V FROMKIN

Studies on Production and Perception of Tones

ucla working papers in phonetics 33

july 1976

UCLA Working Papers in Phonetics 33

August, 1976

The papers presented in this volume have been discussed at the UCLA tone seminar during the academic year 1975-76. This seminar was organized by Jean-Marie Hombert with the participation of Saeed Ali, Steve Anderson, Cathe Browman, Baruch Elimelech, Victoria Fromkin, Jack Gandour, Louis Goldstein, Steve Greenberg, S. Robert Greenberg, Richard Harshman, Dan Kahn, Peter Ladefoged, Philippe Lafage, Carl LaVelle, Ian Maddieson, Russ Schuh, William Welmers and Eric Zee.

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

This volume has been typed by Renee Wellin.

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

'Counterfeit Tones' in the Speech of Southern Thai Bidialectals

Jack Gandour

[University of California, Los Angeles]

Abstract

Evidence from tone errors produced by Southern Thai bidialectal reveals that the disordering mechanisms (perseveration, anticipation, transposition) that have been proposed to handle consonant and vowel errors may be extended without modification to tone errors. The tone errors further reflect or support the independence of tone features and some proposed universals regarding tone rules. Although some of the tone errors may be directly attributed to 'dialect interference', the majority of errors are better explained in terms of more general physiological and/ or perceptual factors (among others) that may intrude in the production of an utterance.*

0. Introduction

In recent years there has been a growing interest in speech errors because of the evidence they provide on questions concerning internalized grammars and the production of speech. All of the earlier studies, with the exception of Fromkin (1971), have discussed errors that involve consonants and vowels--i.e. 'segmental' errors, but have not disucssed errors of stress, tone or quantity--i.e. 'suprasegmental' errors. In this paper I will mainly discuss errors that involve tonal phenomena.

---

*The data for this paper was collected from December 1974 to September 1975 in Phuket province in southern Thailand. The research was made possible by fellowships granted by the Social Science Research Council - American Council of Learned Societies and the National Science Foundation. I wish to thank Sathorn Hararak, Kannika Pokapol, Suchow Pongsanon, Maitree Chantra, and Jariya Chingsakul for their help in collecting the data; Vicki Fromkin for her invaluable contributions to the paper both substantively and editorially; members of the 1975-1976 UCLA Tone Seminar--especially Peter Ladefoged, Ian Maddieson, and Steve Anderson for their helpful comments and suggestions. I alone, of course, take complete responsibility for the content of the paper; the views expressed herein do not necessarily reflect those of any of the above-mentioned persons or institutions. An earlier version of this paper was presented at the 1975 Winter Meeting of the Linguistic Society of America, San Francisco.

The earlier studies of consonant and vowel errors show rather convincingly that the errors are not merely random substitutions but instead are contextually conditioned and constrained by the structure of the language. Since tones, like consonants and vowels, also play a distinctive phonological role, we might expect tone errors to be subject to similar processes. I will attempt to show that tone errors can be accounted for the same kinds of explanations that have been proposed for consonant and vowel errors, and further, that they similarly provide evidence for linguistic constructs.

The 350 tone errors that were analyzed were produced by bidialectal speakers in Phuket province in southern Thailand when speaking Standard Thai (Siamese). They were collected by myself and by two native speakers of Phuket Thai and one native speaker of Standard Thai.[1]

The aim of the study was not to compare differences between tonal errors produced by bidialectal and monodialectal speakers but primarily to see if tonal errors were similar to or different from segmental errors. Since almost all speakers in the Phuket province speaking Standard Thai are bidialectal, a collection of errors in this region of necessity involved bidialectals. A comparison of types of errors must therefore await future collection of errors produced by native Standard Thai speakers. Since these unintentional errors which occurred in spontaneous speech were produced by bidialectals it was necessary to distinguish between errors due to dialect interference and those of a more general character.

The speakers of Southern Thai dialects are well aware of the mistakes they produce when speaking Standard Thai, as is shown by the fact that they characterize such errors as 'counterfeit speech'. This is an apt description since the errors constitute a deviation from the standard language; analogously, counterfeit money constitutes a deviation from a standard currency. We can thus call the tonal errors counterfeit tones.

As will become apparent, some of the errors, especially those that result from the misapplication of a tone rule of the Phuket dialect, can be accounted for in terms of dialect interference. These provide evidence for the 'psychological reality' of such phonological rules. However, the largest number of errors can be better explained by more general mechanisms causing a serial disordering of whole tones or tonal features, as suggested in Lashley's (1951) seminal paper on serial order in behavior.

---

[1] The tones in error, plus their context, were written down in Standard Thai orthography. This method of transcribing errors made it possible to unambiguously identify the substituted tones with any of the tones of either Standard Thai or Phuket Thai. It did not, however, allow for a reliable transcription of tone errors that could not be categorized in either of the two dialects. Such errors, if they do occur in any systematic way, were simply not accessible to this method of investigation.

Figure 1 presents a comparison of the tones of Standard Thai and Phuket Thai.

(1)		Standard Thai		Phuket Thai	
khaa	'leg'	[_/]	'rising'	[‾\]	'high-falling'
khaa	'galangal'	[_]	'low'		
paa	'forest'			[⌒\]	'mid-rising-falling'
paa	'to throw'	[—]	'mid'		
naa	'rice field'			[—\]	'low-falling'
....					
khaa	'value'	[\]	'falling'	[_/]	'rising'
paa	'aunt'			[—]	'high-mid-level'
chaa	'slow'	[‾]	'high'	[_]	'low-level'
mot	'used up'	[-]	'low'	[‾]	'high'
kop	'frog'			[-]	'mid'
mot	'ant'	[‾]	'high'	[_]	'low'
saat	'to splash'	[_]	'low'	[—]	'high-mid-level'
baat	'monetary unit'				
phaap	'picture'	[\]	'falling'	[_/]	'rising'

In non-stopped syllables Standard Thai has five contrastive tones: rising, low, mid, falling, and high; Phuket Thai has six distinctive tones: high-falling, mid-rising-falling, low-falling, rising, low-level, and high-mid-level.

If the tonal errors were simply due to dialect interference one would show a substitution of the Phuket tone for the Standard Thai tone in cognate words. Sixty-five percent of the errors cannot be accounted for in this way.

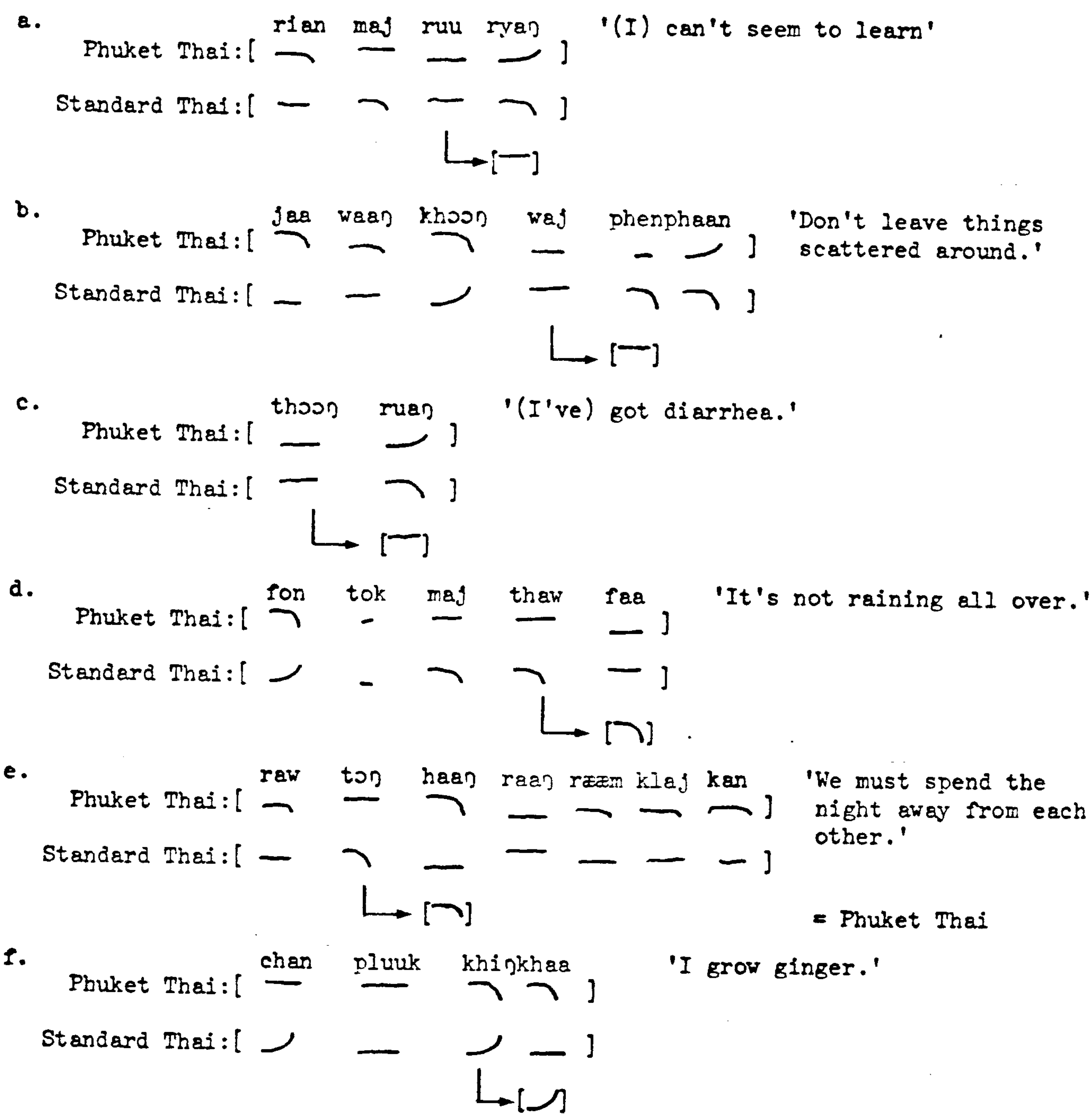
Since there are only five phonetically contrastive tones in Standard Thai, one might argue that the tone substitution errors produced by the Phuket bidialectal are due to chance. A chance hypothesis would be confirmed if we found that the errors were unsystematic and not subject to explanation: that is, if we found, for example, that the substituted tones were generally insensitive to the surrounding phonological and/or phonetic context of the utterance. But, as we shall see, these tone errors are not random and can be explained by the same principles that account for segmental errors--perseveration, anticipation, and transposition.

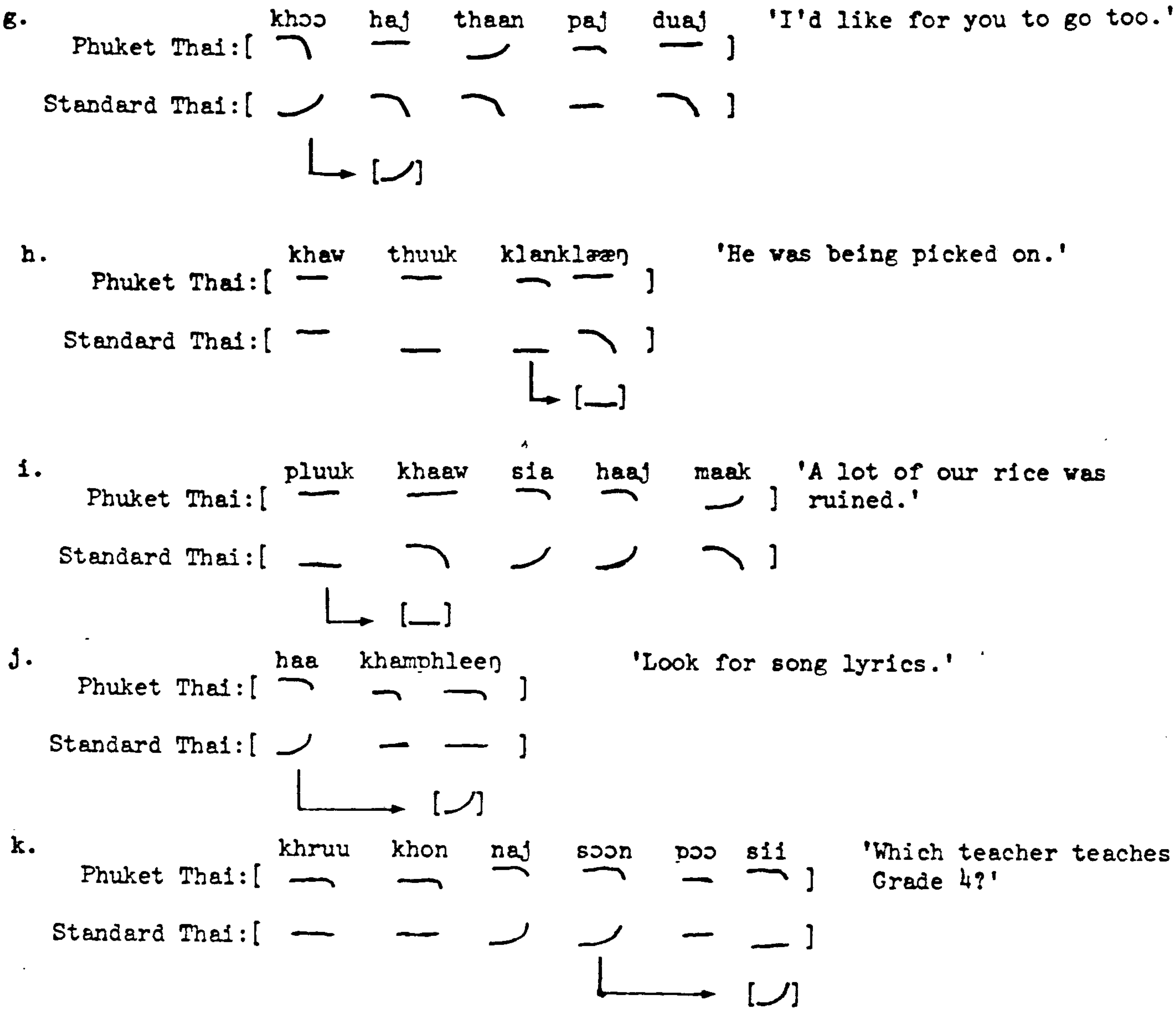
1. Kinds of Tone Errors

1.1 Errors of Perseveration

Errors of perseveration result in the substitution of one tone in perseveration of an identical or similar tone which occurs earlier in the utterance. The examples in 2 illustrate such errors.

(2)





2a-c show perseveration of a high tone, 2d-e a falling tone, 2f-g a rising tone, and 2h-i a low tone. No examples of perseveration of a mid tone occur in the present corpus. 2j-k further illustrate that the substituted tone need not occur on a contiguous syllable. As we shall see, this also holds true for errors of anticipation and transposition, similar to consonant and vowel errors.

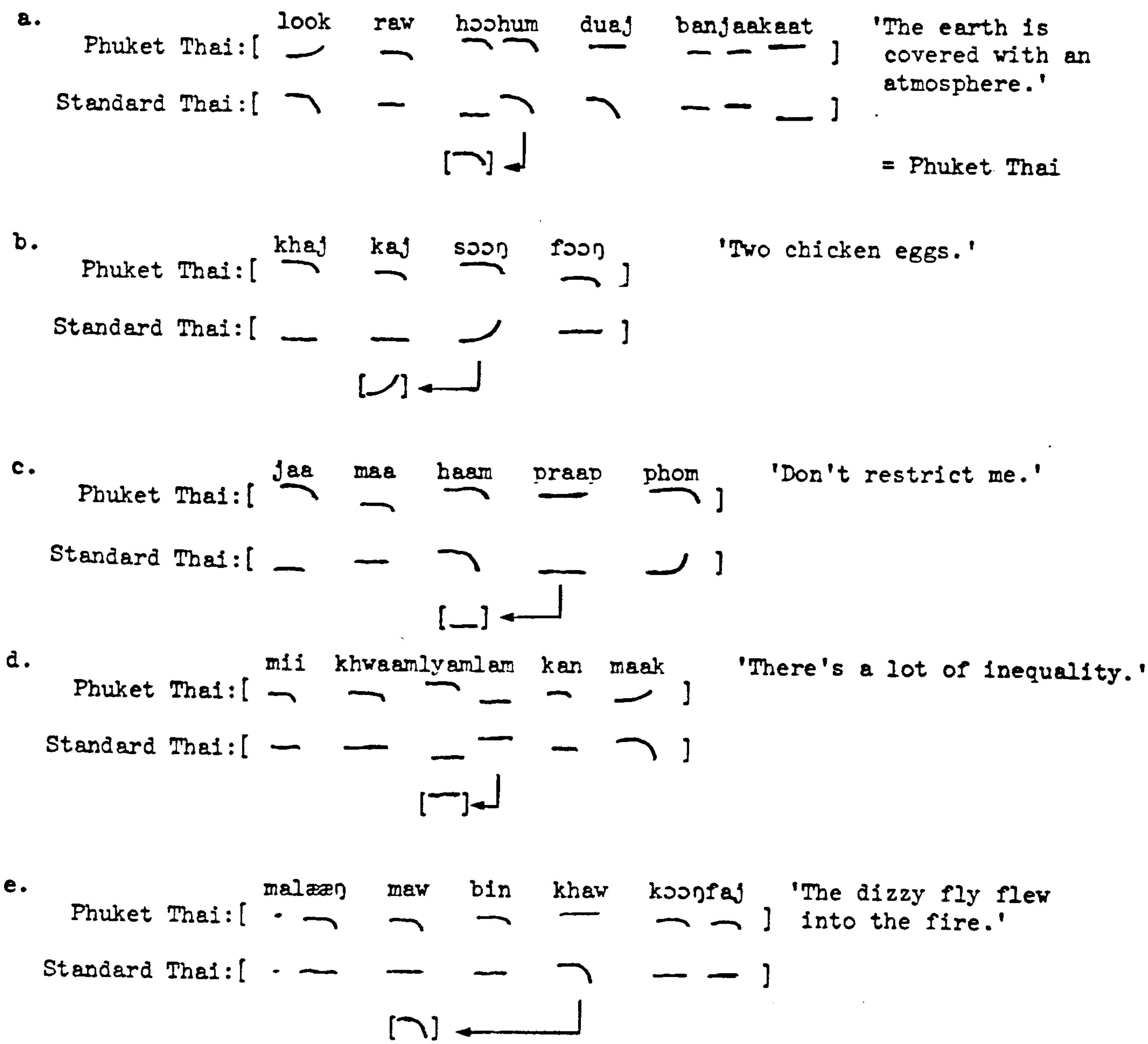
Each example contains the following: a broad phonetic transcription of the tones of the utterance in both Phuket Thai and Standard Thai; the counterfeit tone(s), which appears directly below the intended Standard Thai tone(s), together with an indication of whether the tone in error is identical to the tone for that particular lexical item in Phuket Thai (=Phuket Thai) or not (blank=Standard Thai); an arrow that points from the tone(s) that caused the disordering to the counterfeit tones produced in the error; a broad phonetic transcription of the Standard Thai segments of the utterance (the Phuket Thai consonant and vowel segments are not given where they differ from Standard Thai, since they are not immediately relevant to the issues under discussion); and finally, a rough English translation of the intended Standard Thai utterance.

These perseverative tone errors in 2 are similar to a segmental error like 'Chomsky and Challe' for the intended 'Chomsky and Halle' (Fromkin 1971).

1.2 Errors of Anticipation

Errors of anticipation result in a substitution of one tone in anticipation of an identical or similar tone which occurs later in the utterance, as shown in 3.

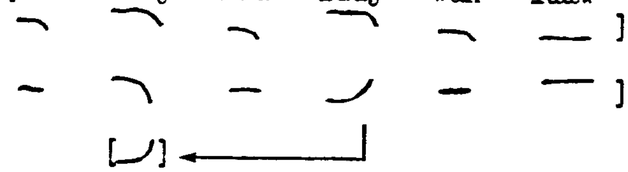
(3)



f. pen khaj maa laaj wan lææw — '(I've) had a fever for several days now.'

Phuket Thai: [ ]

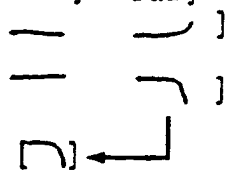
Standard Thai: [ ]



g. thɔɔŋ ruaŋ — '(I've) got diarrhea.'

Phuket Thai: [ ]

Standard Thai: [ ]



3a-d, in that order, show anticipation of a falling, rising, low, and high tone. No examples of anticipation of a mid tone are found in the present corpus. 3e-f show that the substituted tone need not occur on a contiguous syllable. As is true of segmental errors, the fact that tone errors may involve non-contiguous syllables further suggests that speech production is not based on a simple left-to-right Markov processing of successive elements, but instead involves the processing of a larger syntactic/phonological organizational unit prior to the issuing of motor commands. These tone errors then are not errors in motor control but rather errors in the linguistic message. 3g, in comparison with 2c, shows that the same sequence of words may be subject to either an error of anticipation or perseveration.

These anticipatory tone errors are also similar to segmental errors, as in the example 'alsho share' for the intended 'also share' (Fromkin 1971). However, unlike the results of investigations of segmental errors in which the majority were found to be anticipatory (Fromkin 1971:30), in the present corpus of tone errors, perseverative errors outnumber anticipatory errors by a ratio of about 2 to 1.

