean values for the initial time intervals of 15 repetitions of that syllable. The second point is the mean of the mean for the second set of time intervals and so on. Each bar in Figure 3 and Figure 4 represents the mean of 15 tokens for each tone class. The intensity value of the low tone was taken as reference (0 db) to compute the intensity values of other tones.

3. Results

3.1. Intensity

As shown in Figure 1 and Figure 2, there is a good correlation between fundamental frequency and intensity levels. In general, the intensity curves fall and rise with the fundamental frequency curves. However, we can see that the onset points in the intensity curve for both the high and the mid tones do not correlate so well, although, as shown in Figure 3 and Figure 4, the average RMS value is still correlated with the average fundamental frequency for both speakers as the high tone has a higher overall intensity than the mid tone and the mid tone has a higher overall intensity than the low tone. As far as the contour tones are concerned, the intensity curves again correlate closely with the direction of fundamental frequency change, although the offset points on the intensity curve for the low-rising tone do not correlate so well for both speakers. For these contour tones, as shown in Figure 3 and Figure 4, the average RMS value is also correlated with the average fundamental frequency as the high-falling tone has a higher overall intensity than the low-rising tone.

3.2. Duration

As shown in Figure 3, for Speaker 1, the high tone has a longer duration than the mid tone, and the mid tone has a longer duration than the