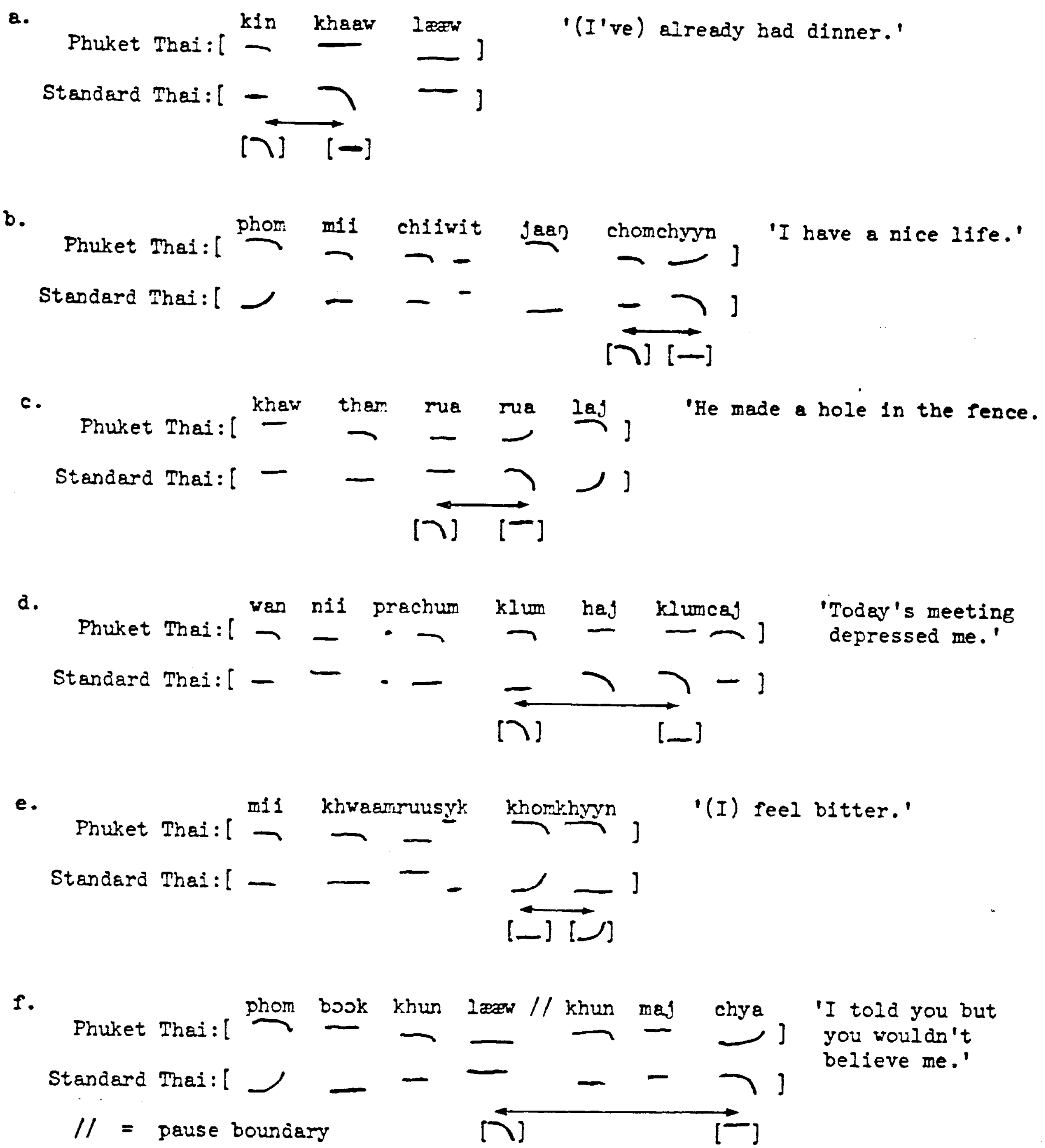
Hyman and Schuh (1974), in their discussion of universals of tone rules based on evidence from tone languages spoken in West African, observe that rules that spread tones from left-to-right (i.e. perseverative tone rules) are quite common in these languages whereas rules that spread tones in the opposite direction (i.e. anticipatory tone rules) rarely occur. They further observe that these perseverative tone rules are most likely to occur when the interval between the two tones is greatest.

The predominance of perseverative tone errors might reflect this tendency for tones to spread from left-to-right, possible reflecting some built-in constraint in the speech production mechanism. The absence of mid tone in perseverative tone errors might reflect the tendency for tones to spread when the interval between the two is greater.

1.3 Errors of Transposition.

Errors of transposition, commonly called 'Spoonerisms' when they involve segments, result when two tones reverse or switch their original position in an utterance. Transposition errors are illustrated in 4.

(4)



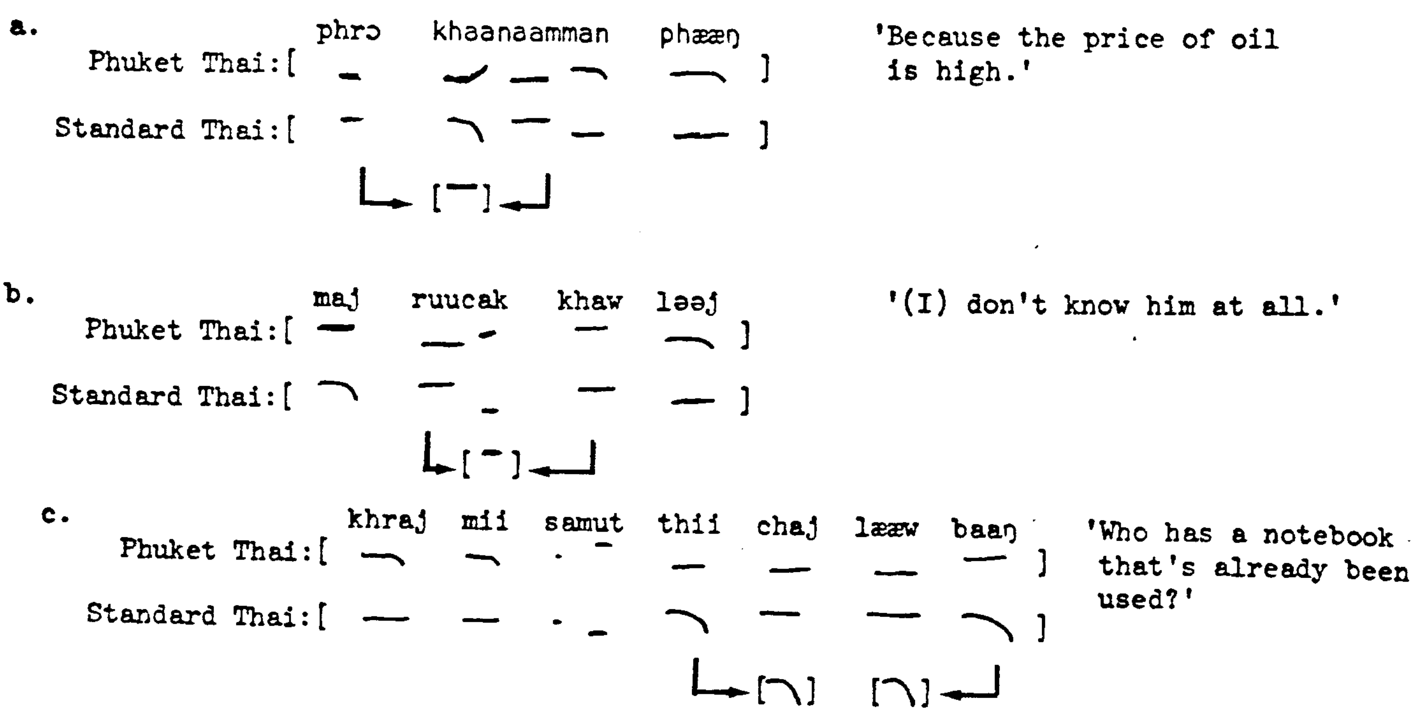
Note, too, that unlike perseverative and anticipatory tone errors, transposition errors may involve a mid tone as evidenced in 4a-b.

These errors are also similar to transposition errors involving segments, as in the example 'teep a cape' for the intended 'keep a tape'(Fromkin 1971).

1.4 Further Similarities Between Tone and Segmental Errors

As is true for consonant and vowel errors, a number of tone errors cannot be unambiguously classified.

(5)



In 5a-b we cannot tell whether the substituted high tone results from anticipation of the following high tone or perseveration of the preceding high tone. In 5c we cannot tell whether the falling tones are errors of anticipation or perseveration, or both. It appears that the probability of a tone error increases when an unlike tone intervenes between a sequence of like tones. This too is found in segmental errors.

One interesting fact about these tonal errors is that unstressed syllables are not normally involved. This is in agreement with the finding of Nooteboom (1969) who points out that 'in significantly more cases than is to be expected in a random distribution the elements involved in a speech error belong to stressed syllables'. MacKay (1969) and Boomer and Laver (1968) also found this to be true for segmental errors. We find here a further similarity between tone and segmental errors.

1.5 Implications for Linguistic Theory

1.5.1 Orthogonal Tone Features

All of the errors that we have discussed so far point to the reality of the independence of whole tones or tone features; that is, the tones may be copied or shifted while the consonants and vowels of the syllable remain in their original position. Fromkin (1971:35-38) employs a similar argument to show the reality of properties or features of consonants and vowels that are smaller than the segment, as in the example 'glear plue sky' for the intended 'clear blue sky' (Fromkin 1973). Note, however, that although tone features function similarly to segmental features in that they can be independently disordered, the evidence from these tone errors argues neither for nor against assigning tones to a suprasegmental as opposed to a segmental unit. Since segmental features are assigned to segmental units, tone features too could be assigned to segmental units on the basis of the tone errors. Clearly, other linguistic evidence must be sought to force a decision on such a theoretical issue.

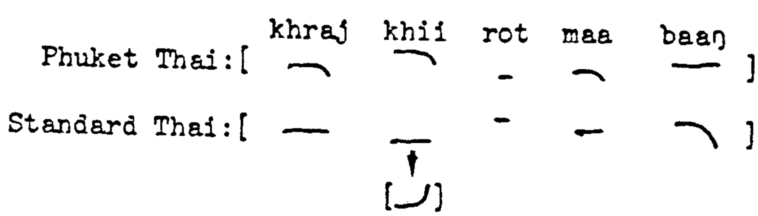
1.5.2 Representation of Contour Tones

As for the phonological representation of contour tones, Wang (1967) proposes a set of distinctive features of tone that includes unit contour tone features (e.g. FALLING, RISING) while Woo (1969), Maddieson (1970), and Leben (1973) among others maintain that all contour tones are to be analyzed as sequences of level tones (e.g. HL = FALLING, LH = RISING). Unfortunately, the data are inconclusive on the question of whether contour tones are to be analyzed as indivisible units or as sequences of level tones. However, the fact that the overwhelming majority of errors reveal that the tones are copied or shifted as unitary wholes, although not evidence against a sequential interpretation of contour tones on a lexical level, certainly necessitates unit contour tone features at some level of representation. Furthermore, most of the substituted tones do not appear to be sensitive to the begin-point and/or end-point of tones that are anticipated or perseverated.

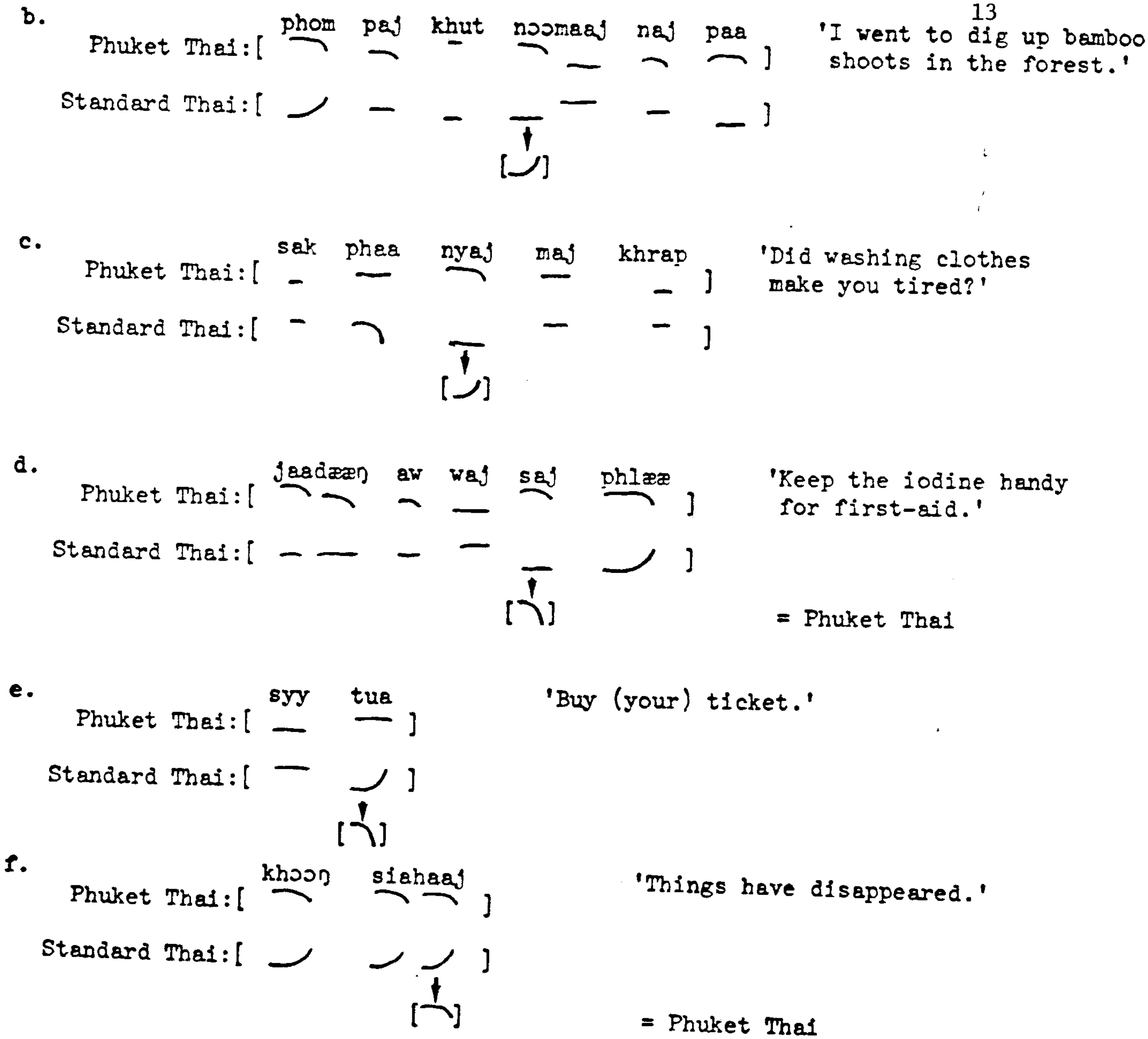
We do find, however, some errors in the present corpus which lead to the conclusion that contour tones be analyzed as sequences of level tones. These errors are given in 6.

(6)

a.



'Who all came by car?'



6a-c suggest that the counterfeit rising tones be analyzed as a sequence of L-H level tones: the rising tones result from the preceding L tone (begin point) anticipating the transition to the following H tone (end point). Similarly, 6d-f suggest that the counterfeit falling tones be analyzed as a sequence of H-L level tones: the falling tones result from the spreading of a preceding H tone (begin point) into a following syllable that carries a L tone (end point). What seems to have happened in these tone errors is shown in 7.

(7) a. L H → $\widehat{\text{LH}}$ H (6a-c)

b. H L → H $\widehat{\text{HL}}$ (6d-f)

The use of unit contour tone features as in 8a-b, would obscure the nature of these tone substitutions.

(8) a. L H → R H

b. H L → H F

It should also be pointed out that a process such as the one depicted in 7a is almost unattested in the tone language of West Africa (Hyman and Schuh 1974). It is possible that the rare occurrence of 7a reflects a typological property of those languages, and not a universal constraint on tone spreading. The process in 7a certainly seems reasonable from a phonetic standpoint. A rising tone facilitates the transition from a lower tone to a higher tone.

The evidence from tone errors in the speech of Phuket bidialectals then suggests that contour tones need to be analyzed as both indivisible units and as sequences of level tones. Segmental errors have been shown to occur at either a more abstract level or at the phonetic level.

The fact that the great majority of contour tone errors involve perseveration, anticipation, or transposition of the whole contour, suggests that even at a more abstract phonemic level contour features may be present.

1.5.3 Tone: Production and Perception

In about fifty percent of the errors the tone substituted for the target or intended tone is a high-falling tone. That is, in half the errors the contaminating tone(s) which perseveres or is anticipated, or transposed is a high-falling tone. This cannot be due to an overwhelming occurrence of high-falling tones in Phuket Thai since the mid-rising falling and low-falling tone also occur widely in the lexicon. The fact that a high-falling tone is substituted for Standard Thai mid and high tones almost as frequently as rising and low tones also makes it difficult to explain in terms of dialect interference alone. It is not the case that the Phuket bidialectal is only substituting a high-falling tone for cognate low and rising tones. However, the distribution of tones in Phuket Thai would partially account for the unevenness in the kinds of tones substituted in these errors since the tones above the dotted line in (1) (see Section 0) occur more frequently than those below.

I would like to suggest that the preponderance of high-falling tones in these errors may be due to language-universal phonetic tendencies instead of language-specific phonological characteristics of the Phuket dialect. First, I know of no tone language that has a phonologically low-falling tone without also having a high-falling tone, whereas numerous Asian tone languages do have only a high-falling tone in their tonal inventory. Li and Thompson (1975), in their investigation of tone acquisition in Mandarin children, found also that the high-falling tone (tone 4), along with the high-level tone (tone 1), was acquired earlier than either the rising (tone 2) or low (tone 3) tones. Ohala and Ewan (1973), based on tone production data, suggest that falling tones are easier to produce than rising tones. Brown (1965, 1975), from evidence on diachronic tone changes among Thai dialects, suggests that a tone gravitates to the most natural height for its contour, tone space permitting: the most natural height for a falling tone is claimed to be mid-high as compared to mid-low for a rising contour tone. These errors thus seem to show how physiological and/ or perceptual factors may influence the kind of errors that occur. Just as in other kinds of speech errors a multiplicity of factors converge. The errors cannot simply be attributed to motor control or performance factors; interaction with other factors must be taken into account in order to explain these errors on the performance level (Fromkin 1975).

Another interesting asymmetry in the distribution of the tones substituted in the errors is that contour tones, ie. falling and rising tones, outnumber level tones by a ratio of 4 to 1. Among the level tones the high tone clearly predominates; few instances of low and mid tones occur in the errors. This distribution in frequency of occurrence of the substituted tones tentatively suggests the rank order in 9.

(9)	'Contour'	1. FALLING
		2. RISING
		3. HIGH
	'Level'	4. LOW
		5. MID

It would appear that errors are more likely to produce dynamic pitch contours than static pitch contours; among the latter, errors are more likely to result in more 'extreme' pitches. Speculatively, this hierarchy might reflect a scale of 'perceptual salience' whereby (comparatively) dynamic tonal contours are perceptually more salient than (comparatively) static tonal contours, and whereby the 'extreme' pitches are perceptually more salient among the static tonal contours. This perceptual bias may intrude in the production of pitch as evidenced by the errors in the present corpus.

2. Dialect Interference

Let me now turn to tone errors that can be directly attributed to dialect interference to see what they tell us about phonological rules.

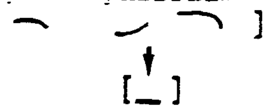
2.1 Phuket Tone Sandhi

Phuket Thai has a tone rule, not found in Standard Thai, by which the rising tone becomes a low-level tone in non phrase-final position at normal speech tempo. The application of this rule is illustrated in 10.

(10)

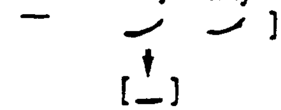
a. pen phiisaaw '(She's) my older sister.'

Phuket Thai:[ ]



b. khaw naŋ niŋ ' He's sitting still.'

Phuket Thai:[ ]



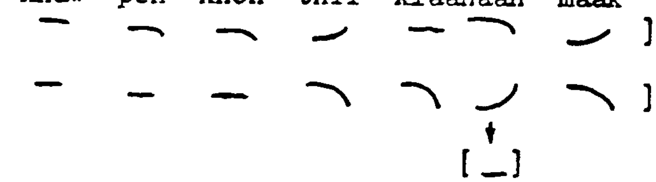
Examples of how the application of this rule may affect a Standard Thai utterance produced by a Phuket bidialectal are given in 11.

(11)

a. khaw pen khon thii klaahaan maak 'He's a very brave person.'

Phuket Thai:[ ]

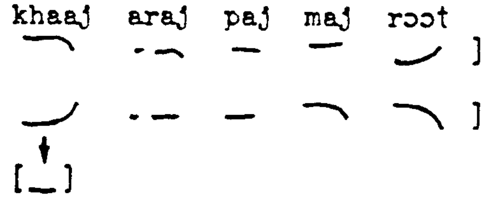
Standard Thai:[ ]



b. khaaj araj paj maj rɔɔt 'I can't seem to sell anything.'

Phuket Thai:[ ]

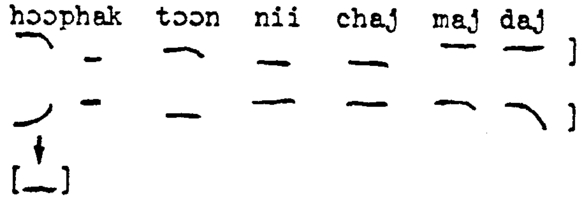
Standard Thai:[ ]



c. hɔɔphak tɔɔn nii chaj maj daj 'These days the dormitory is a mess.'

Phuket Thai:[ ]

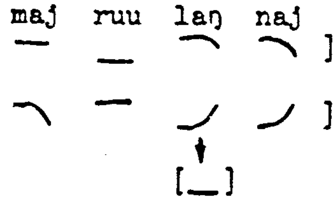
Standard Thai:[ ]



d. maj ruu laŋ naj '(I) don't know which house it is.'

Phuket Thai:[ ]

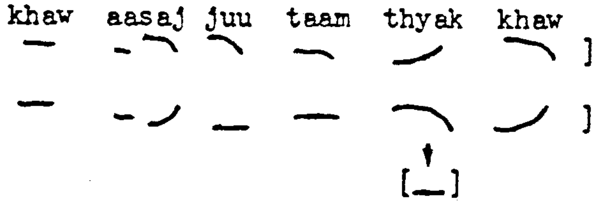
Standard Thai:[ ]



e. khaw aasaj juu taam thyak khaw 'They live in the mountains.'

Phuket Thai:[ ]

Standard Thai:[ ]



= Phuket Thai

Although it is possible that these errors are random substitutions, a more non ad-hoc explanation is that the tone rule is not only part of the Phuket dialect grammar which linguistic evidence necessitates but is actively utilized in performance. This is demonstrated by the fact that the Phuket speaker may incorrectly apply this rule to a Standard Thai underlying form as in 11a-d, and sometimes to a Phuket Thai underlying form as in 11e. Note that in 11a-d the tone rule would not apply to the Phuket Thai underlying forms. In 11c, for example, the underlying form for hɔɔ has a rising tone in Standard Thai, a high-falling tone in Phuket Thai. The tone rule does not apply to the high-falling tone.

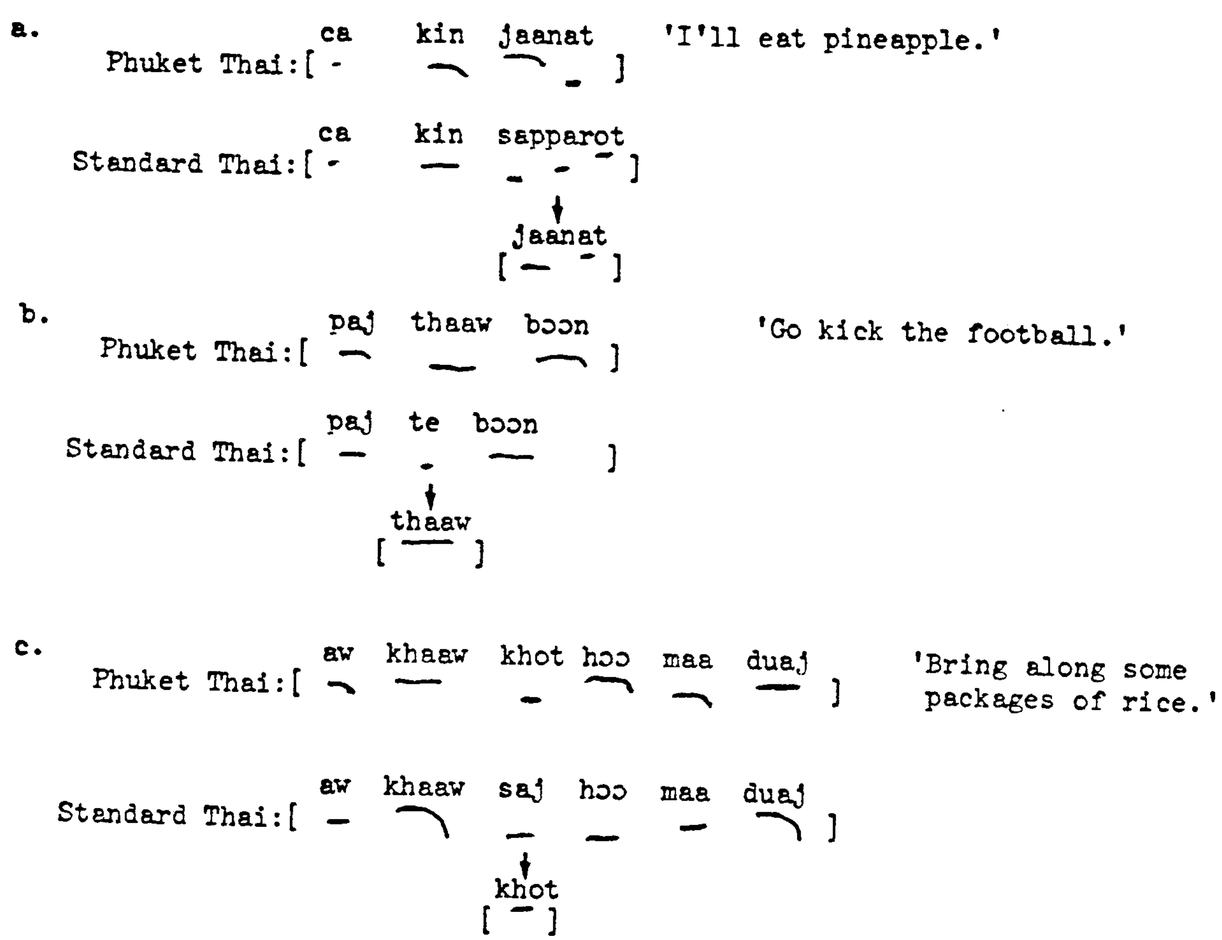
11e further shows that an error may occur in the selection of the underlying cognate form before the application of the tone rule; otherwise, the low-level tone cannot be accounted for. The corresponding Standard Thai underlying form has a high-falling tone and, as we have already seen, the tone rule applies to the rising tone only.

What is important here is that the errors show that the interference is rule-governed. A rule of one grammar is applied incorrectly at a point in another grammar at which it would otherwise apply.

2.2 Non-Cognate Lexical Items

A number of examples can be found in which the Phuket bidialectal substitutes a non-cognate Phuket Thai word in a Standard Thai utterance but pronounces the word with its Standard Thai tone. These are shown in 12.

(12)

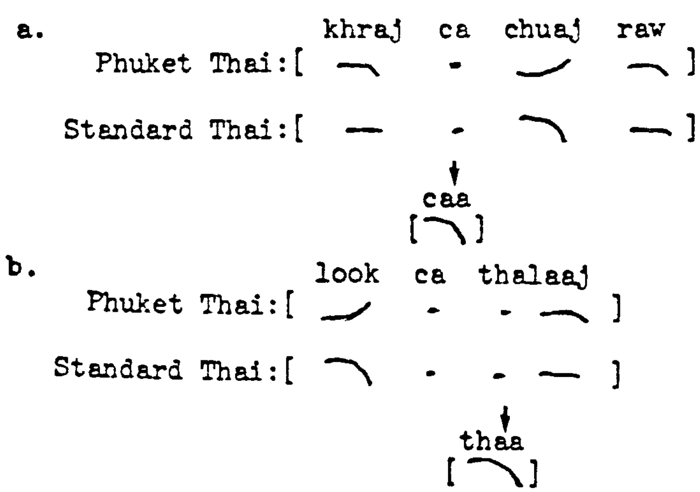


In 12b the word for 'to kick' is te in Standard Thai, thaaw in Phuket Thai. The word thaaw means 'foot' in Standard Thai. Here the speaker substituted the Phuket Thai word but pronounced it with its high tone of the Standard Thai word. These errors clearly result from interference of the Phuket dialect.

2.3 Language-Specific Phonological Constraints

A few errors in the corpus reveal the phonological constraint that disallows contour tones on syllables containing a single sonorant segment.

(13)



a. 'Who will help us?'

b. 'The world is going to crumble.'

In both 13a-b a normally unstressed short vowel is lengthened to accommodate the falling contour tone. 13a shows that ca (future auxiliary) anticipates the falling tone of chuaj; 13b shows that the falling tone of look perseveres into the first syllable of thalaaj, normally pronounced with a short vowel on a mid pitch level. Here we see that unstressed syllables may sometimes be involved in tone errors, as also happens in segmental errors. 13a-b also tell us something about the sequence of events in the production of an utterance. The disordering of the tone must occur before or simultaneous with the lengthening of the vowel. That is, the substitution of the falling tone necessitates a long vowel. If the vowel is not lengthened as a result of the tonal change, a phonological constraint of both Standard Thai and Phuket Thai would be violated. As with segmental errors, tone errors provide evidence for the posited phonological rules of a language.

3. Conclusion

To sum up, the evidence from tone errors made by Phuket bidialectals shows us that (1) the disordering mechanisms that have been proposed to account for consonant and vowel errors may be extended without modification to handle tone errors (2) tones function independently of other segmental features (3) language-specific phonological rules and constraints on the distribution of tones are evidenced in the kinds of errors which occur and (4) tone errors seem to reflect or support some proposed universals regarding tone rules. Tone errors then do provide relevant data for linguistic theory -- to paraphrase Fournié--giving us if not a window, at least a peep-hole, through which we can view the workings of the human mind.

References

Boomer, D.S. and J.D.M. Laver. (1968) 'Slips of the tongue'. *British Journal of Disorders of Communication* 3.1-12.

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

Brown, J.M. (1975) 'The great tone split: did it work in two opposite ways? *Studies in Tai linguistics in honor of William J. Gedney,* ed. by J.G. Harris and J.R. Chamberlain, 33-48. Bangkok: Central Institute of English Language.

Fromkin, V.A. (1971) 'The non-anomalous nature of anomalous utterances.' *Language* 47.27-52.

Fromkin, V.A. (ed.) (1973) *Speech errors as linguistic evidence.* The Hague: Mouton.

Fromkin, V.A. (1975) 'Wheres and whyfores of preach seduction.' Paper presented at the 90th meeting of the Acoustical Society of America, San Francisco.

Hyman, L.M. and R. Schuh. (1974) 'Universals of tone rules: evidence from West Africa.' *Linguistic Inquiry* 5.81-116.

Lashley, K.S. (1951) 'The problem of serial order in behavior.' *Cerebral mechanisms in behavior,* ed. by L.A. Jeffress, 112-136. New York: Wiley.

Leben, W.R. (1973) Suprasegmental phonology. MIT Ph.D. dissertation.

Li, C.N. and S.A. Thompson. (1975) 'Tone acquisition in Mandarin children.' Paper presented at the 8th International Conference on Sino-Tibetan Languages and Linguistics. University of California at Berkeley, also this volume.

MacKay, D.G. (1969) 'Forward and backward masking in motor systems.' *Kybernetik* 6.57-64.

Maddieson, I. (1970) 'The inventory of features required for handling tone in transformational phonology. *Research Notes* (University of Ibadan, Nigeria) 3.2-3:3-18.

Nooteboom, S.G. (1969) 'The tongue slips into patterns.' *Nomen: Leyden Studies in Linguistics and Phonetics,* ed. by A.G. Sciarone et al., 114-132. The Hague: Mouton.

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

Wang, W. S-Y. 'Phonological features of tone.' *International Journal of American Linguistics* 33.93-105.

[Dissertation abstract]

Aspects of Thai Tone

Jackson Thomas Gandour

This dissertation consists of seven sections, each corresponding to a paper that has been published or accepted for publication. These papers investigate a number of hypotheses within the theoretical framework of generative phonology that deal with varous phonological and phonetic aspects of tone. Thai, a member of the Austro-Thai language family is the principal language used for testing the hypotheses. Experimental phonetic data is used to explore both diachronic and synchronic questions concerning tone.

Section 1 explores the relationship between various consonant types and tone in an attempt to provide a phonetic basis for well documented cases of tonal development among languages of Southeast Asia.It is an acoustical investigation of the effects on the fundamental frequency of a vowel caused by preceding consonants of different phonation types. It is found that (1) the fundamental frequency contour in transition to the vowel is relatively high and falling after voiceless consonants, and relatively low and rising (-falling) after voiced consonants (2) the fundamental frequency is initially higher after voiceless unaspirated stops than for voiceless aspirated stops (3) the perturbations on the fundamental frequency caused by preceding consonants are short in comparison to data obtained from languages without lexically contrastive tones and (4) "plain" and "breathy" allophones of the voiceless aspirated stops have a differential effect on the tone of the following vowel. Implications for theories of pitch production and tonal development are discussed.

Section 2 investigates the question of the phonological representation of contour tones, i.e. whether contour tones should be represented lexically as single contour tones or as sequences of level tones. A case of tone sandhi in Lue is found that seems to require not only the contour tone feature RISING, but also the tone feature CONTOUR. Alternative solutions that do not permit underlying contour tone features are shown to obscure a general process of tonal dissimilation; that is, only by postulating classificatory contour tone features can we capture what appears to be the correct linguistic generalization.

Section 3 examines the proposal that predictable segments should always be derived by phonological rule. The glottal stop in Thai, although predictable by rule, suggests that this proposed constraint on a theory of phonology is too strong. The principal evidence in favor of

including the glottal stop in underlying forms involves co-occurrence restrictions on tones and consonants. These generalizations are left unexpressed unless one permits the glottal stop to be present in underlying forms.

Section 4 takes up the issue of whether some classificatory features should be binary or multivalued. In order to provide a satisfactory explanation of the diachronic tone splits conditioned by the phonation types of syllable-initial consonants among languages and dialects of Southeast Asia, it is found necessary to posit a multivalued scalar feature 'glottal width' on the classificatory level of representation. The feature 'glottal width' consists of a linearly ordered set of terms along a single physical continuum that extends from the widest open position of the glottis to the fully closed position (i.e. glottal stop). Other proposed sets of laryngeal features are tested against this evidence and found to be inadequate. A binary feature 'vibrating' is also proposed within this theoretical framework.

Section 5 examines the issue of whether tones are to be represented segmentally or suprasegmentally in Thai. An argument that has been advanced in favor of analyzing tone segmentally in Thai - namely, the neutralization of contour tones in fast speech - is shown to be untenable. A systematic acoustic investigation of these putative cases of tone neutralization shows that RISING and FALLING contour tones are not neutralized to level tones in fast speech. Other arguments, however, can be given in favor of assigning tones to segments on the phonological level of representation as well as analyzing the contour tones as sequences of level tones.

Section 6 consists of a detailed analysis of some tone rules in Thai. It is shown that tonal neutralization on CV syllables - whereby an underlying contrast between HIGH and LOW is neutralized to MID - depends not only on the position of the CV syllable within the words, but also on the presence of an internal word boundary in the underlying forms of certain polysyllabic words. The alternate pronunciations of these polysyllabic words (one alternant exhibits the tonal neutralization whereas the other one does not) is then accounted for simply as a difference in the underlying forms. The presence of an internal word boundary blocks application of the rule that neutralizes the underlying tonal contrast.

Section 7 investigates tone errors produced by Southern Thai bidialectals when attempting to speak Central Thai, the national language of Thailand. It is shown that most of the errors are not due to dialect interference. Instead, the errors reveal that the disordering mechanisms (perseveration, anticipation, transposition) that have been proposed to handle consonant and vowel errors may be extended without modification to tone errors. The tone errors further reflect or support the independence of tone features and some proposed universals regarding tone rules.

<u>Publications corresponding to these sections:</u>

1. "Consonant types and tone in Siamese" *Journal of Phonetics* (1974) 2:337-350.
2. "The features of the larynx: n-ary or binary?" *Phonetica* 32:241-253.
3. "The glottal stop in Siamese: predictability in phonological description" *Phasaa* 4.2:66-78.
4. "On the representation of tone in Siamese" In *Studies in Tai Linguistics*, edited by J. G. Harris and J. R. Chamberlain. Central Institute of English Language: Bangkok, Thailand.
5. "Evidence from Lue for contour tone features" *Phasaa* 5.2:39-52.
6. "Counterfeit tones in the speech of Southern Thai bidialectals" to appear in *Lingua*.

Phonetic explanation of the development of tones from prevocalic consonants

Jean-Marie Hombert

[Department of Linguistics

University of California, Los Angeles]

I. Introduction

The development of contrastive tones on vowels due to the loss of a voicing distinction on obstruents in prevocalic position is probably the most well documented type of tonogenesis. When such a development occurs a relatively lower pitch register develops on vowels following the previously voiced series and a relatively higher pitch is found after the previously voiceless or voiceless aspirated series. This process can lead to a multiplication by two of the number of tones. If the language was atonal it will have two tones after this development; an already existing two-tone system will be transformed into a four-tone system, and so on. The correlation between initial consonant and pitch register was noticed at the beginning of this century by Maspero (1912) and Karlgren (1926) for Chinese, and later extended to other East Asian languages by Haudricourt and Martinet (1946), Haudricourt (1954, 1961), Matisoff (1973a) and Mazaudon (1975). This correlation is also found in other linguistic groups, e.g., in Hottentot (South Africa) as described by Beach (1938). Although it did not give rise to tonal development, a similar correlation between consonant types and fundamental frequency height is found in certain African languages (Hyman 1973 a,b; Hyman and Schuh 1974).

I will assume that when similar sound changes occur in languages genetically, geographically and chronologically distant, these changes should be explained in terms of physiological constraints (articulatory and/or auditory). In order to show that the development of tones due to the loss of a voicing distinction in prevocalic position is phonetically motivated, I will present production and perception data and I will show to what extent they overlap.

II. Production data

a. Previous studies

Phonetic studies by House and Fairbanks (1953), Lehiste and Peterson (1961), Mohr (1968), Lea (1973), and Löfqvist (1975) among others, show how a voicing distinction in pre-vocalic position can affect the fundamental frequency (Fo) of the following vowel. Some of the data from these studies are summarized in Table 1.

	p	t	k	b	d	g
House and Fairbanks (1953)	127.9	127.1	127.2	120.9	120.6	122.8
Lehiste and Peterson (1961)	175	176	176	165	163	163
Mohr (1968)	130.7	129.8	131.1	125.1	124.8	125.0

Table 1. Fundamental frequencies (in Hz) of vowels as a function of the preceding consonant as determined by three studies.

Although the number of subjects and the methods used to measure and average the data differ in these studies, it is clear that the Fo values of vowels are higher after voiceless (aspirated) than after voiced stops and that these values do not vary in any consistent way as a function of the place of articulation of the stops.

Unfortunately, these data give only an averaged or a peak value for Fo, making it impossible to deduce the time course of the Fo perturbation caused by the preceding consonant.

b. Experimental conditions

In order to remedy this, the following data were collected. Five American subjects without speech disorders or history of hearing pathology, speaking some form of general American English dialect were used. They spoke 6 cv nonsense words where C = [p,t,k,b,d,g] (and for three subjects [w,m] as well),and V = $\{i\}$. The word list spoken consisted of ten tokens of each test word arranged in random order. Each test word was uttered in the frame "say ___ again". The recording was done in a sound treated room. Measurements were made on a mini-computer by means of a hardware pitch extractor (Krones, 1968), with a reference point at the onset of the vowel, Fo values were measured at onset and 20, 40, 60, 80 and 100 msec after this onset.

c. Results

The results are given in Figures 1 and 2, Figure 1 showing Fo curves on the vowels following the voiced and voiceless stops, averaged over all speakers' samples, and Figure 2, showing the Fo curves, including those for vowels following sonorants, for three individual speakers.

Although the greatest difference in the Fo curves in figure 1 exists at vowel onset, statistical analysis (analysis of variance followed by Duncan's test) reveals they are still significantly different 100 ms after vowel onset. These two curves differ from each other in two ways: direction of Fo change and average relative value. The data in Figure 2 show that individual speakers' Fo curves exhibit one or both patterns.

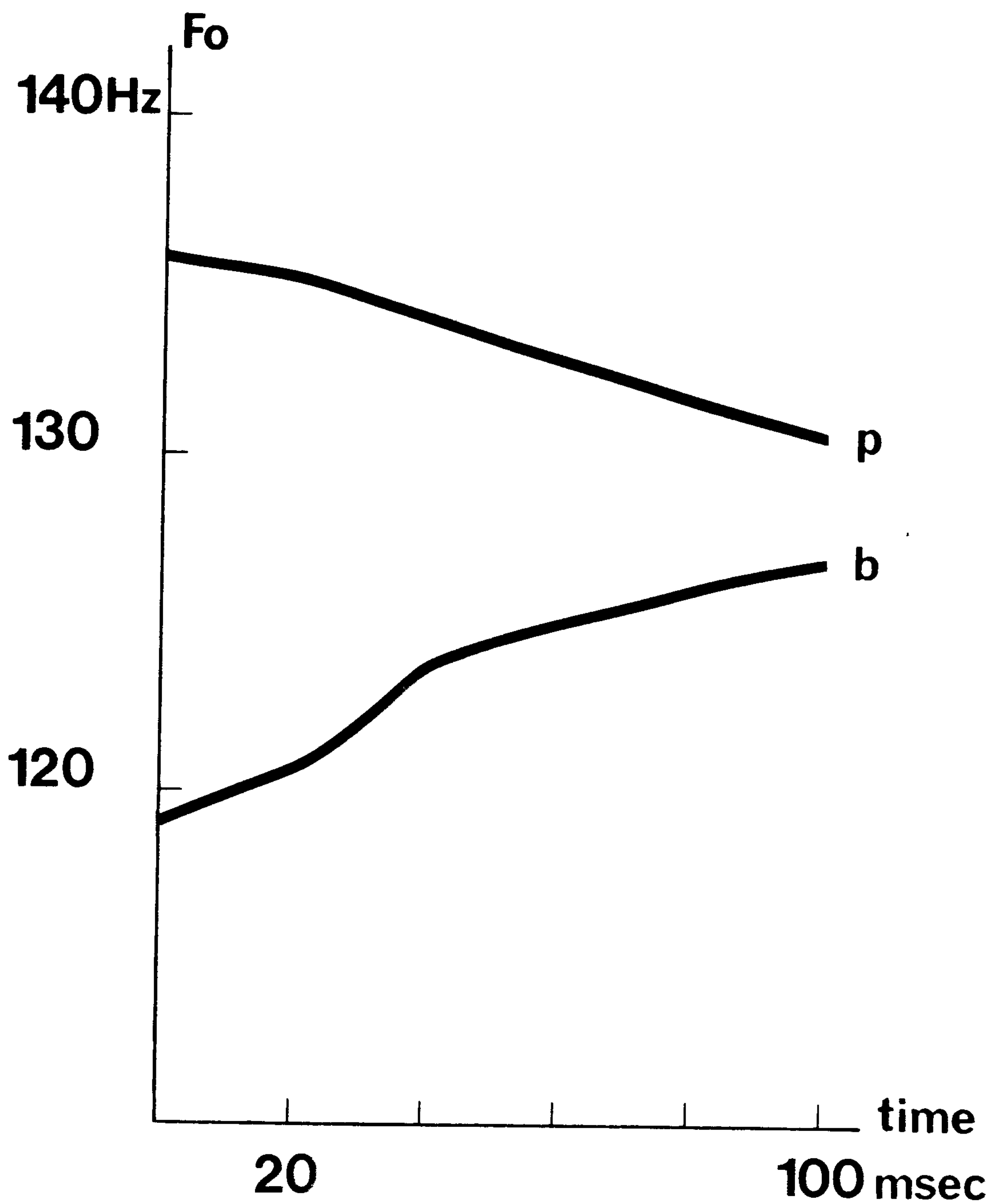


Figure 1. Fundamental Frequency values of vowels after voiced and voiceless aspirated stops - [p] and [b] represent the voiceless aspirated and the voiced series respectively. Fundamental frequency (vertical axis) is measured as a function of time (horizontal axis) -(5 subjects).