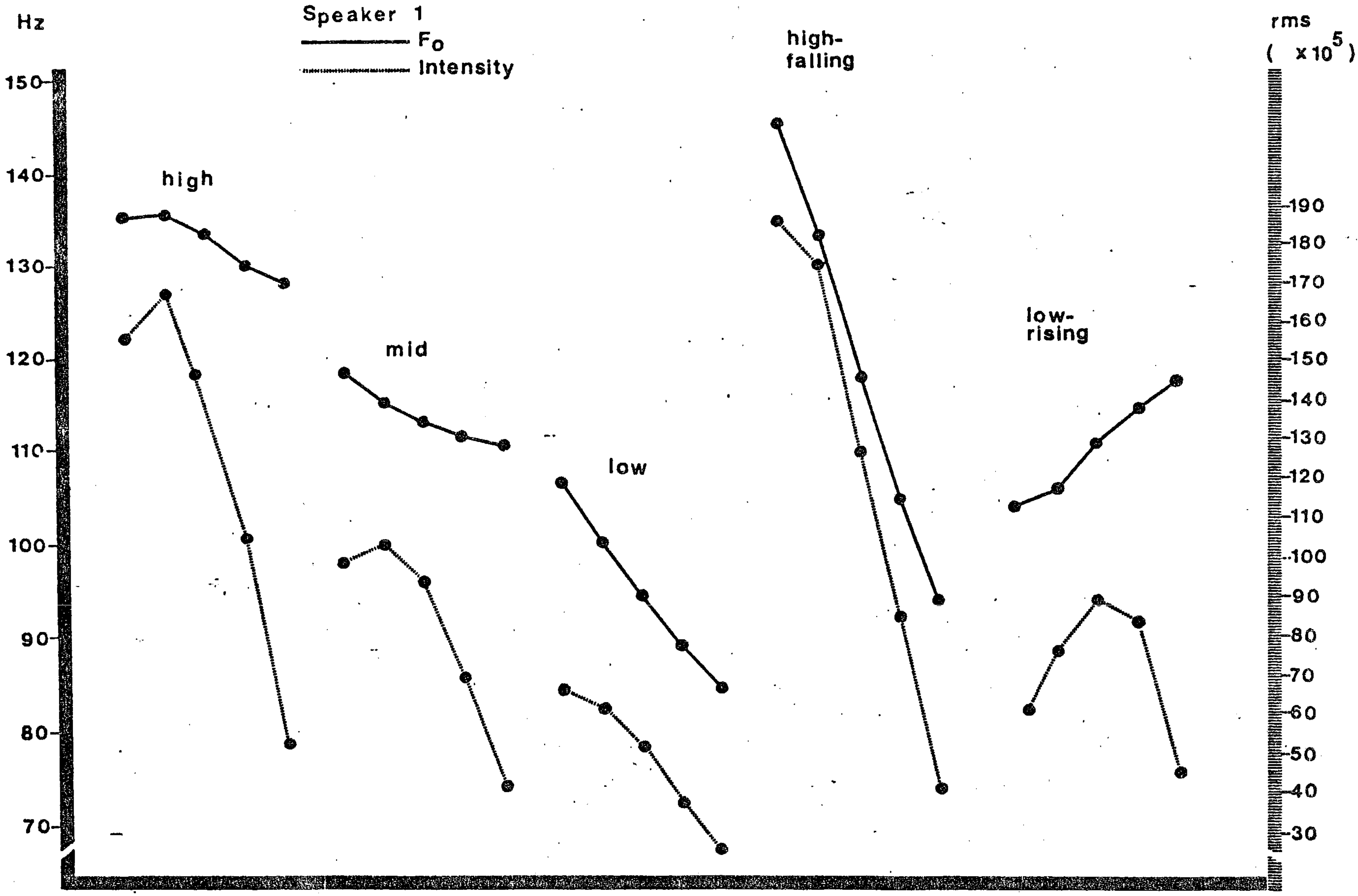
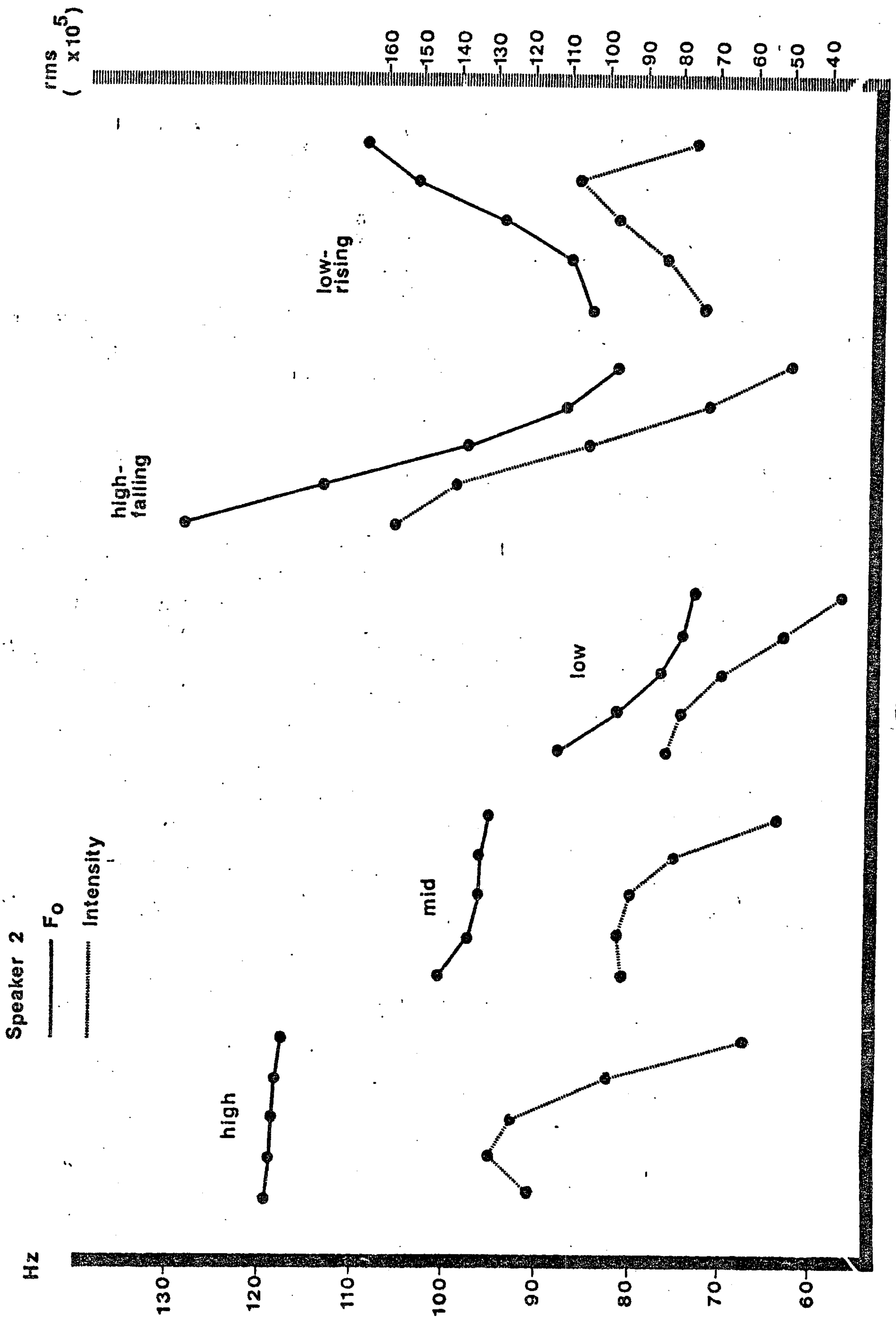

Hz
Speaker 1
Fo
Intensity
rms
(x10^5)
150
140
130
120
110
100
90
80
70
high
mid
low
high-
falling
low-
rising
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30

Figure 1



Speaker 2
F_o
Intensity
high
mid
low
high-falling
low-rising
Hz
130
120
110
100
90
80
70
60
rms (x10^5)
160
150
140
130
120
110
100
90
80
70
60
50
40

Figure 2

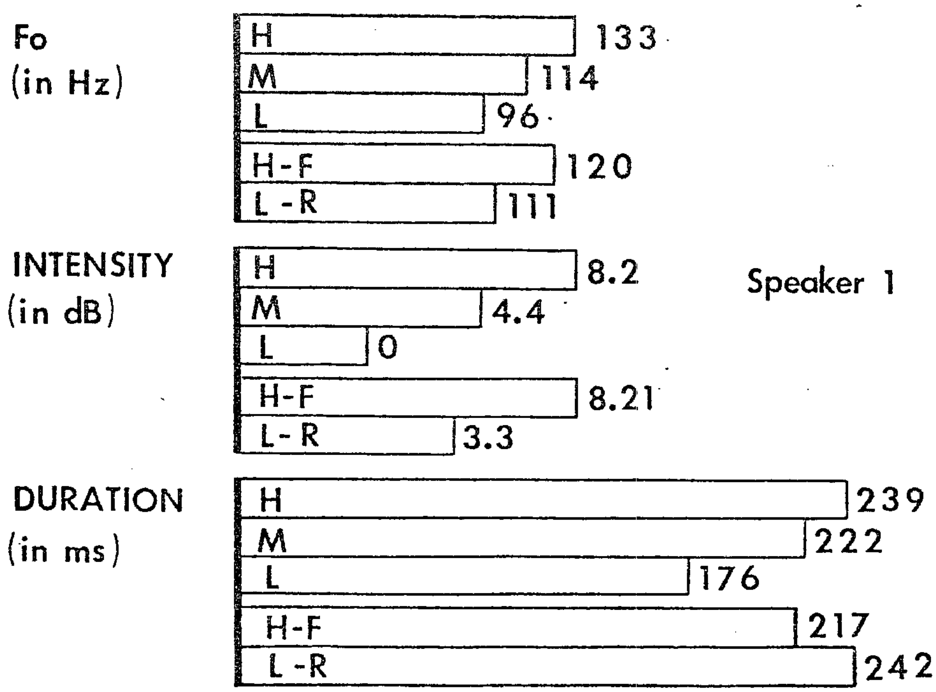


Figure 3

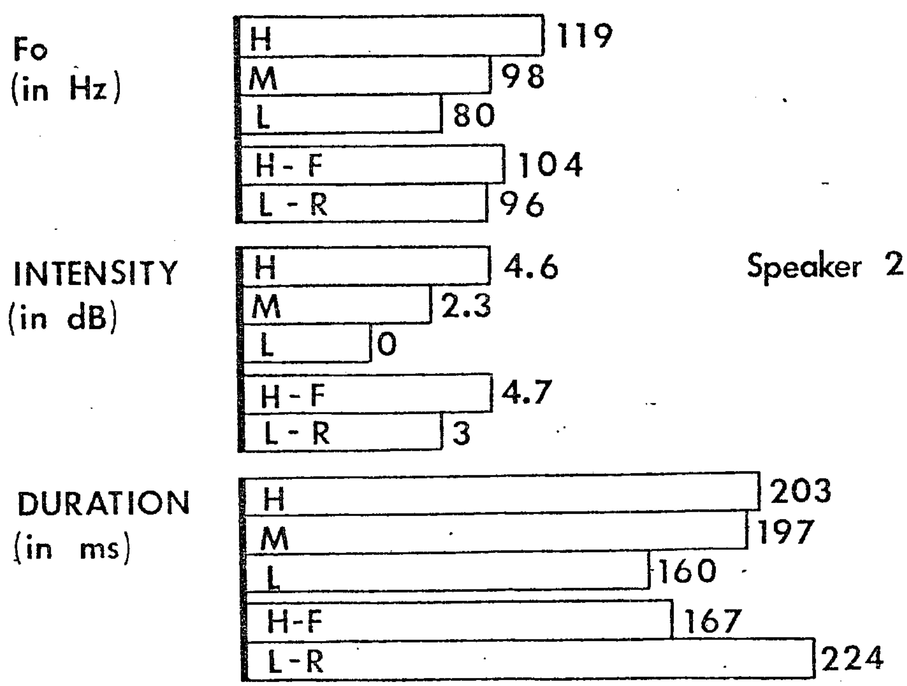


Figure 4

low tone. The low-rising tone has a longer duration than the high-falling tone. T-test results reveal that:

(a) the difference in duration between the high and the mid tone is non-significant;

(b) the difference in duration between the high and the low tones and between the mid and the low tones are significant at the .01 level for both speakers;

(c) the difference in duration between the low-rising tone and the high-falling tone is significant at the .01 level.

Table I shows the results of grouped data T-test for durations of the different tones for Speaker 1.

Similar results were also found in Speaker 2 as shown in Figure 4. The high tone has a higher overall intensity and longer duration than the mid tone, and the mid tone has a higher overall intensity and longer duration than the low tone. The high-falling tone has a higher overall intensity than the low-rising tone, although the low-rising tone has a longer duration. Table II shows the results of grouped data T-test for durations for Speaker 2.

3.3. The low tone as a low-falling tone

It should be noticed that the low tone may be considered a low-falling tone, as the fundamental frequency range covered by the low tone is substantially greater than that covered by the high or the mid tone for both speakers as shown in Table III.

For Speaker 1, the fundamental frequency range difference is 7 Hz for the high tone, 7 Hz for the mid tone and 21 Hz for the low tone. For Speaker 2, the fundamental frequency range difference is 1 Hz for the high tone, 5 Hz for the mid tone and 14 Hz for the low tone. It seems appropriate to consider the high and the mid tones the level tones and the low tone the low-falling tone.