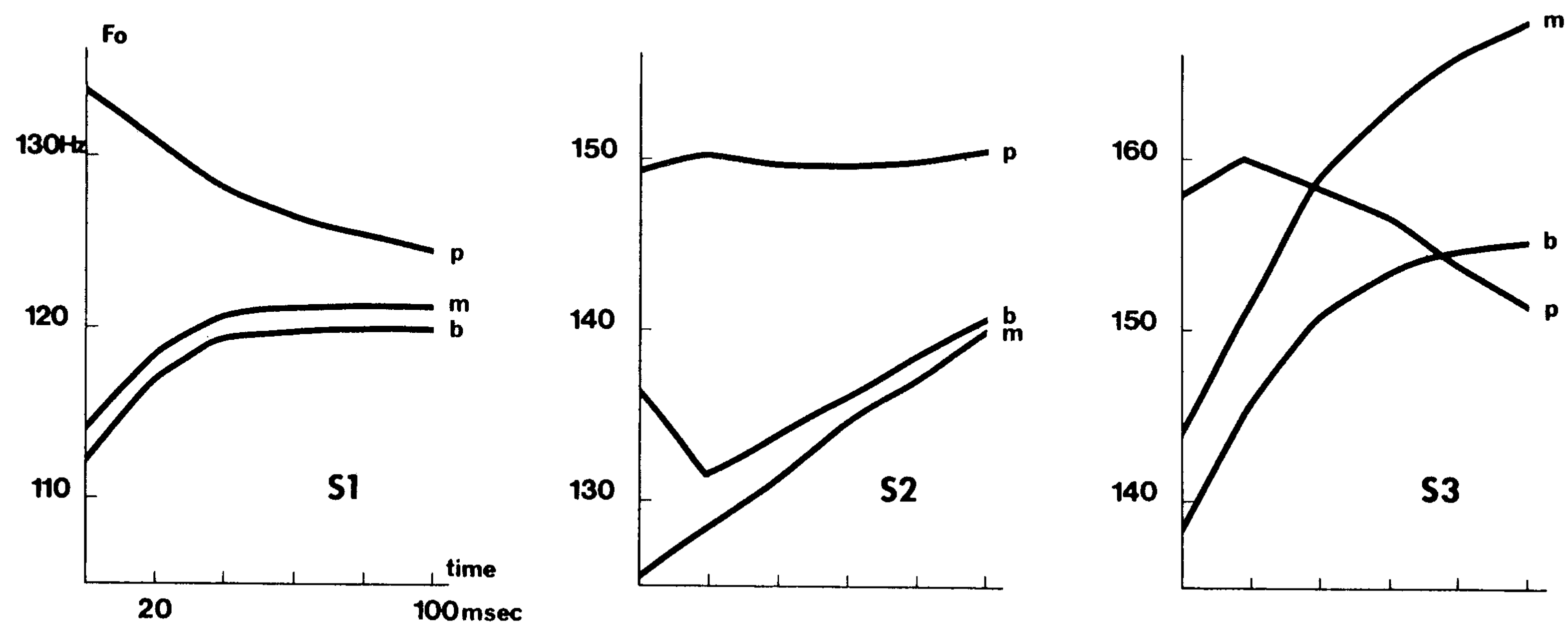


Figure 2. Individual Fo patterns after voiced, voiceless aspirated stops and sonorants (3 subjects).

The explanations proposed to explain these facts (Fo raising after voiceless consonants vs Fo lowering after voiced consonants)can be divided into two categories. The first attributes these Fo perturbations to aerodynamic effects and the second to differences in vocal cord tension.

Researchers following the first theory (i.e. aerodynamic) would explain the phenomenon in the following terms. After the closure of a voiced consonant, voicing continues, but since the oral pressure increases (because of the closure), the pressure drop decreases, leading to a lower frequency. The fundamental frequency then rises after the release until it reaches the "normal" value of the vowel which is being realized. In the case of a voiceless consonant, since the rate of air flow is supposed to be high, a strong Bernoulli effect will draw the vocal folds together very rapidly; they will be pushed apart very rapidly as well because the subglottal pressure is still high. Consequently, the rate of vibration of the vocal folds will be high at the onset of the vowel and will return gradually to the intrinsic value of the vowel being realized.

The experimental data presented earlier as well as earlier studies (Löfqvist 1975), show that a consonant still affects the fundamental frequency of the following vowel at least 100 msec after vowel onset. Proponents of the second theory (vocal cord tension) claim that this perturbatory effect is too long to be attributed to aerodynamic factors. Halle and Stevens (1971) suggest that these intrinsic variations are the result of horizontal vocal cord tension and they propose the features [stiff] and [slack] vocal cords to capture the relationship between low tone and voiced consonants (where the vocal cords are supposed to be slack in order to facilitate voicing) on the one hand, and high tone and voiceless consonants on the other hand. Studies by Ohala (1972), Ewan and Krones (1974), as well as more recent work in progress by Ewan suggest that the Fo perturbation is caused at least partially by vertical tension (i.e. larynx height).

Both of these explanations (horizontal and vertical tension) fail to account for the fact that postvocalic consonants do not have the same effect on F_0 as prevocalic consonants do (however see below). Lea (1972, 1973) suggests that both voiced and voiceless consonants lower the Fo of the preceding vowel. Other studies (Mohr 1971, Slis 1966) indicate that postvocalic consonants have a similar effect on Fo as prevocalic consonants but with a much smaller magnitude. The counterargument I just presented based on different influences of pre- and post-vocalic consonants can be weakened if one considers that postvocalic consonants are less "strongly" articulated than their prevocalic counterparts (Slis 1967, Fromkin 1969).

Nevertheless, Halle and Stevens' position is not supported by experimental data; electromyographic recordings by Hirose, Lisker and Abramson (1973) and Hirose and Gay (1972)do not show obvious differences in the tension of the laryngeal muscles during the production of voiced/voiceless distinctions. Ewan and Krones' claim however, is in agreement with experimental data showing a correlation between Fo and larynx height (Ohala 1972, Ohala and Ewan 1972). Ewan and Krones (1974) also show a correlation between voiced sounds and low larynx position as opposed to voiceless sounds and high larynx position. It was also indicated that the larynx was in lower position at the end than at the beginning of a voiced consonants. This suggests that the larynx is actively lowered during a voiced consonant in order to increase the volume of the oral cavity. Warren and Hall (1973) and Bell-Berti (1975) show that this is, at least partially, an active process. If this is the case one would expect to find a perturbed (lowered) Fo after voiced consonants as opposed to a nonperturbed Fo after sonorants and voiceless consonants. Unfortunately this does not seem to be the case in the data presented in Figure 2, where it is shown that sonorants pattern similarly to voiced obstruents. Although it seems that theories based on muscular tension cannot account for some empirical data, we are in an even more difficult situation with theories based only on aerodynamic factors.

Klatt, Stevens and Mead (1968) present air flow data in which a high rate of air flow lasts only about 50-60 msec into the vowel; comparable but uncalibrated data are presented by Frøkjær-Jensen, Ludvigsen and Rischel (1971). Moreover, van Hattum and Worth (1967) as well as Isshiki and Ringel (1964) show that oral airflow is momentarily lower after voiceless aspirated consonants than it is after voiced consonants. These data are in agreement with the results of a mathematical model of aerodynamics proposed by Ohala (1975a). All these results seem to favor the theories based on muscular tension.

III. Perceptual data

a. Previous studies

There is evidence from experiments using synthesized speech that small fundamental frequency perturbations can be used as cues to discriminate between sonorants and voiced obstruents and between voiced and voiceless obstruents (Chistovitch 1969, Haggard et al 1970, Fujimura 1971, Abramson 1974). The perception of stimuli with changing frequency contours has been investigated for pure tones (Brady et al 1961; Heinz et al 1968, Nabelek and Hirsh 1969; Nabelek et al 1970; Pollack 1968; Sergeant and Harris 1962; Tsumura 1973) as well as for vowels (Klatt 1973; Rossi 1971). From these studies it is difficult to conclude to what extent the perception of a changing frequency contour would be affected by a steady-state frequency immediately following the contour.

b. Experimental conditions

In order to get these data, the following study was carried out.[1] Ten subjects, native speakers of American English, with normal hearing, participated. Acoustic stimuli consisting of 10 instances of the vowel [i] were synthesized with different fundamental frequency patterns.

As shown in Figure 3, each stimulus was composed of a slope followed by a level tone maintained constant at 120 Hz. The onset frequency was either 110 or 130 Hz (i.e. Fo = ±10 Hz). The duration of the slope was varied at 40, 60, 100, 150 and 250 ms. In other words, 5 stimuli (with Fo onset = 130 Hz) had a falling fundamental frequency and 5 stimuli (with Fo onset = 110 Hz) had a rising fundamental frequency. The overall duration of each stimulus was fixed at 250 ms. Each time a stimulus was presented it was followed by a 500 ms pause and a second vowel [i] with a steady-state fundamental frequency. The duration of this vowel was also 250 ms. The level of its fundamental frequency was adjustable by a knob controlled by the subject. The task was to match the pitch of the second vowel to the pitch of the beginning of the first vowel. The rate of stimulus presentation as well as the number of trials for a given presentation were controlled by the subject. Each one of the 10 stimuli was presented 3 times in a randomized order. The subjects heard the stimuli through earphones at a comfortable level (about 70 dB). The parameter values were chosen in order to simulate the effects of consonants on neighboring vowels.

c. Results

The results[2] are presented in Figure 4, subjects' responses are plotted as a function of the duration of the slope. Responses to stimuli with a falling Fo at the onset (from 130 Hz to 120 Hz) are indicated by a circle ("O"), responses to stimuli with a rising Fo (from 110 to 120 Hz) are indicated by a cross ("X"). A statistical analysis of these data (analysis of variance followed by Duncan's test) indicates that the two curves are already perceived as significantly different when the onset slope (from Fo onset to level Fo) is 60 ms long.[3]

This graph suggests that 1) falling patterns (i.e. vowels with fundamental frequency onset above 120 Hz) are perceived more accurately than rising patterns (i.e. vowels with fundamental frequency onset below 120 Hz; 2) the longer the slope, the more accurate the matching, but correlation between slope duration and accuracy of matching is not linear.

d. Discussion

These results can be explained by forward masking. If we extrapolate the results obtained with steady state tones to contours, (i.e. masking a higher frequency by relatively lower frequency) we can understand why the onset region of the rising ramp was not accurately perceived, since each frequency was masked by the previous lower frequency (since the frequency is going up). This is not the case for the falling tone in which each frequency is followed by a lower frequency. Data from the psychoacoustic literature (Brady et al 1961, Heinz et al 1968, Pollack 1968, Nabelek and Hirsh 1969, Nabelek et al 1970, Tsumura 1973) can be interpreted as supporting my claim concerning the role of masking in the perception of changing frequency contours.

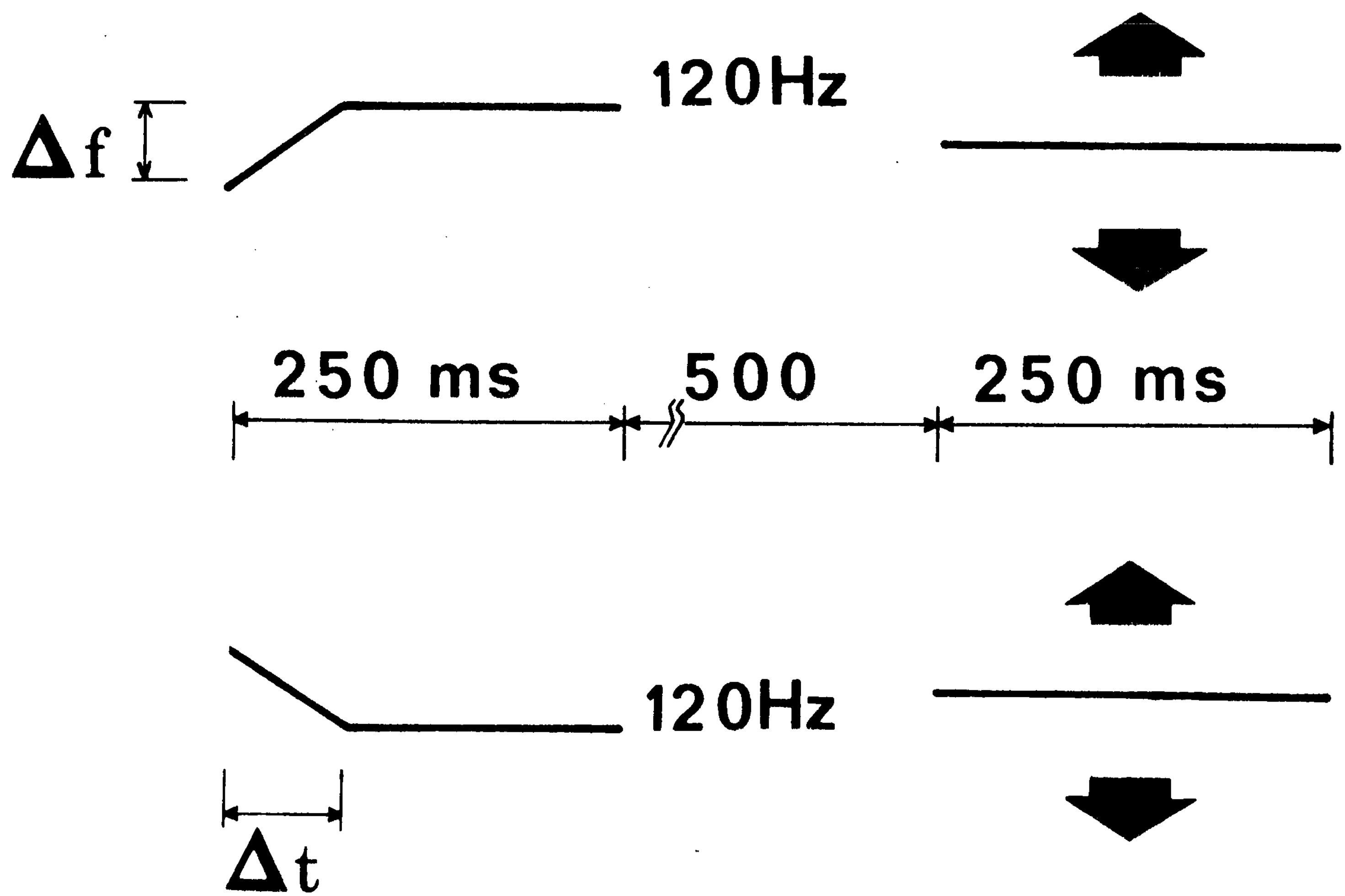


Figure 3. Stimulus presentation.

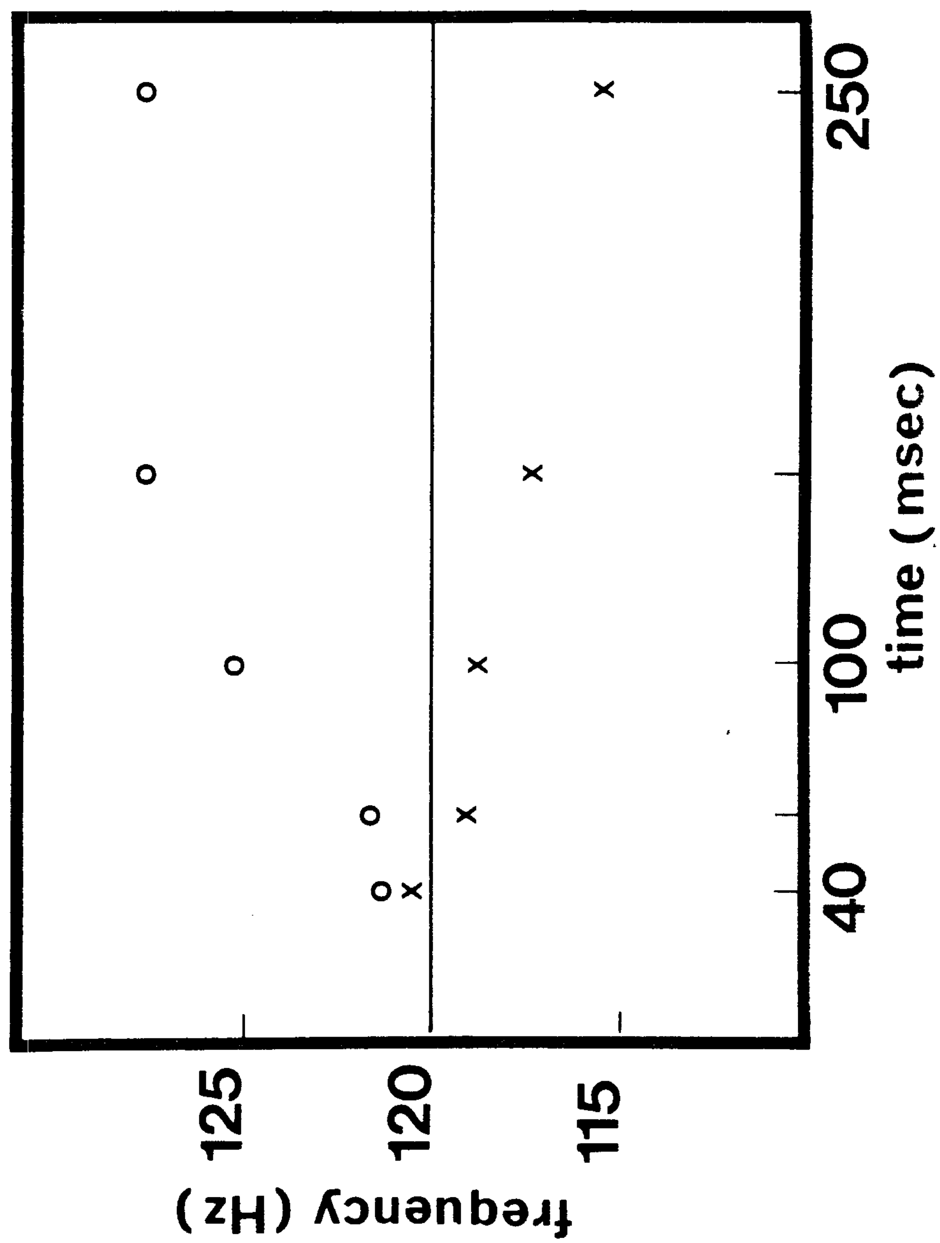


Figure 4. Perceived fundamental frequency (vertical axis) as a function of the duration of the slope (horizontal axis) - (O) indicates responses to falling Fo stimuli (Fo onset = 130 Hz). (X) indicates responses to stimuli with risingFo (Fo onset = 110 Hz).

In these experiments, subjects were asked to match the pitch of a steady state signal with a changing frequency signal. They consistently adjusted their steady state tone closer to the final point of the contour. This fact already shows the role of masking which attenuates the effect of the onset region in favor of the offset region, but furthermore there is a tendency to match closer to the final point when the stimulus is a rising contour (as opposed to a falling contour). This indicates, as I have suggested, that the masking of the onset is more effective in the case of rising contours (as opposed to falling contours), and consequently, this leads to the perception of an averaged pitch closer to the offset frequency.

These data are also in agreement with the study of Brady et al (1961) with respect to the role of the rate of frequency change. They found that the matching of a steady state frequency with a contour frequency is closer to the end point of the contour when the rate of change is high; in other words, the onset region is less salient at high glide rates. This is shown on Figure 4 by responses close to 120 Hz when the slope duration of the stimulus is short (i.e. the rate of frequency change is high). This is also in agreement with Pollack (1968) and Nabelek and Hirsh (1969), whose results indicate that optimum discriminability of relatively small frequency changes is obtained at relatively slow glide rates.

Finally, it should be pointed out that this study was limited to the comparison of Fo differences vowels following voiced and voiceless obstruents. Further research should investigate the perceptual role of Fo during voiced consonant immediately preceding vowels vs.the absence of Fo during voiceless consonants.

IV. Conclusion

In the first part of this paper, I showed that the consonantally-induced Fo perturbations on vowels in such non-tone languages as English and Swedish persist for some 100 ms after vowel onset. The perceptual data just presented show that listeners start hearing significant differences in the Fo onset of our synthesized stimuli when the slope of the Fo contour is 60 ms long. Thus, there is at least 40 ms between the time we start hearing the differences and the time the real consonant-related Fo variations cease to be significantly different.

These data, then, allow us to define the narrow limits, both perceptual and articulatory, within which the development of tones from a former voiced/voiceless stop contrast is likely to occur.

Acknowledgments

I wish to thank John Ohala, Bob Krones, Hector Javkin for their useful advice throughout this study. My thanks also to the members of the UCLA tone seminar for their comments on an earlier version of this paper. This paper is a revised version of a paper presented under the title "Production and Perception: Their respective roles in one case of tonal development" at the 49th Annual meeting of the Linguistic Society of America, Dec. 27-30, 1974, New York.

Footnotes

[1]For a more complete description of this experiment, see Hombert 1975b.

[2]The fact that a linear scale is used does not distort the results within this narrow portion of the frequency range (see Stevens, Volkman, and Newman 1957). In any case my claim about the asymmetry between rising and falling contours would have been even more obvious with a logarithmic scale.

[3]In reporting the results of the perceptual experiments we have given those values of ΔFo and Δt of the pitch ramps which were detectably different *on the average*. It is an open question whether for the purpose of explaining how these differences could be detected in order to give rise to a sound change we should report instead those minimum differences detected by the best, not the average listener. If so, the values given above should be regarded as very conservative estimates of minimal detectable pitch perturbations (Hombert, in preparation).

References

Abramson, A.S. (1974) 'Pitch in the Perception of voicing states in Thai: diachronic implications.' Paper delivered at the Annual Meeting of the Linguistic Society of America, New York City.

Atkinson, J.E. (1973) Aspects of intonation in speech: implications from an experimental study of fundamental frequency. University of Connecticut doctoral dissertation.

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

Bell-Berti, F. (1975) Control of pharyngeal cavity size for English voiced and voiceless stops. *Journ. of the Acoust. Soc. of America* 57:456-461.

Brady, P.T., A.S. House and K.N. Stevens (1961) Perception of sounds characterized by a rapidly changing resonant frequency. *Journ. of the Acoust. Soc. of America* 33:1357-1362.

Broad, D.J. (1968) Some physiological parameters for prosodic description. Speech Communications Research Laboratory, Inc. *SCRL Monograph 3.* Santa Barbara, Calif.

Chistovitch, L.A. (1969) Variation of the fundamental voice pitch as a discriminatory cue for consonants. *Soviet Physics-Acoustics, 14* (#3): 372-8.

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

Ewan, W.G. (1975) Laryngeal behavior in speech. To be published in the *Proceedings of the 8th International Congress of Phonetic Sci.* Leeds, August 1975.

Ewan, W.G. (forthcoming) Laryngeal behavior in speech. Univ. of Calif. Berkeley, doctoral dissertation.

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

Fromkin, V.A. (1966) Neuro-muscular specification of linguistic units. *Language and Speech,* 9.3:170-199.

Frøkjær-Jensen, B., C. Ludvigsen and J. Rischel (1971) A glottographic study of some Danish consonants, in *Form and substance- Phonetic and linguistic papers presented to Eli Fischer-Jørgensen.* Akademisk Forlag.

Fujumura, O. (1971) Remarks on stop consonants: synthesis experiments and acoustic cues, in *Form and Substance,* Akademisk Forlag, pp. 221-232.

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

Haggard, M., Ambler, S. and Callow, M. (1970) Pitch as a voicing cue. *Journ. of the Acoust. Soc. of America,* 47:613-617.

Halle, M. and Stevens, K.N. (1971) A note on laryngeal features. *Quart. Progress Reports,* Res. Lab of Electronics, MIT, 101:198-213.

Hanson, R.J. (1975) Fundamental frequency dynamics in VCV sequences. To be published in the *Proceedings of the 8th International Congress of Phonetic Sciences,* Leeds, August 1975.

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

Haudricourt, A.G. (1961) Bipartition et tripartition des systémes de tons dans quelques languages d'extreme orient. *Bulletin de la Société Linguistique de Paris*, 56:163-180.

Haudricourt, A.G. (1971) On tones in Punjabi. *Pakha Sanjam*, 4:1-3.

Haudricourt, A.G. (1972) L'apparition des registres des langues à tons ponctuels. *Proceedings of the 7th International Congress of Phonetic Sciences*, Montreal, (Mouton), pp. 895-896.

Haudricourt, A.G. and Martinet, A. (1946) Propagation phonétique ou evolution phonologique? Assourdissesment et sonorisation d'occlusives dans l'Asie du Sud. Est. *Bulletin de la Société Linguistique de Paris*, 43:82-92.

Heinz, J.M., B. Lindblom, J. Lindqvist (1968) Patterns of Residual Masking for sounds with speech like characteristics. *IEE Transactions on Audio and Electroacoustics* Vol. AU-16, 1, 107-111.

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

Hirose, H., Lisker, L. and Abramson, A. (1973) Physiological aspects of certain laryngeal features in stop production. (Abstract) *Journ. Acoust. Society of America*, 53:294-295. And *Status Report on Speech Research* SR 31/32 - (1972) Haskins Labs.

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

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

Hombert, J-M. (1975b) Perception of contour tones: an experimental investigation. *Proceedings of the 1st Annual Meeting of the Berkeley Linguistics Society* 221-232.

Hombert, J-M. (1975c) Consonant types, vowel height and tone in Yoruba. Paper presented at the 6th Annual Conference on African Ling. Ohio St. U., April 1975). (Also, in this volume).

Hombert, J-M. (1975d) Development of tone from segmentals: evidence from contour tone perception. To be published in the *Proceedings of the 8th International Congress of Phonetic Sciences*, Leeds, August 1975.

Hombert, J-M. (1975e) Development of tones from vowel height. Paper presented at the Annual Meeting of the Linguistic Society of America, San Francisco, 1975. (In this volume).

Hombert, J-M. (in preparation) The mechanism of phonetically motivated sound changes.

Hombert, J-M. and Greenberg, S. (1975) Contextual factors influencing tone discrimination. Paper presented at the 90th ASA meeting, San Francisco, Nov. 1975. (In this volume).

Hombert, J-M. and Ladefoged, P. (1976) The effect of aspiration on the fundamental frequency of the following vowel. Paper presented at the Spring meeting of the ASA , Washington, D.C.

House, A.S. and Fairbanks, G. (1953) The influence of consonant environment upon the secondary acoustical characters of vowels. *Journ. of the Acoust. Soc. of America*, 25,1:105-13.

Hyman, L.M. (ed) (1973a) Consonant types and tone. Southern Calif. *Occasional Papers in Ling.* 1, Univ. of Southern Calif.

Hyman, L.M. (ed) (1973b) The role of consonant types in natural tonal assimilations. In *Consonants Types and Tone.* Southern Calif. *Occasional Papers in Linguistics*, 1, Univ. of Southern Calif.

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

Isshiki, N., R. Ringel (1964) Air flow during the production of selected consonants. *Journ. of the Speech and Hearing Res.* 7,3:233-244.

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

Karlgren, B. (1926) Etudes sur la phonologie Chinoise. *Archives d'Etudes Orientales*, 15.

Klatt, D.H. (1973) Discrimination of fundamental frequency contours on synthetic speech: Implications for models of speech perception. *Journ. of the Acoust. Soc. of Amer.* 53, pp. 8-16.

Klatt, D.H., K.N. Stevens and J. Mead (1968) Studies of Articulatory Activity and Airflow During Speech in Sound Production in Man. A. Bouhuys (ed.) 42-55. Annals of the New York Academy of Sciences. New York.

Krones, R. (1968) Pitch Meters. *Monthly Internal Memorandum.* UC Berkeley. May and July.

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

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

Lea, W. (1972) Intonational cues to the constituent structure and phonemics of spoken English. Purdue Univ., doctoral thesis.

Lea, W. (1973). Segmental and suprasegmental influences on fundamental frequency co tours. In: *Consonant Types and Tone* (Ed. L. Hyman) pp. 15-70.

Lehiste, I. and Peterson, G. (1961). Some basic considerations in the analysis of intonation. *JASA* 33:419-425.

Lindblom, B. E. F. and Sundberg, J. E. F. (1971) Acoustical consequences of lip, tongue, jaw, and larynx movement. *JASA* 50:1166-1179.

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

Maddieson, I. (1974). A note on tone. *UCLA Working Papers in Phonetics* 27:18-27.

Maspero, H. (1912) Etudes sur la phonetique historique de la langue annamite: les initiales. *BEFEO* 12.

Matisoff, J. (1970) Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study. *JAOS* 90:13-44 (#1).

Matisoff, J. (1972) The Loloish tonal split revisited. Research Monograph No. 7, Center for South and Southeast Asia Studies, U. Calif. Berkeley.

Matisoff, J. (1973) Tonogenesis in Southeast Asia. In: *Consonant Types and Tone* (Ed. L. Hyman), pp. 71-96.

Mazaudon, M. (1975) Tibeto-Burman tonogenetics. To appear in *Current Trends in Sinotibetan Linguistics* (Ed. L. R. Maran).

Mohr, B. (1968) Intrinsic fundamental frequency variation: II. Monthly Internal Memorandum, Phonology Lab., Univ. of Calif., Berkeley, June, pp. 23-32.

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

Møller, A. R. (1973) Coding of amplitude and frequency modulated sounds in cochlear nucleus. Paper presented at the Symposium on Auditory Analysis and Perception of Speech, Leningrad, 21-24 August, 1973.

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

Nabelek, I. V., A. K. Nabelek and I. J. Hirsh (1970) Pitch of tone bursts of changing frequency. *JASA* 48:536-553.

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

Ohala, J. (1972) How is pitch lowered? (Abstract) *JASA* 52:124.

Ohala, J. (1973) The physiology of tone. In: *Consonant Types and Tone* (Ed. L. Hyman), pp. 1-14.

Ohala, J. (1974a) Phonetic explanations in phonology. In: *Papers from the Parasession on Natural Phonology* (Eds. A. Bruck, R. A. Fox, and M. W. LaGaly), Chicago Linguistic Society, pp. 251-274.

Ohala, J. (1974b) Experimental historical phonology. In: *Historical Linguistics II. Theory and Description in Phonology.* (Proceedings of the First International Conference on Historical Linguistics, Edinburgh), (Eds. J. M. Anderson and C. Jones), Amsterdam: North Holland Publ. Co., pp. 353-389.

Ohala, J. (1975a) A model of speech aerodynamics. To appear in *Proceedings of the 8th International Congress of Phonetic Sciences*, Leeds.

Ohala, J. (1975b) Phonetic explanations for nasal sound patterns. In: *Nasalfest; Papers from a Symposium on Nasals and Nasalization* (Eds. L. A. Ferguson, L. M. Hyman and J. J. Ohala), Language Universals Project, Stanford University.

Ohala, J. and Ewan, W. (1973) Speed of pitch change. (Abstract) *JASA* 53:345.

Peterson, G. E. and Barney, H. L. (1952) Control methods used in a study of the vowels. *JASA* 24:175-184.

Pollack, I. (1968) Detection of rate of change of auditory frequency. *J. Exp. Psychol.* 77:535-541.

Pollack, I. and Ficks, L. (1954) Information of elementary multidimensional auditory displays. *JASA* 26:155-158.

Rossi, M. (1971) Le seuil de glissando ou seuil de perception des variations tonales pour les sons de la parole. *Phonetica* 23:1-33.

Sergeant, R. L. and Harris, J. D. (1962) Sensitivity to unidirectional frequency modulation. *JASA* 34:1625-1628.

Slis, I. H. (1966) A model for the distinction between voiceless and voiced consonants. *IPO Annual Progress Report* 1:40-44.

Stevens, K. N. (1975) Modes of conversion of airflow to sound and their utilization in speech. To appear in *Proceedings of the 8th International Congress of Phonetic Sciences*, Leeds.

Stevens, S. S. and Volkmann, J. (1940) The relation of pitch to frequency: a revised scale. *Am. J. of Psychology* 53:329-53.

Sundberg, J. (1973) Data on maximum speed of pitch changes. *Speech Trans. Lab., Quarterly Progress and Status Reports* (Stockholm) 4:39-47.

Tsumura, T., T. Sone and T. Nimura (1973) Auditory detection of frequency transition. *JASA* 53,1:17-25.

Van Hattum, R. J. and Worth, J. H. (1967) Air flow rates in normal speakers. *The Cleft Palate Journal* 4,2:137-147.

Warren, D. W. and Hall, D. J. (1973) Glottal activity and intraoral pressure during stop consonant productions. *Folia Phoniatrica* 25:121-129.

Whitfield, I. G. and Evans, E. F. (1965) Responses of auditory cortical neurons to stimuli of changing frequency. *J. Neurophysiology* 28:655-72.

Consonant Types, Vowel Height and Tone in Yoruba

Jean-Marie Hombert

Department of Linguistics

University of California, Los Angeles

1. Introduction

In this paper I shall address myself to three different issues: consonant types, vowel height, and distribution of tones. In the first part, I want to show how a voiced vs. voiceless consonant can affect the fundamental frequency of the following vowel depending on the tone of this vowel. These data are needed in order to be able to show that F_0 perturbation caused by voiced/voiceless consonants in prevocalic position are important enough to allow a reinterpretation into a tone distinction if a voicing contrast disappears. In the second part, the intrinsic pitch of vowels as a function of tone will be investigated; the relevance of these data for rejecting or accepting current theories of intrinsic pitch of vowels will be examined. I will also try to show why the development of tones from the reinterpretation of intrinsic pitch of vowels is so rarely attested in languages. Finally, I will indicate that it is often the case that other cues than steady state fundamental frequencies are used to identify tones. The so-called Yoruba low level tone which is realized as a low falling tone, at least in some environments, is a good example of this.

2. Consonant types

In an earlier study (Hombert 1976a) I asked five American speakers to read a word list consisting of ten tokens of each test word arranged in random order. These test words were 6 CV nonsense words using the consonants /p, t, k, b, d, g/. Only the vowel [i] was used since I was not interested in variations caused by vowel height in this experiment. With a reference point at the onset of the vowel, F_0 values were measured at the onset and 20, 40, 60, 80 and 100 msec after this onset. Under these conditions, data presented in Figure 1 were obtained. As to be expected, the largest differences between the two curves, obtained respectively after voiceless and voiced consonants, were found at the vowel onset, with these differences decreasing as time increases. A statistical analysis[1] of these data showed that these two curves were still significantly different[2] at 100 msec after vowel onset.

In the second part of this study using synthesized stimuli I showed that these intrinsic perturbations can be perceived even when they are only 60 msec long. In relating these perceptual data and the production data to each other, we can see that there is an overlap of at least 40 msec between the time we start hearing significant differences and the time these differences cease to be significant.

These experimental data validate and explain the well-attested development of tone due to the loss of some voicing distinction in prevocalic position.

Tone systems are not static. A language can acquire tones and then increase the complexity of this tone system but it can also decrease the number of its tones and ultimately become non-tonal. These two processes, acquisition and recession of tones, have been termed tonogenesis (Matisoff, 1970, 1973) and tonoexodus (Lea, 1973). Cases of tonoexodus are rare[3] and it is not clear what the intermediate historical stages between the tonal and non-tonal stages are.

In the case of tonogenesis, the complete "scenario" is well attested. The development of contrastive tones on vowels due to the loss of a voicing distinction on obstruents in prevocalic position is probably the most well-documented type of tonogenesis. When such a development occurs, a relatively lower pitch register develops on vowels following the previously voiced series and a relatively higher pitch is found after the previously voiceless or voiceless aspirated series. Phonetic studies by House and Fairbanks (1953), Lehiste and Peterson (1961), Mohr (1971), Lea (1972, 1973), Löfqvist (1975) and Hombert (1976a) show how a voicing distinction in prevocalic position can affect the fundamental frequency of the following vowel.

Unfortunately, all these phonetic studies are based on non-tonal languages. Thus, they can explain how a non-tonal language can acquire two tones from the loss of some voicing distinction in prevocalic position but we do not have any strong basis for believing that these data can be extended to languages which are already tonal. Namely we do not know what the behavior of voiced or voiceless consonants is going to be at different frequency registers. Is a voiced consonant still going to affect the onset frequency of a vowel with low tone? Is a voiceless consonant going to perturbate the frequency of a high tone?

In order to answer these questions, two Yoruba subjects were asked to read[4] a word list of 42 CV tokens (2 consonants[5] x 7 vowels x 3 tones) put in the frame: sō CV sókè[6] ("say CV louder"). Each token was read five times. The results are presented in Figure 2. On this graph each point represents the average of 70 measurements (7 vowels x 5 repetitions x 2 subjects). The thin lines represent the fundamental frequency of vowels after [k] or [g]. The wide lines represent the average F_0 after voiced and voiceless velar stops for the three Yoruba tones.[7]

My purpose was to compare the effect of voicing contrast at different frequency registers. From the data presented in Figure 2, three

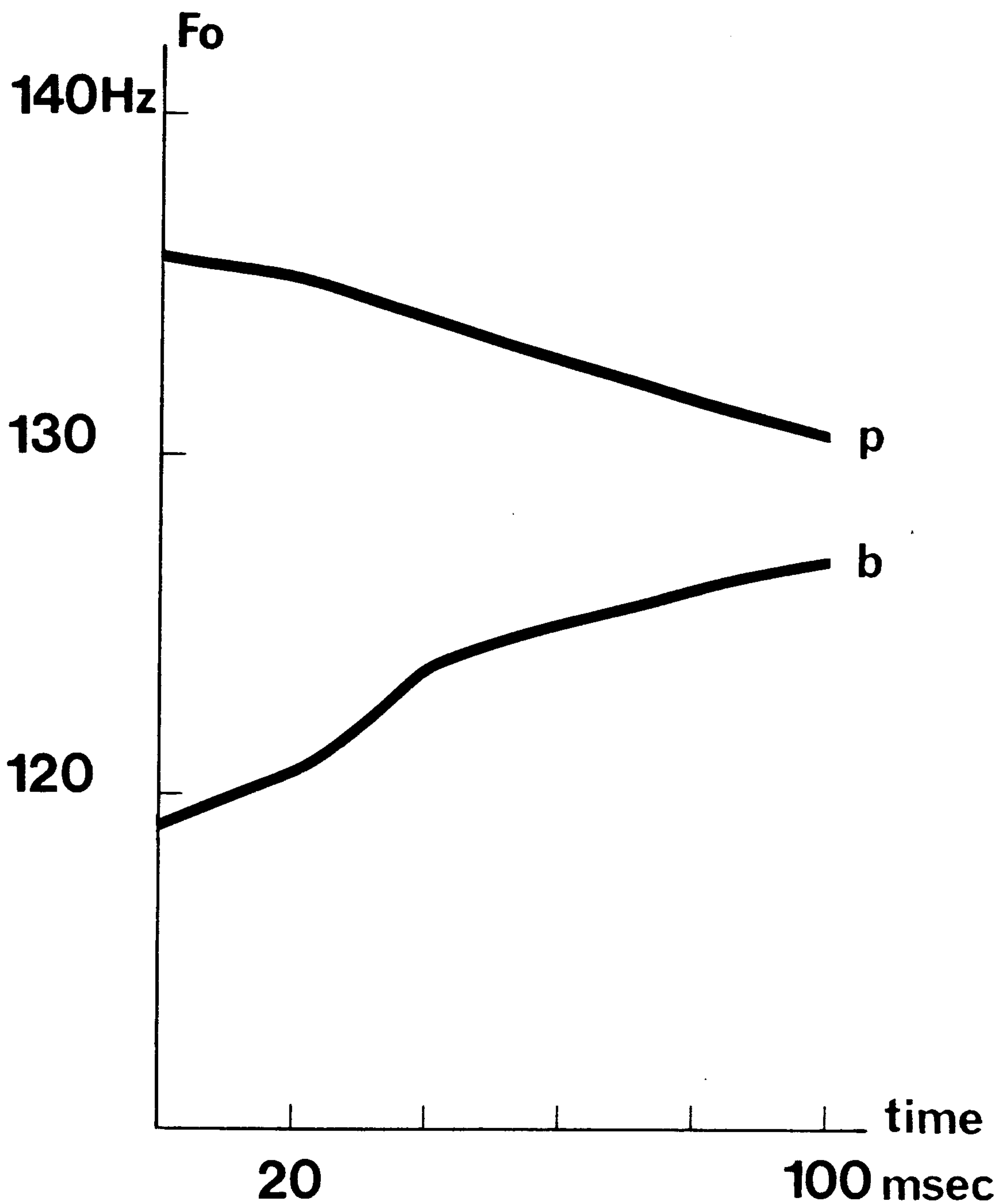


Figure 1: Averaged Fo of vowels after voiced and voiceless stops (from Hombert, 1976a) ([p] represents the voiceless stops and [b] represents the voiced stops)

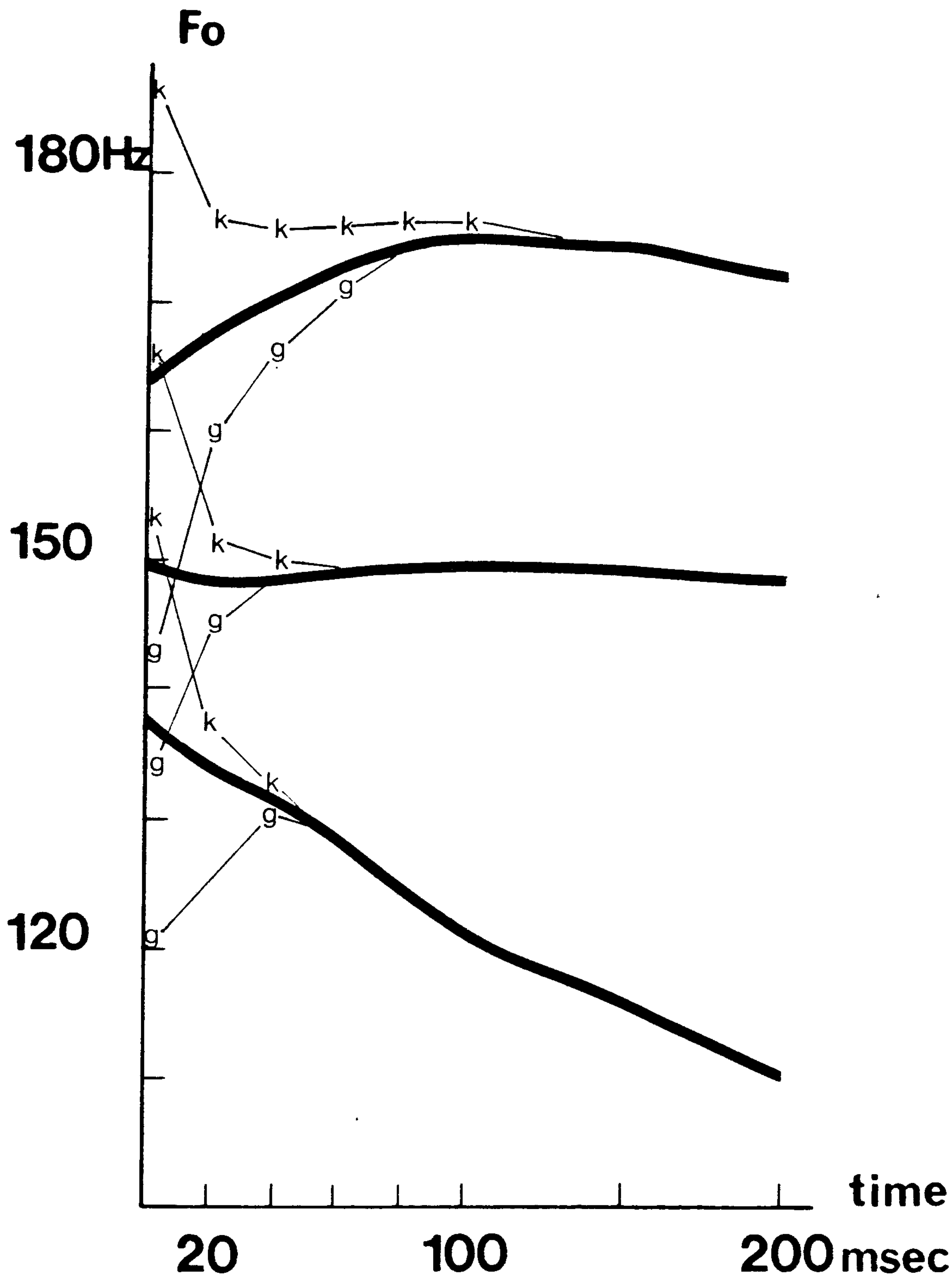


Figure 2: Influence of prevocalic voiced and voiceless stops on the fundamental frequency of the three Yoruba tones (2 subjects)

points should be emphasized:

1. The perturbation caused by a voiced consonant on a following high tone or by a voiceless consonant on a following low tone is greater than the effect of these two series of consonants on a mid tone.
2. The effect of a voiced consonant on a high tone is greater than the effect of a voiceless consonant on a following low tone.
3. The duration of the perturbations caused by prevocalic consonants on the fundamental frequency of vowels is shorter in a tone language (Yoruba) than in a non-tonal language (see Figure 1).

It is interesting to point out that these results are in agreement with Gandour's findings (1974) in his investigation of "Consonant types and tone in Siamese". Gandour found that a shorter part of the vowel was affected by the preceding vowel (about 30 msec for voiceless consonants and about 50 msec for voiced consonants). It seems that there is a tendency in tone languages (which does not exist in non-tonal languages) to actively minimize the intrinsic effect of prevocalic consonants probably in order to keep the different tones maximally different perceptually.