3.4. The high-falling tone vs. the low-rising tone

Figure 1 and Figure 2 show that the high-falling tone covers a much greater range than the low-rising tone: 52 Hz for the high-falling tone vs. 14 Hz for the low-rising tone for Speaker 1, and 46 Hz for the high-falling tone vs. 24 Hz for the low-rising tone for Speaker 2. Since the duration of the high-falling tone is shorter than that of the low-rising tone for both speakers (Figure 3 and Figure 4), the rate of change of fundamental frequency is thus substantially greater for the high-falling tone than for the low-rising tone for both speakers. For Speaker 1, the linear rate of fundamental frequency change is 235 Hz/sec in the high-falling tone and 57 Hz/sec in the low-rising tone. For Speaker 2, the linear rate of change in the high-falling tone is 275 Hz/sec and 152 Hz/sec in the low-rising tone. These are shown in Table IV.

4. Discussion

4.1. That the high tone has a higher overall intensity than the mid tone and the mid tone has a higher overall intensity than the low tone may be because higher frequency correlates with both greater tension of vocal cords and higher subglottal pressure. The vocal cords will be driven by a heightened subglottal pressure which will lead to higher intensity. Increased subglottal pressure (Ladefoged, 1963) and increased tension of the vocal cords (Stevens, 1973, 1975) have been considered as separate factors, either one of which acting alone is capable of producing an increase of fundamental frequency. In natural speech increased subglottal pressure and increased tension of the vocal cords may operate together to get maximum efficiency. An increase of each may be capable of producing a great increase in fundamental frequency. However, there are cases when increased tension of the vocal cords must operate alone to increase fundamental frequency. As we have indicated in 3.1., the onset points of the intensity curve for both the high and the mid tones and the offset points of the low-rising tone do not correlate so well with the directions of fundamental frequency change. In fact, in both the high and the mid tones, the second half of the intensity curve falls sharply although the fundamental frequency curve falls in a much lesser degree for both speakers (Figure 1 and Figure 2). These indicate that a higher fundamental frequency may be produced without a corresponding higher subglottal pressure. Particularly in the case of the low-rising tone where the intensity curve falls sharply while the fundamental frequency curve rises. Thus, increased tension of the vocal cords alone, perhaps with the assistance of larynx elevation, is capable of producing a higher fundamental frequency.

4.2. For both speakers, the low-rising tone is longer than the high-falling tone and the range covered by the high-falling tone is greater than the range covered by the low-rising tone. As can be seen in Tables III and IV, the rate of change of fundamental frequency is greater for the high-falling tone than for the low-rising tone. The reason may be that it requires more time for the vocal cords to increase the rate of vibration than to decrease it. This observation is supported by the result obtained by Ohala and Ewan (1972) that there is a marked tendency for pitch lowering to be faster than for pitch raising for a comparable pitch interval. Similarly, Sundberg (1973) has demonstrated that unlike singers, untrained subjects perform pitch drops considerably faster than pitch elevations. Ohala's explanation for this phenomenon is that it requires more physiological effort to raise a pitch than to lower it. As shown in Figure 1 and Figure 2, the intensity curve falls sharply at the offset point, whereas the fundamental frequency curve rises, that is, fundamental frequency rises without the presence of a higher subglottal pressure. Perhaps it is the lack of an accompanying higher subglottal pressure that slows down the speed of fundamental frequency change in the low-rising tone. It is likely that the extra physiological effort that Ohala refers to is needed when the subglottal pressure begins to drop during the process of fundamental frequency raising. To compensate for the loss of acoustic energy, such extra physiological effort is

Table I. Results of grouped data T-test for durations (in msec) of (a) high and mid tones, (b) high and low tones, (c) mid and low tones, and (d) high-falling (H-F) and low-rising (L-R) tones (Speaker 1)

	(a)		(b)		(c)		(d)	
	High	Mid	High	Low	Mid	Low	H-F	L-R
MEAN	243	229	243	176	229	176	218	246
STD DEV	28	17	28	14	17	14	19	27
T-SCORE	1.63		8.22		9.34		3.12	

Table II. Results of grouped data T-test for durations (in msec) of (a) high and mid tones, (b) high and low tones, (c) mid and low tones, and (d) high-falling (H-F) and low-rising (L-R) tones (Speaker 2)

	(a)		(b)		(c)		(d)	
	High	Mid	High	Low	Mid	Low	H-F	L-R
MEAN	203	196	203	159	196	159	166	224
STD DEV	18	14	18	20	14	20	17	14
T-SCORE	1.07		6.07		5.63		9.73	

Table III. Fundamental frequency (in Hz) ranges of high, mid and low tones and their range differences (Speaker 1 and Speaker 2)

	Speaker 1		Speaker 2	
	F_0 range	Difference	F_0 range	Difference
High	135-128	7	119-118	1
Mid	118-111	7	101- 96	5
Low	107- 86	21	88- 74	14

Table IV. Fundamental frequency (in Hz) ranges and linear rate of fundamental frequency change in high-falling and low-rising tones (Speaker 1 and Speaker 2)

	Speaker 1		Speaker 2	
	F_0 range	Linear rate of F_0 change	F_0 range	Linear rate of F_0 change
High-falling	54	235 Hz/sec	46	275 Hz/sec
Low-rising	14	57 Hz/sec	24	152 Hz/sec

Table V.

	Conditioning		Phonetic realization
	tone 1	→	duration/intensity 1
	tone 2	→	duration/intensity 2
intrinsic duration/ intrinsic intensity of a vowel ➡	tone 3	→	duration/intensity 3
	tone 4	→	duration/intensity 4
	tone 5	→	duration/intensity 5
	.		.
	.		.
	.		.

needed for further increasing the tension of the vocal cords and further raising the larynx in order that the fundamental frequency may be raised.