3. Intrinsic fundamental frequency of vowels

In the second part of this paper, I shall address myself to the issue of intrinsic fundamental frequency of vowels. Essentially, four theories have been proposed to explain why high[8] vowels have a higher intrinsic fundamental frequency than lower vowels.[9]

The first theory, proposed by Taylor (1933) and adopted by House and Fairbanks (1953) is called the "dynamo-genetic" theory. Taylor claims that the muscular tension of the tongue, required for the realization of high vowels, is transferred to the muscles of the larynx "via a kind of sympathetic resonance or radiation". This is not a viable theory anymore, since we know that electrical insulation in muscles and nerves is good enough to prevent serial contraction of adjacent muscles triggered by osmotic spread of excitability (Atkinson, 1973).

The second theory presented by Mohr (1971) relates width of the pharynx and pressure build-up behind the point of constriction to explain the fundamental frequency differences between low and high vowels. Since the width of the pharynx is about one-fourth as big for low back vowels as for corresponding high vowels, Mohr tries to relate smaller cavity and constriction further back with higher supraglottal pressure, leading to smaller pressure drop across the glottis and consequently lower fundamental frequency. Unfortunately Mohr's data do not support his theory. In any case, as Ohala (1973a) mentions

> "...the air flow during vowels is rapid enough and the magnitude of the pressure small enough that whatever back pressure is caused by the vowel constriction will be manifested equally rapidly for all vowels".

The next theory, known as the source tract coupling theory, was first proposed by Flanagan and Landgraf (1968). This theory assumes a possible coupling between the vocal cords and the vocal tract so that a low first formant (characteristic of high vowels) would attract and consequently raise the fundamental frequency. This sucking effect does not occur when the first formant is further away from the fundamental frequency (as it is in the case of low vowels). The intrinsic pitch difference between low and high vowels would then be explained. Unfortunately, predictions made by this theory do not receive empirical support. It would predict, for instance, that the difference in pitch between high and low vowels would be reduced when speaking with a helium-air mixture (since a property of helium (or other light gases) is to raise formants and consequently to increase the distance between F_0 and F_1). Beil (1962) showed that this was not the case.

The tongue pull theory (Ladefoged, 1964; Lehiste, 1970) is based on the assumption that when the tongue is in high position for the realization of high vowels, it exerts an extra tension transmitted to the larynx via the hyoid bone. This vertical pull increases the tension of the vocal cords (Ohala, 1972) and gives rise to a higher pitch for these high vowels. This theory ran into great difficulty when Ladefoged et al. (1972) provided data showing that tongue height and hyoid bone height were inversely proportional. Ohala (1973a) admits that such findings show that the pulling is not done through the hyoid bone but he maintains that the tongue pull theory is still a viable theory provided that the pulling is done through other tissues.

I would like to show in this paragraph that tone languages such as Yoruba can bring some very useful data in this controversy. If we assume a correlation between larynx height and Fo (Ohala and Ewan, 1972; Ewan and Krones, 1974), it seems that the tongue pull theory would predict that the F_0 difference would be smaller with vowels realized with high tones as opposed to vowels realized with low tones. Since the larynx is in higher position for high tones than for low tones, we expect that the tension exerted by the tongue will be less. This assumes a linear relationship between tension and larynx elevation (which would have to be tested). Figure 3 displays the averaged fundamental frequency values of the 7 Yoruba vowels depending on the tone under which they were realized. The measurements were made 100 msec after vowel onset. Each point is an average of 20 tokens (2 consonants x 5 repetitions x 2 subjects). From these data it is clear that the prediction made by the tongue pull theory is not verified; in fact the opposite is found, namely that the fundamental frequency difference between high and low vowels is more pronounced with high tone than with low tone.

Although it is obvious that more data are needed before refuting either the tongue pull theory or accepting the source/tract coupling theory, the data I just presented seems to be difficult to account for by the tongue pull theory.

The second point I would like to discuss with respect to intrinsic pitch of vowels has to do with tonogenesis. Phonetic studies indicate that the intrinsic perturbations caused by prevocalic consonants on the pitch of the following vowel are of about the same order of magnitude as the intrinsic differences depending on vowel height. Lehiste (1970) mentions

> "the influence of an initial consonant could counterbalance the influence of intrinsic pitch: the average for /kæ/ sequences was 171 Hz while that of /gi/ sequences amounted to 170 Hz."

However it is very rare to find languages making an extrinsic use of these intrinsic variations. In fact the only example I know of is Pilszczikowa-Chodak's study of Hausa (1972), in which she shows that tone assignment of verbs and noun plurals is largely predictable from the height of the final vowel in Hausa; high vowel predicting a high tone and vice versa. This confirms the prediction we would have made supposing that the intrinsic variations described earlier would become phonologized. Unfortunately this correlation has been criticized by Hausa scholars (Leben (personal communication), Newman, 1975). It seems important then to explain why the two kinds of tonal developments (from prevocalic consonants and from intrinsic F_0 of vowels) are not similarly attested.

First I want to show that different loudness levels cannot explain this asymmetry between the two potential possibilities of tonal developments:

We know that low vowels are perceived as louder than higher vowels (Fant, 1960) and we also know that loudness can affect our perception of pitch. If we can show that the *produced* intrinsic differences between high and low vowels are not *perceived* because of factors (such as loudness) affecting our perception of pitch, then we would have an explanation to the question of why languages do not develop tone from these intrinsic fundamental frequency differences.

Although the magnitude of the effect of loudness on pitch differs from one study to the other (Zurmühl, 1930; Stevens, 1935; Snow, 1936; Cohen, 1961), it is generally accepted that the effect of increased loudness (if any) (see Cohen, 1961) will be to lower the pitch (at least within this frequency region). Thus, this effect will lead to an *increase* in the pitch difference between high and low vowels and would make it more difficult to explain why tone developments based on these differences did not occur historically. In fact, instead of comparing the overall amplitude values of vowels just by stating that low vowels are louder than high vowels, it would be more appropriate to compare the amplitude values of different vowels within the frequency region relevant for pitch discrimination. It has been shown (Plomp, 1967) that the frequency region around the fourth harmonic is more important for pitch perception than the fundamental frequency region. Under the best conditions (around 300 Hz), high vowels are 10 dB louder than low vowels (Fant, 1960). This amplitude difference is not enough to cancel out the intrinsic differences obtained in production. Thus, this explanation based on loudness is not satisfactory.

The second suggestion I want to propose is based on the well-known fact that whatever the sensory modality, our auditory mechanism is more sensitive to a signal varying from state 1 to state 2 (with a variation $\Delta S = S_2 - S_1$) than to a static difference between the same two steady states S_1 and S_2 (Whitfield and Evans, 1965; Møller, 1973). If we apply

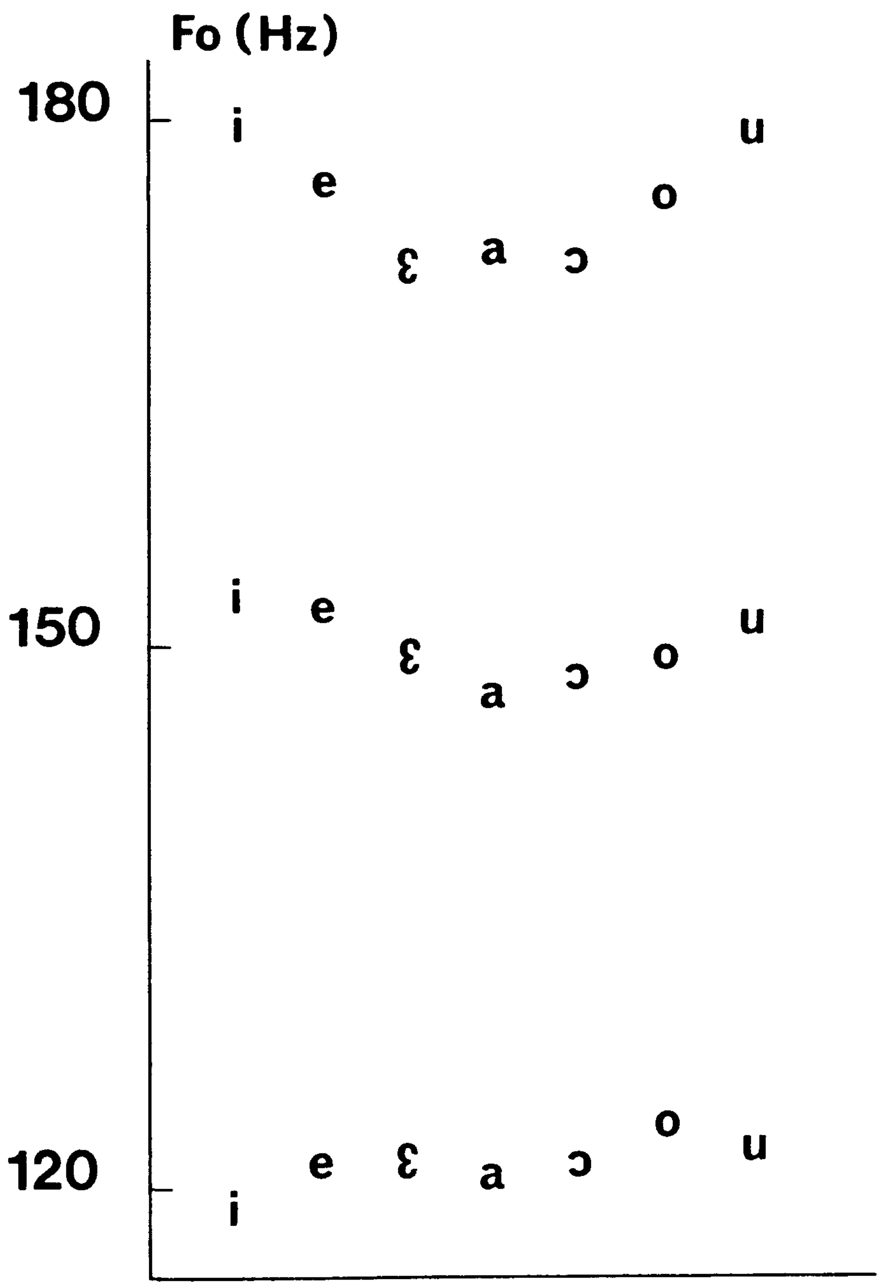


Figure 3: Intrinsic Fo of vowels for the three Yoruba tones (High, Mid and Low) (2 subjects)

this principle to pitch perception we can understand why tones develop from prevocalic consonants (where the intrinsic effects are realized as either rising or falling contours) but not from vowel height differences (where the intrinsic differences are in terms of steady state differences). A third possibility which is currently being tested (Hombert, 1976b) is that our perception of pitch of vowels is affected by vowel quality. Since intrinsic pitch and vowel quality are always associated for a given vowel (as opposed to the case of consonantal perturbation where the consonant can be removed) it is possible that the pitch of a high (vs low) vowel is lower (vs higher) than its intrinsic fundamental frequency as a result of a process of normalization done by our auditory system. If this is the case, the pitch differences between low and high vowels would be smaller than their fundamental frequency differences and this would account partially at least for the lack of development from vowel height.

4. Distribution of tones

It has been shown that if we consider intensity and frequency as two independent parameters, about 350,000 different pure tones can be discriminated by the human ear over the whole auditory area (Stevens and Davis, 1938; Wever, 1949; Licklider, 1956; Winckel, 1968). About 1500 of these tones are discriminated from pitch differences only. The fact that it is very rare to find tone languages with ten or more distinctive tones can be explained by the following reasons:

-- the amplitude parameter is not completely independent of the frequency parameter in speech (Hombert, 1975).

-- the fundamental frequency range is a much smaller range than the auditory range.

-- the historical unlikeliness of developing a huge number of tones in any one language.

-- the most important point is the difference between discrimination and identification. In contrast with our amazing ability to decide if two tones presented successively are similar or not, our identification ability (i.e., our ability to identify and name sounds) is rather poor. For speech and music it has been estimated that a trained listener can identify about 50 sounds presented individually. Winckel (1938) indicates that fluctuations of less than 20 Hz are imperceptible, in noise (i.e., in everyday situations).[10] Figure 2 indicates that the three tones in Yoruba are from 20 to 30 Hz apart.

Pollack (1952) shows that a maximum of five level tones can be distinguished under laboratory conditions[11] (when loudness cues have been removed).

Few languages have been reported as having as many as five level tones (Longacre, 1952). In fact it seems that when a language has more than two level tones, other cues than steady state fundamental frequency are used to identify some of the tones.

It has been shown by Pollack and Ficks (1954) that our auditory system is more efficient (i.e., transmits more information) at processing a stimulus with multiple encoding (i.e., more than one cue is used) as opposed to the processing of the same amount of information using a more sophisticated coding of only one parameter. This finding can be applied to pitch perception indicating that secondary cues such as loudness, duration and direction and speed of change, and phonation types are likely to be used in order to facilitate our tone identification (Hombert, in preparation). From Figure 2, we can see that the so-called "low" tone in Yoruba is in fact a falling tone. The fact that the falling contour is a more important cue than the F_0 level has been shown by LaVelle (1974). LaVelle argues that Yoruba has a rule which lowers a low tone in phrase final position. I propose that the "unmarked" realization of the Yoruba low tone is in fact a low falling tone and that the falling tone is realized as a low level tone when followed by a non-low tone (i.e., when followed by either a mid or a high tone). When a low tone occurs in isolation or in final position, it is realized with its more distinctive perceptual characteristics. When it is followed by a higher tone, articulatory constraints (the complexity of going down for a falling tone and then up for the next tone) take over and the low falling tone is realized as a low level tone. But in that environment perception is made easier by the following tone which can be used as a reference for comparison. It has been shown by Han and Kim (1974) that "the pitch information in the neighbouring syllables serves as a basis on which the retrieval of the phonemic status of a tone is made".

5. Conclusion

In this paper, I have shown:

1. How the three Yoruba tones were affected by voiced/voiceless consonants in prevocalic position. Voiced consonants have a greater effect on high tones and voiceless consonants on low tones. The duration of these perturbations is smaller in Yoruba than in a non-tonal language such as English.

These data show that an already existing tone system can multiply the number of its tones if the voicing contrast in prevocalic position is to disappear.

2. The fact that the intrinsic F_0 differences between high and low vowels are smaller when these vowels are realized with a low tone as opposed to a high tone seems to be a counter-argument to the tongue-pull theory generally proposed to account for these intrinsic F_0 differences between high and low vowels.

In the discussion it was also suggested that the fact that tones do not develop from intrinsic pitch of vowels as they do from intrinsic F_0 perturbations caused by prevocalic consonants cannot be explained by loudness but may be by the dynamic effect of prevocalic consonants as opposed to the static effect of vowel height.

3. It is often the case that other cues than steady state F_0 are used

to identify tones. It seems that Yoruba "low" tone is in fact a low falling tone, but that this low falling tone is changed into a low level tone when followed by a non-low tone.

Acknowledgements

I would like to thank L. Hyman, J. Ohala and the members of the UCLA tone seminar for their helpful comments and suggestions on this paper. I also want to thank A. Bamgboṣe, B. Elugbe, M. Lindau and K. Williamson for their help when I was collecting my data at University of Ibadan, Nigeria. Special thanks to Will Leben for discussing this paper with me and reading it at the 6th Conference on African Linguistics, Columbus, Ohio, April, 1975.

Footnotes

1. Analysis of variance followed by Duncan's test.

2. With a 1% confidence level.

3. Some cases can be found among Eastern Bantu languages.

4. The recording was done at University of Ibadan, Nigeria, and the data analysis was done in the Phonology Laboratory, University of California, Berkeley.

5. Only 2 consonants [k] and [g] were used in this experiment to represent voiced and voiceless stops. Recent data seem to indicate that velar stops have a more important perturbatory effect on the F_0 of the following vowel than other stops (Hombert and Ladefoged, 1976; Meyers, 1976).

6. The three Yoruba tones are represented as follows:

   High Tone ´
   Mid Tone ¯
   Low Tone `

7. When the F_0 values after voiced or voiceless stops were too close to the average value (wide line), they were not represented on Figure 2.

8. High refers to the tongue-height parameter.

9. See Atkinson (1973) and Ohala (1973a) for a more detailed review of these different theories.

10. 20 Hz is probably an overestimation considering on the one hand the richness of certain tone languages (in terms of number of tones) and on the other hand the maximum frequency range of vocal cord vibrations.

11. Pollack's experiment was done using a 100-5000 Hz frequency range, but his data show that the number of identifiable tones is negligibly affected by a lowering of the upper limit of the frequency range.

References

Atkinson, J. E. (1973) "Aspects of intonation in speech: Implications from an experimental study of fundamental frequency" Ph.D. Thesis. University of Connecticut.

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

Beil, R. G. (1962) "Frequency analysis of vowels produced in Helium-rich atmosphere" *JASA* 34:347-49.

Capell, A. (1949) "Two tonal languages of New Guinea" *Bulletin of the School of Oriental and African Studies* 8:184-99.

Chistovich, L. A. (1969) "Variation of the fundamental voice pitch as a discriminatory cue for consonants" *Soviet Physics - Acoustics* V. 14, no. 3. pp. 372-8.

Cohen, A. (1961) "Further investigation of the effects of intensity upon the pitch of pure tones" *JASA* 33:1363-76.

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

Fant, G. (1960) *Acoustic Theory of Speech Production* Mouton: The Hague.

Flanagan, J. L. and L. Landgraf (1968) "Self-oscillating source for vocal tract synthesizers" *IEEE Trans. Audio* 16:57-64.

Fujimura, O. (1971) "Remarks on stop consonants" In *Form and Substance* Akademisk Forlag. pp. 221-32.

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

Haggard, M., S. Ambler and M. Callow. (1970) "Pitch as a voicing cue" *JASA* 47, No. 2, 613-7.

Han, M. S. and K. O. Kim (1974) "Phonetic variation of Vietnamese tones in disyllabic utterances" *Journal of Phonetics* 2:223-32.

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'extrême-orient" *Bulletin de la Société Linguistique de Paris* 56:163-180.

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. Thesis. University of California, Berkeley.

Hombert, J.-M. (1976a) "Phonetic explanation of the development of tones from prevocalic consonants" In this volume.

Hombert, J.-M. (1976b) "Development of tones from vowel height" In this volume.

Hombert, J.-M. (In preparation) "Production and perception of tone of Yoruba nouns"

Hombert, J.-M. and P. Ladefoged (1976) "Effect of aspiration on the fundamental frequency of the following vowel" *Journal of the Acoustical Society of America* 59(A) S72.

House, A. S. and G. Fairbanks (1953) "The influence of consonant environment upon the secondary acoustical characteristics of vowels" *JASA* 25: No. 3:105-13.

Hyman, L. M. (1973) Ed. "Consonant types and tone" *Southern California Occasional Papers in Linguistics* No. 1.

Hyman, L. M. (1973b) "The role of consonant types in natural tonal assimilations" *Consonant Types and Tone* (Ed. L. Hyman). *Southern California Occasional Papers in Linguistics* No. 1:153-179.

Karlgren, B. 1926. "Etudes sur la phonologie Chinoise" *Archives d'Etudes Orientales* Vol. 15.

Ladefoged, P. (1964) *A Phonetic Study of West African Languages* Cambridge Univ. Press.

Ladefoged, P., J. DeClerk, M. Lindau, G. Papçun. (1972) "An auditory-motor theory of speech production" In *Working Papers in Phonetics* UCLA 22:48-75.

LaVelle, C. R. (1974) "An experimental study of Yoruba tone" *Studies in African Linguistics* Supp. 5:185-194.

Lea, W. (1972) "Intonational cues to the constituent structure and phonemics of spoken English" Ph.D. Thesis. Purdue.

Lea, W. (1973) "Segmental and suprasegmental influences on fundamental frequency contours" In *Consonant Types and Tone* (ed. L. Hyman) pp 15-70.

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

Lehiste, I., and G. Peterson (1961) "Some basic considerations in the analysis of intonation" *JASA* 33:419-25.

Licklider, J.C. (1956) "Auditory frequency analysis" *Proc. 3rd London Symp. on Information Theory* pp.253-68 London.

Longacre, R. (1952) "Five phonemic pitch levels in Trique" *Acta Linguistica* 7:62-82.

Löfqvist, A. (1975) "Intrinsic and extrinsic F_0 variations in Swedish tonal accents" *Phonetica*:31:228-47.

Luce, G. H. (1959) "Introduction to the comparative study of Karen languages" *J. BURM.* R.S. 42,1:1-18.

Maspéro, H. (1912) "Etudes sur la phonétique historique de la langue annamite: les initiales" *BEFEO*, XII.

Matisoff, J. (1970) "Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study" *Journal of the American Oriental Society* 90,1:13-44.

Matisoff, J. (1973) "Tonogenesis in Southeast Asia" In *Consonant Types and Tone* Ed. L. Hyman. pp. 71-96.

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

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

Møller, A. R. (1973) "Coding of amplitude and frequency modulated sounds in the cochlear nucleus" Paper presented at the Symposium on Auditory Analysis and Perception of Speech. Leningrad. Aug. 21-24, 1973.

Newman, P. (1975) "The non-correlation of tone and vowel height in Hausa" *Studies in African Linguistics* 6.2:207-213.

Ohala, J. (1972) "How is pitch lowered" (Abstract) *JASA* 52:124.

Ohala, J. (1973a) "Explanations for the intrinsic pitch of vowels" *MIM* Jan. 1973:9-26.

Ohala, J. (1973b) "The physiology of tone" In *Consonant types and tone*. (Ed. L. Hyman). pp 1-14.

Ohala, J. and W. Ewan (1973) "Speed of pitch change" (Abstract) *JASA* 53:345.

Pliszczikowa-Chodak, N. (1972) "Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa" In *Studies in African Linguistics* 3:399-422.

Plomp, R. (1967) "Pitch of complex tones" *JASA* 41:1526-1533.

Pollack, I. (1952) "The information of elementary auditory displays" *JASA* 24:745-49.

Pollack, I. and L. Ficks (1954) "Information of elementary and multi-dimensional auditory displays" *JASA* 26:155-158.

Shafer, R. (1953) "Etudes sur l'Austroasien" *B.S.L.P.* 48,1:114-115.

Snow, W. B. (1936) "Change of pitch with loudness at low frequencies" *JASA* 8:14-19.

Stevens, S. S. (1935) "The relation of pitch to intensity" *JASA* 6:150-54.

Stevens, S. S. and H. D. Davis (1938) *Hearing: Its Psychology and Physiology* J. Wiley & Sons. New York.

Taylor, H. 1933. "The fundamental pitch of English vowels" *Journal of Exp. Psychol.* 16:565-582.

Wever, E. G. (1949) *Theory of Hearing* J. Wiley & Sons. New York.

Whitfield, E. G. and E. F. Evans (1965) "Responses of auditory cortical neurons to stimuli of changing frequency" *J. Neurophysiology* 28:655-72.

Zurmühl, G. (1930) "Abhängigkeit der tonhöhenempfindung von der Lautstärke und ihre Beziehung zur Helmholtzschen Resonanztheorie des Hörens" *Z. Sinnesphysiol.* 61:40-86.

Development of Tones from Vowel Height

Jean-Marie Hombert

[Department of Linguistics, University of

California, Los Angeles]

Introduction

In this paper I will attempt to provide one of the major reasons why the intrinsic fundamental frequency variations caused by vowel height rarely, if at all, give rise to the development of phonological tones. Explanations of sound changes based on physiological constraints of our articulatory and/or auditory mechanisms have been proposed for a long time (Durand, 1938; Grammont, 1933; Haden, 1938; Passy, 1891; Paul, 1891; Rousselot, 1891; Sweet, 1888). The interest in finding such explanations for widely (i.e. non-language specific) attested sound changes has been revived recently (Hombert,1974, 1975a, b, c; Hombert, Ohala and Ewan, 1975, in preparation; Ohala,1974 a, b, 1975). These explanations imply that the pronunciation intended by the speaker may get distorted by the time it is perceived by the listener. This may be due to the action of articulatory constraints which affect the way the sounds were uttered or by the action of auditory constraints which affect the way the sounds were analyzed by the listener's ear. In order to demonstrate that a sound change is phonetically motivated, one has to demonstrate first that these intrinsic perturbations are present in the speech signal (at least in the case of articulatory motivated sound changes) and second, that their magnitude is sufficient to be perceived.

The role played by experimental phonetics in explaining sound changes is clear from this scheme. But the interaction between phonetics and historical linguistics is not unidirectional. In this paper, I want to illustrate a case in which historical data can lead us to new insights concerning perception and speech processing. I mentioned earlier that a phonetically motivated sound change will be explained if experimental data show that intrinsic perturbations (non-intended by the speaker) are present in the speech signal and of sufficient magnitude to be heard by the listener. Now let us suppose we can demonstrate these two points despite the fact that historical data do not show that this historical change is widely attested (i.e. probably phonetically motivated). At least two reasons can be proposed to account for this discrepancy: (1)

linguists have failed to report this type of sound change perhaps because of a lack of focus in the linguistic literature on such a possibility. (2) This sound change simply did not occur. If this second case is valid, this may imply that our previous investigations of the existence and the perception of the intrinsic perturbations may have been too hasty or too superficial and consequently have to be reconsidered.

Phonetic data

It has been shown (Hombert, 1974) that the most well documented type of tonogenesis, namely the development of contrastive tones on vowels due to the loss of a voicing distinction on obstruents in prevocalic position is phonetically motivated. In this paper I want to address myself to the following question: Why is it that such a development does not occur from the intrinsic fundamental frequencies caused by vowel height? First, let us examine data concerning the intrinsic fundamental frequencies of vowels in American English. Several studies have shown that vowels have an intrinsic fundamental frequency related to their height: high vowels (low F1) have a higher fundamental frequency than low vowels (House and Fairbanks, 1953; Lehiste and Peterson, 1961; Peterson and Barney, 1952) as shown in table 1.

Table 1. Intrinsic fundamental frequency of vowels (in Hz).

		i	a	u
House and Fairbanks (1953)	(10 subjects)	127.9	118.0	129.8
Lehiste and Peterson (1961)	(5 subjects)	129	120	134
Peterson and Barney (1952)	(33 subjects)	136	124	141

The data presented in table 1 indicate that the intrinsic variations of fundamental frequency due to vowel height are of the same order of magnitude as the intrinsic differences caused by prevocalic consonants (Hombert, 1974; House and Fairbanks, 1953; Lea, 1972, 1973; Lehiste and Peterson, 1961; Löfqvist, 1975; Mohr, 1971). In fact, Lehiste (1970:71) mentions "The influence of an initial consonant could counterbalance the influence of intrinsic pitch: the average for /kæ/ sequences was 171 Hz, while that of /gi/ sequences amounted to

170 Hz." Since vowel height and prevocalic obstruents seem to cause comparable intrinsic perturbations, one would expect tonal development resulting in vowel merging to be as frequent as tonal development resulting in the loss of some voicing contrast in prevocalic position, and one would expect the development to show that high vowels give rise to high tones and low vowels to low tones. However, historical data illustrating such developments are not straightforward.

Historical data

Pilzczkowa-Chodak (1972) suggests that tone assignment of verbs and noun plurals in Hausa is largely predictable from the height of the final vowel: a high (vs. low) final vowel predicting a high (vs. low) tone. This analysis, however, has been criticized by Hausa scholars (Newman, 1975; Leben and Schuh, personal communications). It seems that Middle Chinese words with checked tones (i.e. p,t,k endings) and voiceless initial consonants developed a relatively lower tone when the vowel nucleus was [a] than when it was [ə] (Baron, in preparation; Pulleyblank, 1970-1). In some Cantonese dialects, this tone development has sometimes been analyzed as originating from a length contrast. In the Omei dialect of Mandarin, two tones rearranged themselves depending on vowel height; the 'new' high tone regrouping high vowels (Baron, in preparation; Cheung, 1973). In Ngizim (Schuh, 1971) and in Bade, the tone patterns of verbs are partially predictable from the vowel of the first syllable; if the vowel is [a], the verb will have a high tone. The literature offers a few other examples which seem to indicate an interaction between vowel height and tone. In Foochow (Yuan et al, 1960; Chen and Norman, 1965; Maddieson, 1975) the vowel is raised if the tone of some lexical items is replaced by a higher tone. A similar phenomenon occurs in Lahu (Matisoff, 1973) where the rising tone can raise a vowel. These last two cases (Foochow and Lahu) can be interpreted as an effect of tones on vowels. It would seem that the interaction between tones and vowel height works only in one direction: tone can affect vowel height but not vice-versa. As I mentioned earlier, historical data do not suggest that the development of contrastive tones from vowel height is a widely attested process; furthermore, the reverse direction of interaction (i.e. low vowels giving rise to high tones) as observed in Ngizim and Bade seems unexplainable phonetically. It would seem reasonable then to find an explanation for the infrequency of this type of effect.

Several reasons can be proposed. One would expect that such a tonal development would start where the intrinsic differences are maximal, that is from the merging of low and high vowels. This seems unlikely since even very restricted vowel systems contain i, a and u. One can also argue that the intrinsic perturbations caused by prevocalic consonants are of a different nature than those caused by vowel height. A voiced (vs. voiceless) consonant causes a rising (vs. falling) fundamental frequency pattern at the onset of the following vowel. On the other hand, the intrinsic fundamental frequency associated with different vowel qualities is manifested by differences in steady-state fundamental frequency levels. Our auditory system is more "efficient" at detecting changes in varying fundamental frequency signals rather than differences (of the same magnitude) between two steady-state fundamental frequency signals.

Another possibility is that our perception of pitch of vowels is affected by vowel quality. Since intrinsic pitch and vowel quality are always associated for a given vowel (as opposed to the case of consonantal perturbation where the consonant can be removed) it is possible that either the pitch of vowels is affected by the spectrum envelopes (effect of loudness at certain frequencies) or that the pitch of high (vs. low) is perceived lower (vs. higher) as a result of some sort of normalization. The following experiment was designed to test this hypothesis.

Experimental paradigm

Ten subjects (5 females and 5 males) all native speakers of American English participated in this experiment. They were asked to compare the pitch of two synthesized vowels (generated by a software synthesizer) of different quality. Only those who were able to achieve a score of 95% or better in a control experiment involving the comparison of the pitch of two pure tones were selected as subjects. The formant values of the three vowels used in this experiment are given in table 2.

Table 2. Formant values of synthesized vowels i, a, u.

	i	a	u
F1 (in Hz)	270	730	300
F2	2600	1090	870
F3	2800	2440	2240

These values are taken from Peterson and Barney (1952) except for the second and third formants of [i] which have been raised in order to sound more speech-like (Delattre et al (1952)). Three fundamental frequencies, 115 Hz, 120 Hz and 125 Hz (with an accuracy of ± 0.5%) were superimposed on each of these three vowels which had a duration of 250 msec (with a rising and decay time of 20 msec). The interval separating the first (V1) and second vowel (V2) was 500 msec. The fundamental frequency of the second vowel was either 5 Hz below, equal to, or 5 Hz above the fundamental frequency of the first vowel (in other words, the fundamental frequency range of V2 was from 110 Hz to 130 Hz). V1 and V2 always differed in quality (i.e. there were no i-i, a-a, or u-u sequences). Six repetitions of all possible V1-V2 comparisons were presented (excluding cases where V1 and V2 had the same quality) making a total of 324 judgements (3V1 qualities x 3V1 F0 x 2V2 qualities x 3V2 F0 x 6 repetitions). Over-all amplitude levels were equalized for the three vowels.

Subjects were asked to judge whether the first or the second vowel was higher in pitch (i.e. a two-way forced choice). They were instructed to mark the corresponding vowel on their answer sheet. They had three seconds in which to make a response. The experiment was divided into two parts preceded by a short training session in which six pairs were presented. The same stimulus presentation format was adopted in the control experiment in which synthesized vowels were replaced by pure tones. The role of this control experiment was twofold; it was used to determine the accuracy of the subjects' pitch perception and also to investigate the effect of stimulus ordering on the perception of pitch especially when the two tones had the same frequency.

Results and Discussion

Since the criterion used to select subjects was quite strict (minimum of 95% correct in the control experiment), they made very few mistakes when the vowels were 5 Hz apart (90% correct or better). As a result, in the following analysis I will only consider the pairs which had identical F0 on both vowels. The subjects' responses for these pairs are presented in table 3.

Table 3. Number of times each vowel was judged to be the higher in pitch when comparing the pitch of vowels of different quality but equal fundamental frequency (10 Subjects).

i:a comparison (i-a and a-i pairs)	u:a comparison (u-a and a-u pairs)	u:i comparison (u-i and i-u pairs)
a = 257	a = 261	i = 185
i = 103	u = 99	u = 175

From this table, it is clear that the low vowel [a] has a tendency to be judged higher in pitch than the high vowels [i] or [u] (although their fundamental frequencies were in fact equal).

Let us now look in more details at the pairs involving comparisons of low (a) vs. high (i or u) vowels. In table 4, data showing the effect of fundamental frequency level and ordering for these pairs are presented.

Table 4. Number of times each vowel was judged to be higher in pitch when comparing the pitch of vowels of different quality (but equal FO) as a function of the fundamental frequency level and the order of presentation of these two vowels (10 Subjects).

	i:a comparison						u:a comparison					
	i:a			a-i			u-a			a-u		
	FO			FO			FO			FO		
	115	120	125	115	120	125	115	120	125	115	120	125
Low vowel a judged higher in pitch	44	39	48	46	36	44	48	36	42	57	47	34
High vowel i or u judged higher in pitch	16	21	12	14	24	16	12	24	18	3	16	26

These data were subjected to a three-way analysis of variance. The hypothesis being tested (i.e. low vowel being judged higher in pitch than high vowel of equal fundamental frequency) was found to be statistically significant ($p < 0.0005$). The ordering effect was found to be non-significant. It can be noticed in table 4 that the differences in the number of "high" judgements for the low vowel [a] vs. the high vowel [i] or [u] is smaller at 120 Hz and moreover the responses for the a-u pair at 115 Hz deviates considerably from the ratio found under other conditions. These two points suggest that the location of the relevant harmonics with respect to the formants may play a role in pitch perception. In fact, the statistical analysis revealed that the effect of fundamental frequency level was significant at the 0.02 level. Although this last point requires further investigation, the basic point has been demonstrated. Since the low vowel [a] is perceived higher in pitch than the high vowels [i] and [u] when compared at equal fundamental frequency, this implies that vowel quality affects pitch perception of vowels. The pitch difference between high and low vowels is smaller than their fundamental frequency differences. This is an important point for explaining why intrinsic fundamental frequency differences caused by vowel height did not give rise to tonal development as frequently as we would have expected considering the magnitude of these intrinsic differences.

These data demonstrate that the effect exists, that the perceived pitch difference between high (low F1) and low (high F1) vowels is smaller than their fundamental frequency difference would indicate. However, the origin of this effect has not been demonstrated. Two possibilities should be investigated:

1) it is a low level phenomenon which can be explained by vowel spectra characteristics

2) it is a higher level phenomenon involving some normalization of fundamental frequency depending on vowel height.

Although it is clear that a controlled experiment has to be done in order to accept or reject one of these hypotheses, available relevant experimental data do not seem to favor the first possibility. It has been known for quite a long time that intensity affects our pitch perception. Stevens (1935) shows that the pitch of a tone whose frequency is 150 Hz is lowered by as much as 11% when its intensity is raised by 50 dB. More recent data (Cohen, 1961) suggest a smaller effect (2 to 4%). In order to be able to apply these data to vowels, we have to know which frequency region is the most relevant for pitch perception. Ritsma (1967) indicates that for fundamental frequencies in the range of 100-400 Hz (i.e. the relevant range for speech), the frequency region consisting of the third, fourth and fifth harmonics plays a dominant role in the perception of pitch. Since the vowels used in this experiment had a FO of 115, 120 or 125 Hz, the dominant frequency region is located between 345 (3 x 115) and 600 Hz (5 x 125). Since we know (Cohen, 1961; Snow, 1936; Stevens, 1935; Zurmühl, 1930) that by increasing loudness we lower the pitch (assuming that we are dealing with a tone below 1000 Hz), our results would be explained if we could show that the amplitude of the spectrum of [i] or [u] is sufficiently greater than the amplitude of [a] in the relevant frequency region (i.e. approximatively 345-600 Hz). In fact, as can be seen from Figure 1, the intensity of [a] (is higher than the intensity of [i] or [u] in a major part of this frequency region) and consequently would not explain our results.

The second possibility would imply an interaction between Fo and vowel quality at a higher neurological level. Since vowel quality and intrinsic FO are always produced simultaneously, it is possible that our auditory system subjects the speech signal to some form of normalization having the effect of raising the pitch of low vowels (or lowering the pitch of high vowels). It should be remembered that similar effects have been investigated and described: for instance, the size-weight effect by which a small size box is judged heavier than a bigger size box of equal physical weight (Birnbaum and Veit, 1974). It should be noted that this experiment indicates only the direction of the effect. The only quantitative information is that for these selected subjects, under these experimental conditions, the effect was smaller than 5 Hz. Before making definitive conclusions about the relevance of these data for explaining the non-development of tones from vowel height, two stages are required. First, a quantification of the effect under laboratory conditions (i.e. by presenting pairs of vowels differing both in quality and FO) and second, a quantification of the relationship between laboratory conditions and 'real world' conditions.

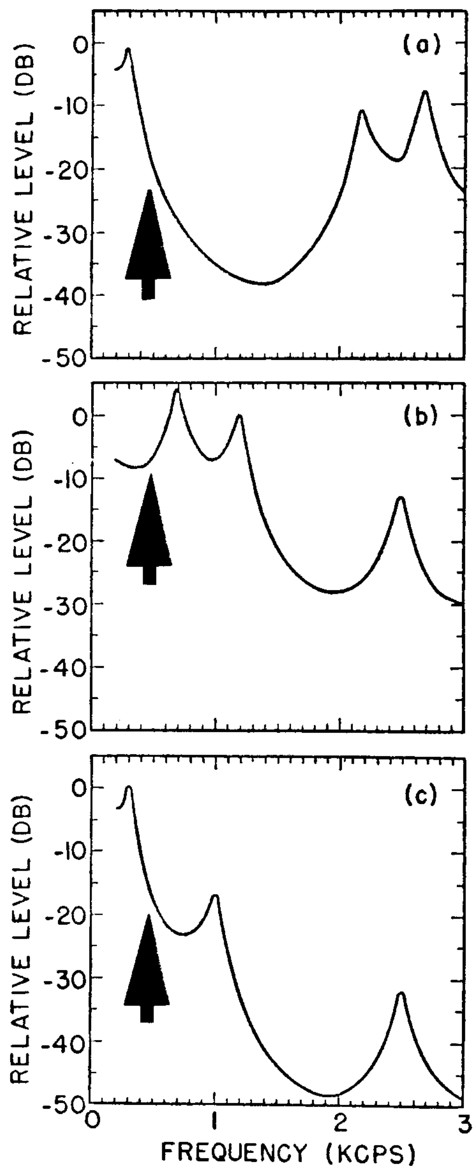


Figure 1. Spectrum of vowels i, a, u from top to bottom (from Stevens and House, 1961). Arrows indicate dominant frequency region for pitch perception.

Conclusion

In summary I have shown that the study of historical developments, or more precisely in this case, the lack of historical developments can lead to new insights concerning perception of speech. I presented data indicating that the pitch differences between high and low vowels were smaller than the corresponding intrinsic fundamental frequency differences. This is taken as a partial explanation for the lack of tonal development from vowel height.

Acknowledgements

The research reported in this paper was supported by a NSF Grant SOC75-07158, made to the UCLA Phonetics Laboratory. I wish to thank the members of the UCLA tone seminar for their valuable comments. Special thanks are owed to Peter Ladefoged, Victoria Fromkin, Hector Javkin, Steve Greenberg and Louis Goldstein, Larry Hyman, Richard Harshman and John Ohala for their criticism and discussion and also to W. Ballard, Steve Baron, Jim Matisoff, Martine Mazaudon and Russ Schuh for discussing the historical data with me.

References

Baron, S. (in preparation) "On tone and vowel height in Chinese."

Birnbaum, M.H. and Veit, C.T. (1974) 'Scale-free tests of an additive model for the size-weight illusion.' *Perception and Psychophysics* 16:2, 278-282.

Chen, L. and J. Norman (1965) "An introduction to the Foochow dialect." San Francisco State College in cooperation with the U.S. Office of Education.

Cheung, S.H.N. (1973) "Tonal redistribution in the Omei dialect." Paper presented at the 6th International Sino-Tibetan Conference, San Diego.

Cohen, A. (1961) 'Further investigation of the effects of intensity upon the pitch of pure tones.' *Journ. of the Acoust. Soc. of America*, 33, 1363-1376.

Delattre, P., A.M. Liberman, F.S. Cooper and L.J. Gerstman (1952) "An experimental study of the acoustic determinants of vowel color; observations on one-and two-formant vowels synthesized from spectrographic patterns. *Word*, 8.3, pp. 195-210.

Durand, M. (1956) "Du rôle de l'auditeur dans la formation des sons du language. *J. Psychol. Norm. Pathology*, 52: 347-55.

Ewan, W.G. (1975) "Explaining the intrinsic pitch of vowels." Paper presented at the Fifth California Linguistics Association Con. San Jose, May 1975.

Ewan, W.G. (forthcoming) "Laryngeal behavior in speech" Univ. of Calif. Berkeley, doctoral dissertation.

Flanagan, J.L. and M.G. Saslow (1958) "Pitch discrimination for synthetic vowels." *Journ. Acoust. Soc. Amer.*30:5, pp. 435-442.

Grammont, M. (1933) "Traité de Phonétique." Paris, Librairie Delagrave.

Haden, E.F. (1938) "The physiology of French consonant changes." *Language Dissertations*. 26.

Hombert, J.M. (1974) "Production and perception: their respective roles in one case of tonal development." Paper presented at the Annual meeting of the Linguistic Society of America. New York. Dec. 1974.

Hombert, J.M. (1975a) "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. To be published in the Working Papers from the Phonology Lab, UC Berkeley.

Hombert, J.M. (1975b) "Perception of contour tones: an experimental investigation." in *Proceedings of the first annual meeting of the Berkeley Linguistics Society*. pp. 221-232.

Hombert, J.M. (1975c) "Development of tone from segmentals: evidence from contour tone perception." To be published in the *Proceedings of the Eighth International Congress of Phonetic Sciences*. Leeds, Aug. 1975.

Hombert, J.M. (1975d) "Consonant types, vowel height and tone in Yoruba." Paper presented at the Sixth Annual Conference on African Linguistics. Ohio State Univ. April,1975. **In this volume.**

Hombert, J.M. J. J. Ohala and W.G. Ewan (1975) "Tonogenesis: theories and queries." Paper presented at the 8th Sino-Tibetan Conf. Berkeley, Oct. 1975.

Hombert, J.M., J.J. Ohala and W.G. Ewan (in preparation) "Phonetically motivated sound changes."

House, A.S. and G. Fairbanks (1953) "The influence of consonant environment upon the secondary acoustical characteristics of vowels." *Journ. of the Acoustical Soc. of Amer*. 25.1. Reprinted in *Readings in Acoustic Phonetics*, edited by Ilse Lehiste, M.I.T. Press.

Klatt, D.H. (1973) "Discrimination of fundamental frequency contours in synthetic speech: implications for models of pitch perception." *Journ. of the Acoust. Soc. of Amer.* 53.1, pp. 8-16.

Lea, W. (1972) "Intonational cues to the constituent structure and phonemics of spoken English. Purdue Doctoral Dissertation.

Lea, W. (1973) "Segmental and Suprasegmental influences on fundamental frequency contours." in *Consonant types and tone* (Ed. L. Hyman) p. 15-70.

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

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

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

Maddieson, I. (1975) "The intrinsic pitch of vowels and tones in Foochow" In this volume.

Matisoff, J.A. (1973) The Grammar of Lahu. *Publications in Linguistics* 75, Univ. of Calif. Press. Berkeley and Los Angeles.

Mazaudon, M. (1975) "Tibeto-Burman tonogenetics." To appear in *Current trends in Sinotibetan Linguistics*. L.R. Maran (ed.).

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

Møller, A.R. (1973) "Coding of amplitude and frequency modulated sounds in the cochlear nucleus. Paper presented at the Symposium on Auditory Analysis and Perception of Speech. Leningrad Aug. 21-24, 1973.

Newman, P. (1975) "The non-correlation of tone and vowel height in Hausa." *Studies in African Ling.* 6:2, 207-213.

Ohala, J. (1973) "Explanations for the intrinsic pitch of vowels." Monthly Internal Memo, Jan.

Ohala, J. (1974a) "Phonetic explanation in phonology" in *Papers from the parasession on natural phonology, pp.* 251-274, Bruck A. R.A. Fox and M.W. Lagaly (eds.). Chicago Linguistic Soc., Chicago, Ill.

Ohala, J. (1974b) "Experimental historical phonology." in *Historical Linguistics II. Theory and description in phonology.* Anderson, J.M. and C. Jones (eds). (Proceedings of the First Intern. Conference on Historical Linguistics, Edinburgh). pp. 353-389. Amsterdam, N. Holland Publ. Co.