4.3. That the high tone and the mid tone have a longer duration than the low tone may be because the low tone is actually a low-falling tone as indicated in 3.3. Statistical results have shown that the difference in duration between the high and the mid tones is non-significant, which may be due to the fact that they are both level tones. It seems that the difference in duration between different tone classes is primarily determined by the shape, rather than the level, of the fundamental frequency curve. As the level (high and mid) tones are slightly falling, we would expect the duration of these level tones to come in between those of the low-rising tone and of the falling (high-falling and low-falling) tones. As a matter of fact, for both speakers, the level tones are shorter than the low-rising tone and longer than the falling tones.

4.4. Lehiste refers to intrinsic duration as duration of a segment as determined by its phonetic quality (1970:18); and to intensity of a vowel as intensity considered in relation to its phonetic quality (1970: 120). These definitions are only partially applicable to a tone language, such as Taiwanese, for in a tone language the phonetic realization of a vowel must include both its segmental quality and its relative pitch or tone. It may be than that the 'intrinsic duration and intensity' of a tone interacts with the intrinsic duration and intensity constrained by the vowel quality, or that the intrinsic duration and intensity of a vowel in a tone language is determined by its phonetic quality as further conditioned by its tone. (See Table V.)

It seems then the intrinsic duration or the intrinsic intensity of a vowel in a tone language always remain recessive and will only manifest itself as a variant.

5. Summary of results

The purpose of this experiment was to determine whether fundamental frequency correlates with both intensity and duration. A tone language, Taiwanese, was used for investigation. The results from analyzing the data produced by two male Taiwanese speakers and discussion of these results indicate:

(a) The high tone has an overall higher intensity than the mid tone and the mid tone has an overall higher intensity than the low tone.

(b) The high tone and the mid tone have a longer duration than the low tone. The high tone has a longer duration than the mid tone, but the difference between them is non-significant.

(c) The high-falling tone has an overall higher intensity than the low-rising tone, although the low-rising tone has a longer duration.

(d) If the low tone is considered a low-falling tone, the low-rising tone is longer than the level (high and mid) tones and the level tones are longer than the falling (high-falling and low-falling) tones.

(e) The linear rate of fundamental frequency change is greater for the

high-falling tone than for the low-rising tone.
(f) The high-falling tone covers a greater range of fundamental frequency than the low-rising tone.
(g) Regardless of the tone classes, the intensity tends to drop at the offset point of a tone.
(h) The difference in duration between different tone classes is primarily determined by the shape of the fundamental frequency contour.
(i) The intrinsic duration and the intrinsic intensity of a vowel in a tone language are conditioned by the tone that the vowel carries.

Acknowledgment

I would like to thank the members of UCLA Tone Seminar for many helpful comments. Special thanks are owed to Jean-Marie Hombert, Peter Ladefoged, Victoria Fromkin and Steve Greenberg. This research was supported by NSF.

References

Chen, T.M. (1972) Review of Tone in Taiwanese, an instrumental investigation, by F.F. Weingartner. *Lingua* 30: 2, 3, 285-294.

Chiang, H.T. (1967) Amoy-Chinese tones. *Phonetica* 17, 100-115.

Ladefoged, P. (1963) Some physiological parameters in speech. *Language and Speech* 6, 109-119.

Lehiste, I. (1970) *Suprasegmentals* Cambridge, Massachusetts: M.I.T. Press.

Ohala, J. and W. Ewan (1972) Speech of pitch change. Paper presented at the 94th ASA-Meeting, Autumn,]972, Miami Beach, Florida.

Stevens, K.N. (1973) Theory of vocal cord vibration and its relation to laryngeal features. Paper presented at the 86th ASA-Meeting, Autumn, 1973, Los Angeles, California.

Stevens, K.N. (1975) Physics of laryngeal behavior and larynx modes. Paper presented at the Seminar on Larynx and Language, International Congress of Phonetics, August, 1975, Leeds, England.

Sundberg, J. (1973) Data on maximum speed of pitch changes. *STL-QPSR*, 4/1973, 39-47.

Van den Berg, J. (1957) Subglottalic pressures and vibrations of the folds. *Folia Phoniatrica* 9, 65-71.

Weingartner, F. (1970) Tones in Taiwanese, an instrumental investigation. Taipei, Taiwan: National Taiwan University, College of Arts, Monograph No. 2.

The effect of F_o on the duration of [s]

Eric Zee

[Department of Linguistics, University of California,

Los Angeles]

1. Introduction

There have been studies concerning the influence of vowels on the duration of an adjacent fricative /s/. Klatt (1974) reported that in English there is a tendency for the duration of prestressed /s/ to be longer than the duration of preunstressed /s/. Schwartz (1970) found that the fricative /s/ in the environment of a high front vowel /i/ is longer than in the environment of a low vowel /a/. The purpose of the present study was to determine whether the duration of the prevocalic fricative /s/ is influenced by the tone on the following vowel. It also examines whether temporal compensation exists between the duration of the fricative and the duration of the vowel.[1] A tone language, Taiwanese, was used for investigation. In this Chinese dialect, all its five contrastive tones; high, mid, low, low-rising and high-falling, occur on the syllable /si/:

/sí/	(high)	meaning	'silk'
/sī/	(mid)	"	'yes'
/sì/	(low)	"	'four'
/sǐ/	(low-rising)	"	'time'
/sî/	(high-falling)	"	'to die'

2. Procedure

A word list containing 75 tokens of these syllables was prepared, that is, 15 repetitions for each tone class. Two male Taiwanese speakers, both in their early thirties, were instructed to read the list at a normal rate of speech. The test word was placed in a carrier frame as follows:

guá tçí-má tʰàʔ ___ taì-keī tʰíá

'I now read ___ everyone listens'

The recording was made in a sound treated room in a single session for each speaker. The data were analyzed, using a PDP-12 computer at the UCLA Phonetics Laboratory. Duration measurements of both segments of the syllable /si/ were made directly on the waveform.

3. Results

Figure 1 shows the results of the duration measurements of both the fricative and the vowel of the syllable produced with different tones for both speakers. Each blank bar on the left represents the mean duration of 15 tokens of the fricative /s/ for each tone class. Each dark bar on the right represents the mean duration of 15 tokens of the vowel /i/ for each tone class. Durations are given in milliseconds.

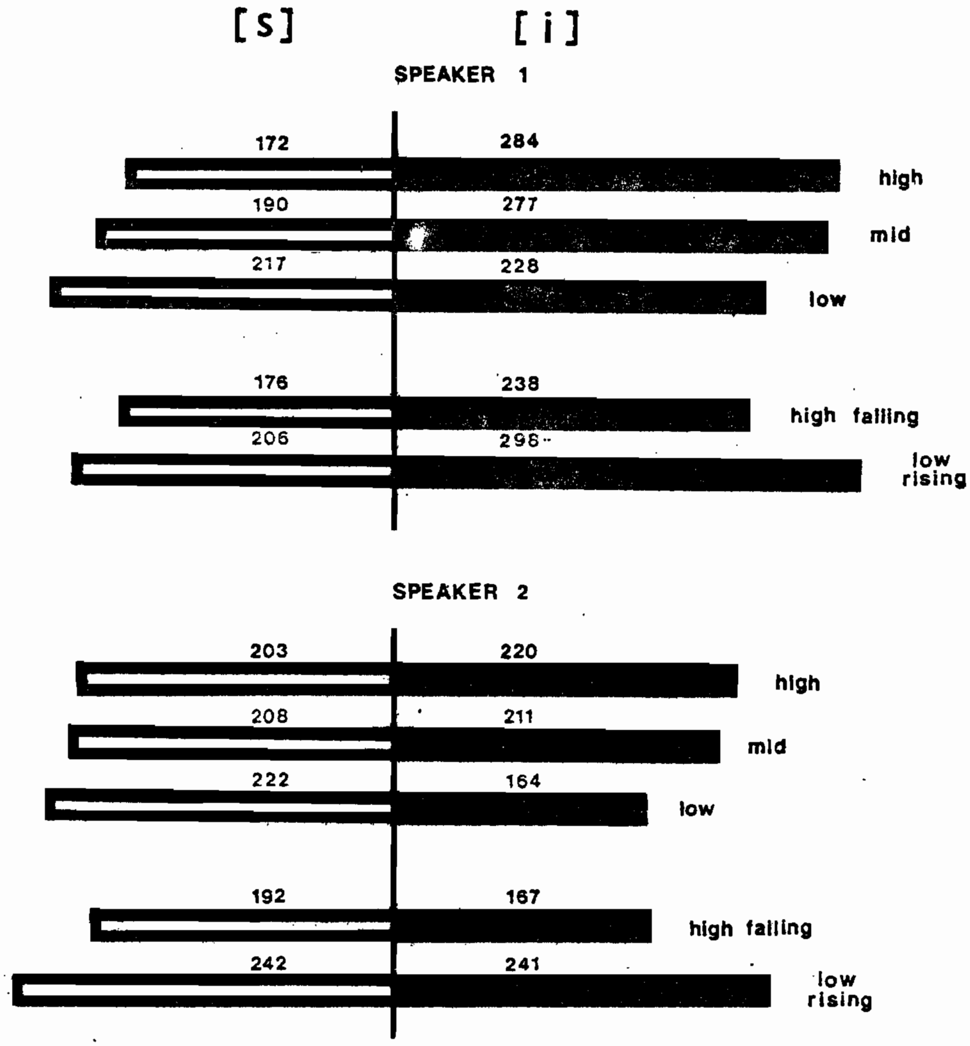
An examination of the durations of /s/ for both speakers shows that the duration of /s/ associated with either the low tone or the low-rising tone is longer than the duration of /s/ associated with the high, the mid, or the high-falling tone.

T-test results reveal that the difference in duration between /s/ associated with the low tone and /s/ associated with the high, the mid, or the high-falling tone is significant.[2] The difference in duration between /s/ associated with the low-rising tone and /s/ associated with the high, the mid, or the high-falling tone is also significant.[3]

The fact that the duration of the fricative /s/ associated with either the low or the low-rising tone is longer than the fricative /s/ associated the high, the mid, or the high-falling tone indicates that there may be a correlation between the duration of the fricative /s/ and the onset value of the tone on the following vowel /i/.

4. Disucssion

In Zee and Hombert (1976), we presented the duration pattern of the vocalic portion of the syllable /si/ produced with the same five Taiwanese tones, which is similar to the duration pattern obtained in the present study as shown on the right side of Figure 1. The low-rising tone is longer than the high and the mid tones, and the high and the mid tones are longer than the low and the high-falling tones. It was claimed that the duration of the vocalic portion of the syllable /si/ is determined by the shape of the tone. Figure 2 shows the duration of the entire syllable produced with five different tones. Each is the mean of 15 tokens. The only difference between Figure 1 and Figure 2 is that in Figure 1 the lineup point is the onset of a vowel and in Figure 2 is the beginning of the fricative /s/. As before, the blank bars represent the fricative portion /s/ and the dark bars represent the vocalic portion /i/. We can see for both speakers the duration of the syllable /si/ associated with the low-rising tone is longer than the duration of the syllable associated with any other tones; and the duration of the syllable associated with the high-falling tone is the shortest of all. Although the duration of the vocalic portion associated with the low tone is similar to the duration of the vocalic portion associated with the high-falling tone, the duration of the syllable is longer for the low tone than for the high-falling tone as the duration of the fricative portion for the low tone is longer than the high-falling tone. Thus, the duration of a syllable, such as /si/, varies according to the tone that the vowel carries, and furthermore, it is the onset of the tone



[s]
[i]
SPEAKER 1
172
284
high
190
277
mid
217
228
low
176
238
high falling
206
296
low rising
SPEAKER 2
203
220
high
208
211
mid
222
164
low
192
167
high falling
242
241
low rising

Fig. 1

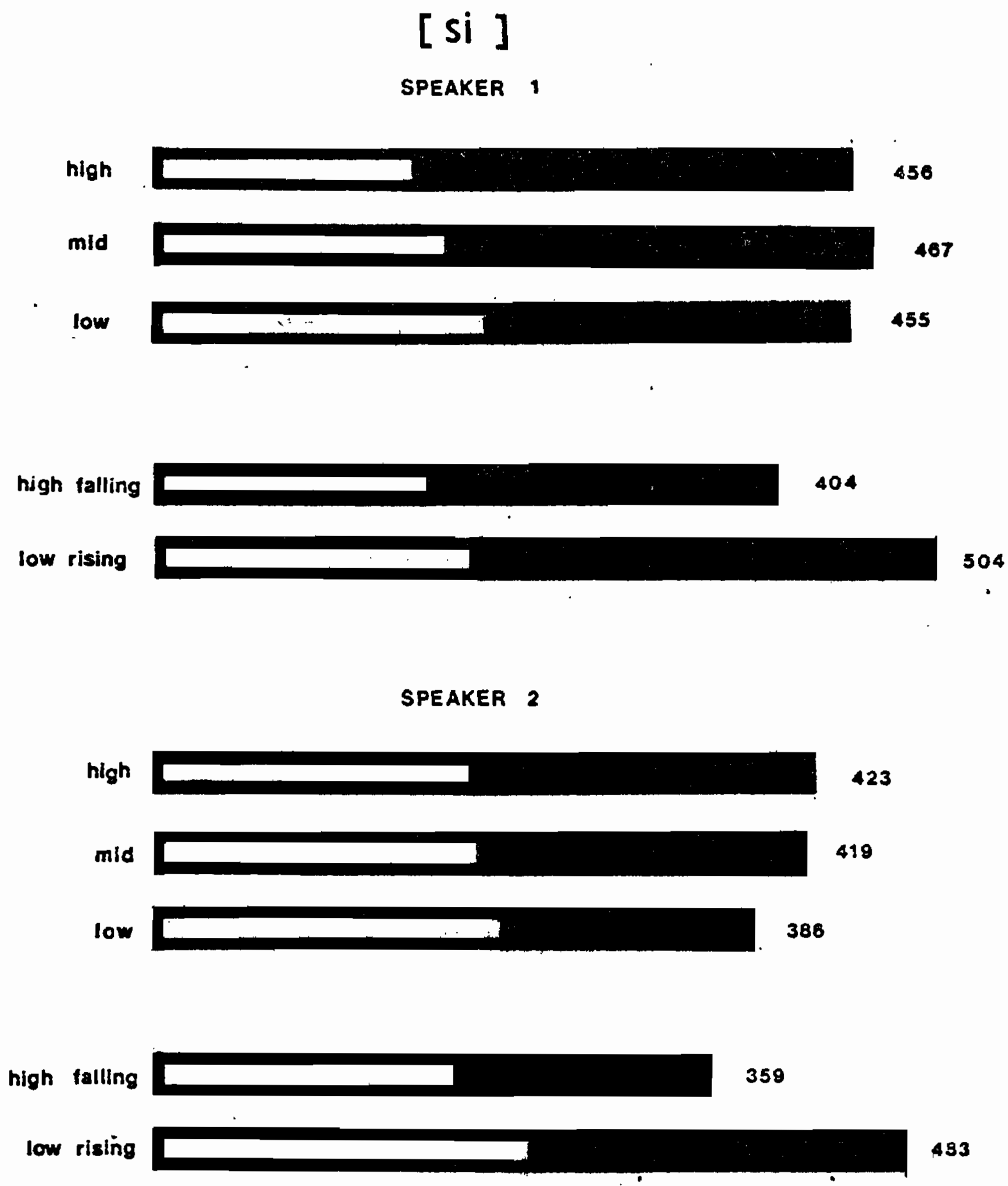

[si]
SPEAKER 1
high
456
mid
467
low
455
high falling
404
low rising
504
SPEAKER 2
high
423
mid
419
low
386
high falling
359
low rising
483

Fig. 2

that determines the duration of the fricative portion and the shape of the tone that determines the duration of the vocalic portion.

Temporal compensation within the sequences of the segments /s/ and /i/ associated with any one of the five tone classes was tested. To see whether there is temporal compensation within two segments is to find out whether the durations of the two segments are negatively correlated (Lehiste, 1970). There is a negative correlation between the durations of two segments if the relative variance of the duration of the sequence of two segments is less than the sum of the relative variances of the segments considered separately. Relative variance is simply variance divided by average duration, a concept introduced by Allen (1969). The result is, within each tone class, the relative variance of the duration of the syllable /si/ as one unit is greater than the sum of the relative variances of the segments /s/ and /i/ considered separately. This is true for both speakers. This indicates that there is not a negative correlation between the durations of the two segments. It follows that, within each tone class, temporal compensation does not exist between the duration of the segment /s/ and the duration of the segment /i/ in the syllable /si/.

A negative correlation between the durations of two segments has been taken to imply that the two are programmed as one unit at some higher level at which articulatory sequences are programmed (Lehiste, 1970). Our result would seem to imply that the segments /s/ and /i/ in the syllable /si/ produced with any tone are programmed not as one unit but as two separate units at some higher level, as there is not a negative correlation between the durations of the two segments /s/ and /i/. However, the fact that the duration of the segment /s/ is determined by the tone on the following vowel /i/ indicates that the segments /s/ and /i/ in the syllable /si/ produced with any tone cannot be programmed as two separate units at some higher level at which articulatory sequences are programmed, but as one unit. Thus, <u>in a tone language</u>, such as Taiwanese a negative correlation between the durations of two segments, a fricative /s/ followed by a vowel /i/, does not constitute a necessary condition for determining whether the two segments are programmed as two separate units, or as one unit, at some higher level at which articulatory sequences are programmed.

5. Summary and Conclusions

The purpose of this study was to determine whether the duration of the prevocalic fricative /s/ is influenced by the tone on the following vowel and to examine whether temporal compensation exists between the duration of the fricative and the duration of the vowel within a certain tone class. A tone language, Taiwanese, was used for investigation. The results from analyzing the data produced by two male Taiwanese speakers indicate that:

(1) the duration of the fricative /s/ is influenced by the tone on the following vowel, and it varies according to the onset value of the tone.

(2) within each tone class, temporal compensation does not seem to exist between the durations of the segments /s/ and /i/ in the syllable /si/.

(3) in a tone language, such as Taiwanese, a negative correlation between the durations of two segments, such as a fricative /s/ followed by a vowel /i/, does not constitute a necessary condition for determining whether the two segments are programmed as two separate units, or as one unit, at some higher level at which articulatory sequences are programmed.

Footnotes

1. It refers to the temporal compensation between segments /s/ and /i/ within one single tone class, not across different tone classes.

2. (a) The difference in duration between /s/ associated with the low tone and /s/ associated with the high tone is significant at the .005 level for Speaker 1 and .025 level for Speaker 2.

(b) The difference in duration between /s/ associated with the low tone and /s/ associated with the mid tone is significant at the .005 level for Speaker 1. However, only a trend in this direction was obtained for Speaker 2 (.1 level).

(c) The difference in duration between /s/ associated with the low tone /s/ associated with the high-falling tone is significant at the .005 level for both Speaker 1 and 2.

3. The difference in duration between /s/ associated with the low-rising tone and /s/ associated with the high, the mid, or the high-falling tone is significant at the .005 level for both speakers.

References

Allen, George D. (1969) Structure of timing in speech production. Paper presented at the 78th Acoust. Soc. of Amer. Meet. Autumn, 1969, SanDiego, Calif.

Klatt, Dennis (1974) The duration of [s] in English words. *Journ. of Speech and Hearing Res.* 17, 51-63.

Lehiste, Ilse (1970) Temporal organization of spoken language. The Ohio State University, *Working Papers in Linguistics,* 4, 96-114.

Schwartz, Martin F. (1969) Influence of vowel environment upon the duration of /s/ and /š/. *Journ. of the Acoust. Soc. of Amer.* 46, 480 (L).

Zee, Eric and Jean-Marie Hombert (1976) Duration and intensity as correlates of F_o. Paper presented at the 92nd ASA meeting, Autumn, 1976, San Diego, Calif.

Acknowledgment

I would like to thank the members of the the UCLA Phonetics Seminar for their valuable comments. Special thanks are owed to Vicki Fromkin, Jean-Marie Hombert and Steve Greenberg.