Ohala, J. (1975) "Phonetic explanations for nasal sound patterns. Papers from the Nasalfest, Stanford.

Passy, P. (1891) *Etudes sur les changements phonétiques.* Paris, Librairie Firmin-Didot.

Paul, H. (1891) *Principles of the History of Language.* London.

Peterson, G.E. (1961) "Parameters of vowel quality." *Journ. of Speech and Hear. Disorders,* 4: 10-29.

Peterson, G.E. and H.L. Barney (1952) "Control methods used in a study of the vowels." *Journ. of the Acoust. Soc. of Amer.* 24, 175-184. Reprinted in *Readings in Acoustic Phonetics,* edited by I. Lehiste, M.I.T. Press (1967).

Pilszczikowa-Chodak, N. (1972) "Tone-vowel height correlation and tone assignment inthe patterns of verb and noun plurals in Hausa." in *Studies in African Linguistics III,* pp. 399-422.

Pulleyblank, E.G. (1970-1) "Late middle Chinese." *Asia Major* 15: 197-239 and 16: 121-168.

Ritsma, R. (1967) 'Frequencies dominant in the perception of complex sounds.' *Journ. of the Acoust. Soc. of Amer.* 42, 191-198.

Rousselot, L'abbé (1891) *Les Modifications Phonétiques du Langage* Paris. H. Welter.

Schuh, R. (1971) "Toward a typology of Chadic vowel and tone system." Unpublished, UCLA.

Slawson, A.W. (1968) "Vowel quality and musical timbre as functions of spectrum enevelope and fundamental frequency." *Journ. of the Acoust. Soc. of Amer.* 43: 87-101.

Snow, W.B. (1936) 'Change of pitch with loudness at low frequencies.' *Journ. of the Acoust. Soc. of Amer.* 8, 14-19.

Stevens, S.S. (1935) 'The relation of pitch to intensity.' *Journ. of the Acoust. Soc. of Amer.* 6, 150-154.

Sweet, H. (1888) *History of English Sounds.* Oxford.

Whitfield, I.G. and Evans, E.F. (1965) "Responses of auditory cortical neurons to stimuli of changing frequency." *J. Neurophysiology* 28, 655-72.

Yuan, Chia-hua et al. (1960) Han-yu fang-yen kai-yao (outling of Chinese Dialects) Peking.

Zurmühl, G. (1930) 'Abhängigkeit der Tonhöhenempfindung von der Lautstärke und ihre Beziehung zur Helmholtzschen Resonanztheorie des Hörens.' *Z. Sinnesphysiol.* 61, 40-86.

WORD GAMES: SOME IMPLICATIONS FOR ANALYSIS OF TONE AND OTHER PHONOLOGICAL PROCESSES

Jean-Marie Hombert

Department of Linguistics, University of California, Los Angeles

1. Introduction

Word games have become very popular among linguists and, in fact, have a current literature of their own. Bamgboṣe (1970) with Yoruba, Chao (1931) with Chinese dialects, Conklin (1956, 1959) with Tagalog and Hanunóo, Haas (1957, 1969) with Thai and Burmese, Hombert (1973) with Bakwiri, and Sherzer (1970) with Cuna give us an idea of the socio-cultural role of word games, the kind of manipulations which can be used, and how they can help us to solve linguistic problems.

In this paper[1], I shall first present some examples where word games can shed light on problems involved in phonological descriptions and then turn to the representation of tone, especially to the segmental vs suprasegmental nature of tone. In order to do this, I will present data from experimental work[2] done with speakers of Asian and African languages. After looking at problems of analysis introduced by these two different representations, I will attempt to define what some of the theoretical implications are. Finally, I will mention the relevance of data from existing word games and also the possibility of inventing word games as a way of gaining insight into phonological systems.

2. Linguistic use of word games

The examples presented in this paper involve a word game which consists of inverting the position of the two syllables in disyllabic words, as in (1).

(1) $C_1V_1C_2V_2 \longrightarrow C_2V_2C_1V_1$

for instance: 'death' k^wélí ⟶ lík^{w}é

The first part of this study will be illustrated with examples from Bakwiri[3], a language spoken on the southern slopes of Mount Cameroun.

2.1 Syllable structure

2.1.1. Word games can be used to define the syllable structure. Let us consider the sequence presented in (2). It is not clear to which syllable V_2 and C_2 belong. Examples presented in (2a, b and c) can suggest a solution.

(2) (C_1) V_1 (V_2) C_2 V_3

a. 'plantain' mɔ̀kɔ̀ ⟶ kɔ̀mɔ̀ BUT NOT ɔ̀mɔ̀k
b. 'throat' ìŋɔ̀ ⟶ ŋɔ̀ʔì BUT NOT ɔ̀ʔìŋ
c. 'wood' wèɔ̀lí ⟶ lììwéó[4] BUT NOT ɔ̀líwé

Considering these data, the Bakwiri syllable structure should be analyzed as (C) V (V). It is interesting to notice the appearance of a glottal stop in transformation 2b. In Bakwiri glottal stops occur only before a vowel in word initial position as shown in (3).

(3) /ìŋɔ̀/ ⟶ [ʔìŋɔ̀]

This glottal stop is preceived with difficulty in initial position but it becomes obvious when the syllables are reversed in the word game as seen in (2b). Since the vowel sequence [ɔi] is not prohibited and, in fact, does exist in certain words in Bakwiri, our example (2b), along with examples such as (4) shows the psychological reality of the initial glottal stop.

(4) a. 'salt' ìkʷá ⟶ kʷàʔí
b. 'sun' ìgʷé ⟶ gʷèʔí

The reason why the glottal stop is transposed with the vowel is to prevent the loss of a syllable. If the vowel were moved to the end without the glottal stop, the resulting unacceptable game form (5) would reduce the word to only one syllable.

(5) ìŋɔ̀ is not transformed into ŋɔ̀ì

Thus, the glottal stop provides one more piece of evidence that the unit moved by the word game is the syllable.

2.1.2 The only consonant clusters occurring in Bakwiri consist of nasals followed by homorganic voiced stops. This restricted distribution might lead us to suspect that these clusters are phonemic units. Their behavior in the word game confirms this suspicion.

(6) a. 'to take care' kɔ́mbà ⟶ mbákɔ̀[5] BUT NOT bákɔ̀m
b. 'rice' kɔ́ndì ⟶ ndíkɔ̀ BUT NOT díkɔ̀n
c. 'father' záŋgó ⟶ ŋgózá BUT NOT gózáŋ

2.2 Glides

One of the problems in the analysis of Bakwiri is establishing whether forms as in (7) are to be analyzed with an intervocalic glide or not.

(7)	a.	'stone'	lìyé	OR	lìé
	b.	'excrement'	lówá	OR	lóá
	c.	'small'	tèyí	OR	tèí
	d.	'village'	mbówà	OR	mbóà

The word game provides a reliable criterion for setting up underlying forms as seen in the left column of (8) and (9).

(8) a. lìyé ⟶ yèlí
    b. lówá ⟶ wáló

On the other hand, some forms are unable to undergo the game transformation, presumably because they have no underlying glide and therefore consist of only one syllable.

(9) a. tèí ⟶ ø
    b. mbóà ⟶ ø

It is possible sometimes to be able to check the correctness of the transformations made by the native speaker. In (10) the plural form of the item given in (7a) is presented, we see that the glide is, in fact, present.

(10) 'stones' màyé

2.3 Vowel length

If we look at (11) we can see that when a word has a long vowel, the length is not transposed with the syllable but rather stays in the same place.

(11) a. 'stomach' lùùŋgá ⟶ ŋgààlú
     b. 'burn' zééyá ⟶ yáázé

If we call S1 the first syllable and S2 the second syllable, the generalization of this transformation can be written as in (12).

(12)
$$\begin{bmatrix} \text{S1} \\ \text{+ length} \end{bmatrix} \begin{bmatrix} \text{S2} \\ \text{- length} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{S2} \\ \text{+ length} \end{bmatrix} \begin{bmatrix} \text{S1} \\ \text{- length} \end{bmatrix}$$

From the examples presented in (11) it could be argued that there is a constraint which does not allow a double vowel in word final position. If we look at (13) we see that all combinations are possible and therefore this argument does not hold.

(13)	a.	'earth'	ʔèzé 1 2	⟶	zèʔé 2 1		
	b	'bone'	ʔèèzé 1 2	⟶	zèèʔé 2 1	BUT NOT	zè ʔéé 2 1
	c.	'it is not'	é zèè 1 2	⟶	zé ʔèè 2 1	BUT NOT	zéé ʔè 2 1

In both cases (13b) and (13c) the rhythmic pattern does not change when the syllables are reversed by the game. We can now generalize as in (14).

(14) $\begin{bmatrix} \text{S1} \\ \alpha \text{ length} \end{bmatrix} \begin{bmatrix} \text{S2} \\ \beta \text{ length} \end{bmatrix} \longrightarrow \begin{bmatrix} \text{S2} \\ \alpha \text{ length} \end{bmatrix} \begin{bmatrix} \text{S1} \\ \beta \text{ length} \end{bmatrix}$

When the long or more exactly, the double vowel is formed by two different vowels, the native speaker is sometimes unable to apply the word game rules, but in other cases he can give the expected form as in (15).

(15) 'door' lìòβá ⟶ βààlíó

The problem encountered by the native speaker in applying the word game rules in this case can have two origins. Either it can be a difficulty in replacing a sequence of two different vowels V_1V_2 by a sequence of two identical vowels V_3V_3 (in the case of a word having the structure $C_1V_1V_2C_2V_3$) or more probably, the speaker feels he is violating the length pattern by putting a double vowel in a position where there was a single vowel before the transformation. All these data seem to suggest that, in Bakwiri, the length pattern is a property of the whole word and consequently is not dependent on the segment to which it is originally assigned. The same kind of evidence is found in Conklin (1956, 1959) for Tagalog as shown in (16) and for Hunanuóo as shown in (17).

(16) dooti ⟶ diito BUT NOT ditoo

(17) buuŋa ⟶ ŋaabu BUT NOT ŋabuu

3. Tone

3.1 As can be seen from the examples presented so far, as well as in (18) below the tone pattern of disyllabic words in Bakwiri is not affected by the word game transformation.

(18)	a.	Low-low	'plantain'	mɔ̀kɔ̀ ⟶ kɔ̀mɔ̀	
	b.	High-high	'death'	kʷélí ⟶ líkʷé	
	c.	Low-high	'one person'	mɔ̀kɔ́ ⟶ kɔ̀mɔ́ BUT NOT	kɔ́mɔ̀
	d.	High-low	'falling'	kʷélì ⟶ líkʷè BUT NOT	lìkʷé

As I suggested earlier for the length pattern, speakers isolate the tone pattern as a property of words, that is, *without a particular tone being attached to a particular segment*, and, in our case, *not even to a particular syllable*. In other words, data from this word game suggest that tone is a *suprasegmental feature in Bakwiri*[6].

It is interesting to note that in Thai and Burmese word games reported in Haas (1969), the reverse is found. These word games consist of interchanging the finals of two successive syllables as shown in (19) for Thai[7] and (20) for Burmese.

(19) 'big bottom' k[ôn] j[àj] ⟶ k[àj] j[ôn]

(20) 'fire place' mí bòw ⟶ mòw bí

The tones move with the transposed segments, suggesting that in Thai and Burmese, tones may be best analyzed as a *segmental feature on syllabic segments*.

3.2. In fact, one of the hotly debated issues on the status of tone is whether tone should be represented as a *segmental or a suprasegmental feature*. Schachter and Fromkin (1968) in Akan, Woo (1969), Maddieson (1971) and Fromkin (1972) argue for segmental representation. But Pike (1948), and McCawley (1964, 1970) view the syllable as the domain of the tone feature. Welmers (1962) claims that tone should be regarded as a feature on the morpmeme in Kpelle. Edmonson and Bendor-Samuel (1966) present evidence for regarding the phonological word as the tone bearing unit in Etung. More recent studies by Elimelech (1973) in Kru, by Leben (1973a) in Mende, Bambara, Maninka and Hausa, and by Mazaudon (1971) in Tamang spoken in Nepal, lead to a similar conclusion: that is in these languages tone should be regarded as a *feature on the phonological word.*

3.3. One persuasive argument for tone as a feature on the phonological word in these languages is that the number and shapes of tone patterns are not dependent on the number of vowels or syllables a given word has. This is what is shown in (21) for Kru[8] and (22) for Tamang.

(21)

	Monosyllabic			Disyllabic		
a.	'pepper'	bá	[‾]	'wine'	númó	[‾ ‾]
b.	'rice'	kɔ̀	[_]	'cup'	tàpè	[_ _]
c.	'child'	jǔ	[⁀]	'inside'	kélě	[‾ ⁄]
d.	'chicken'	sû	[⁀\ ]	'coconut'	kítâ	[‾ \]

(22)		Disyllabic			Trisyllabic		
	a.	'he went'	ni-ci	[ˋ—]	'if he goes'	ni-sami	[ˋ—˧]
	b.	'he sat'	ci-ci	[—^]	'if he stays'	ci-sami	[—¯˧]
	c.	'he pinched'	cii-ci	[—˧]	'if he pinches'	cii-sami	[——˧]
	d.	'he remembered'	cii-ci	[_/\]	'if he remembers'	cii-sami	[_¯˧]

From these data it is clear that the tone patterns on di- or trisyllabic words are just expanded versions of tone patterns on monosyllabic units for Kru and disyllabic units for Tamang, suggesting that tone is mapped over the phonological word without using segmental information.

3.4. Word games can provide important insight into this segmental vs. suprasegmental controversy. I have devised a word game for this purpose and used it with native speakers of African and Asian languages. This word game involves the two manipulations shown in (23).

(23) a. $C_1V_1C_2V_2 \longrightarrow C_1V_2C_2V_1$ (M1)

b. $C_1V_1C_2V_2 \longrightarrow C_2V_2C_1V_1$ (M2)

A disyllabic word $C_1V_1C_2V_2$ was changed into $C_1V_2C_2V_1$ (vowels interchanged) by the first manipulation M1, or into $C_2V_2C_1V_1$ (syllables interchanged) by the second manipulation M2. After Bakwiri, I experimented with speakers of the Dschang and Kru languages of Africa, and Mandarin, Cantonese, Taiwanese and Thai among East Asian languages.

I taught the word game to my informants by using the simple syllable structure CV (and not CVV or CVC) and identical tones on both syllables to illustrate the rules of vowel and syllable interchanging. The output from M2 was quite consistent both in African and Asian languages. As the data in (24) show, the tone pattern was left unchanged by African speakers performing the syllable switching manipulation.

(24) a. Bakwiri 'falling' k^wélì $\xrightarrow{M2}$ lík^wè BUT NOT lìk^wé

b. Dschang 'bone' àk^wē̄ $\xrightarrow{M2}$ k^wèā BUT NOT k^wēà

c. Kru 'axe' tǘwê $\xrightarrow{M2}$ wétû BUT NOT wêtǘ

But apart from very rare exceptions which will be discussed later, the tones were consistently moved with the syllables by the Asian speakers as shown in (25).

(25) a. Thai 'woman' phu˥˩ jiŋ˩˥ ⟶ jiŋ˩˥ phu˥˩ BUT NOT jiŋ˥˩ phu˩˥

b. Cantonese 'convenient' fɔŋ˥˧ pin˧ ⟶ pin˧ fɔŋ˥˧ BUT NOT pin˥˧ fɔŋ˧

c. Taiwanese 'meaning' i˥˧ su˨˩ ⟶ su˨˩ i˥˧ BUT NOT su˥˧ i˨˩

d. Mandarin 'at least' tɕ'i˧˥ ma˨˩˦ ⟶ ma˨˩˦ tɕ'i˧˥ BUT NOT ma˧˥ tɕ'i˨˩˦

From (26) we can see that the output from M1 is also consistent for the African languages considered. The tones are not moved when the vowels are interchanged.

(26) a. kʷélì —M1→ kʷílè BUT NOT kʷìlé
b. àkʷē —M1→ àkʷā BUT NOT ākʷà
c. tawe —M1→ tewu BUT NOT têwú

The data from the Asian languages on M1 are not so straightforward as they were for M2. In Thai, a great majority of items presented (70%) were transformed as shown in (27a), that is, the tones were moved with the segments; but some items were left unchanged as in (27b).

(27) a. 'woman' phu˥˩ jiŋ˧˥ —M1→ phiŋ˧˥ ju˥˩ BUT NOT phiŋ˧˥ ju˥˩
b. 'milk' nam˥ nom˧ —M1→ nom˥ nam˧ BUT NOT nom˧ nam˥

The output of M1 is also subject to variation in Cantonese; 50% of responses were of the form presented in (28a), with the tones being moved with the vowels, and approximately 50% were of the form presented in (28b).

(28) a. 'convenient' fɔŋ˥˧ pin˧ —M1→ fin˧ pɔŋ˥˧ BUT NOT fin˥˧ pɔŋ˧
b. 'salted vegetables' haːm˨˩ tʃɔi˧ —M1→ hɔi˨˩ tʃaːm˧
BUT NOT hɔi˧ tʃaːm˨˩

For Taiwanese and Mandarin, forms such as (29a) are very rare, and in fact, the tone pattern is almost always left unchanged after manipulation M1 (vowel interchanging game) as shown in (29b) and (30).

(29) Taiwanese

a. 'meaning' i˥˩ su˧˩ ⟶ u˧˩ si˥˩ BUT NOT u˥˩ si˧˩
b. 'below' e˧˩ kha˥ ⟶ a˧˩ khe˥ BUT NOT a˥ khe˨˩

(30) Mandarin

'at least' tɕ'i˧˥ ma˨˩ ⟶ tɕ'a˧˥ mi˨˩ BUT NOT tɕ'a˨˩ mi˧˥

3.5. Problems of analysis

3.5.1. Monosyllabicity

While one might be tempted to conclude from the data obtained with M2 that tone is a feature of the word in the African languages presented here, but a feature of a smaller unit in the Asian languages, there is one important problem which must be faced before the data from the syllable switching game can be accepted as conclusive: namely that words in Asian languages tend to be monosyllabic as opposed to the polysyllabic tendency of African languages. In order to be able to have a certain control on the bias introduced by this difference, I tried to have three different kinds of disyllabic compounds for Asian speakers. One group, exemplified in (31), contained compounds of two clearly distinct words, the second group contained compounds which were 'felt' to be one word as in (32), and the last group contained associations of words which make very little sense together as represented in (33).

(31)	a.	Thai	'not to stay'	mai ˦˧	ju ˩
	b.	Cantonese	'cat's tail'	mau ˥	mei ˧˥
	c.	Taiwanese	'red flower'	aŋ ˧	hue ˥
	d.	Mandarin	'black dog'	hei ˥	gou ˨
(32)	a.	Thai	'woman'	phu ˥˩	jiŋ ˧˥
	b.	Cantonese	'monkey'	ma: ˩˧	lau ˥
	c.	Taiwanese	'lonely'	ko ˧	tuã ˥
	d.	Mandarin	'train'	huo ˩	tʂ'e ˥
(33)	a.	Cantonese	'cat wind'	ma:u ˥	fuŋ ˦˧
	b.	Mandarin	'fire pig'	huo ˩	tʂu ˥

It was assumed that words of this last group should not be felt as one word and unexpectedly these nonsense compounds turned out to be the only exceptions we have to the general tendency for Asian speakers to move the tones with the syllables during manipulation M2.

3.5.2. Different behavior with M1 and M2

In fact, the apparently opposite results we got with M2 lead us to the same conclusion. In both types of languages, the speakers keep the tone pattern attached to the word or morpheme unit. In African languages, the tone pattern of a disyllabic word is maintained when the two syllables change position. In Asian languages, the tone is changed with the syllables precisely because each syllable is a word. These results suggest that *the word should be regarded as the tone bearing unit* in the case of these African languages: but they do not rule out the possibility that tone is a feature on a smaller unit in Thai, Cantonese, Mandarin and Taiwanese. As we mentioned earlier, the tone pattern of the African words is not modified by the vowel switching game. This was a predictable result: if tones are not interchanged when the whole syllable is moved, we should not expect them to be affected when a smaller unit (namely the vowel) is moved. We saw in (28) through (30) that variations are possible from language to language and from informant to informant. Tone is not always moved with the vowel in Thai, Cantonese, Taiwanese and Mandarin. But nevertheless, the fact that such an interchange is not attested at all with African speakers has to be accounted for. That segments can also be tone bearing units seems to be a reasonable working hypothesis. Further investigation can be done either by using the manipulation M1 (vowel switching game) and looking for consistency across speakers and across languages, or by trying to find tonological evidence showing whether tone behaves like a suprasegmental or a segmental feature. Leben (1973a) has correctly pointed out that if we assume that tone is a feature capable of being associated with units larger than segments, it should behave like other suprasegmental phenomena.

3.5.3. Level of application of word game rules.

When word games are created by linguists as a method of gaining insight into the native speaker's competence, it is very important to be able to check the level of application of the word game rules. It is obvious that our data will not have the same importance if the word game operates on the surface phonetic form as opposed to a deeper form. It is sometimes easy to know if certain rules have applied before or after the word game. Let us compare the relative ordering of tone sandhi rules and word game rules in Taiwanese. If we consider only the non-checked tones, Taiwanese has five tones as in (35).

(35)		
	High level	(˥)
	High falling	(˥˩)
	Mid	(˧)
	Low falling	(˨˩)
	Low falling rising	(˨˩˧)

When two monosyllabic words are in compound, the first tone of the compound is modified as in (36).

(36)

˥ ⟶ ˧

˥˩ ⟶ ˥

˧ ⟶ ˨˩

˨˩ ⟶ ˥˩

˨˩˧ ⟶ ˧

The tone of the second word is left unchanged. In (37), (38) and (39) three examples, with different ordering of tone sandhi and the word game rules of M2, are shown. The word presented in (37a): i ˥˩ su ˨˩ is transformed into su ˨˩ i ˥˩ . In fact the original tones are low falling-low falling and the first tone (low falling) has been transformed into high falling by tone sandhi. Then, if we proceed as proposed in (37b), we will get the incorrect output: su ˥˩ i ˨˩ (high falling-low falling). But rather, if we consider that the word game rules apply on the surface form only *after the tone sandhi rule* without having access to the underlying form, then we will get the correct derivation presented in (37c).

(37) a. 'meaning' i ˥˩ su ˨˩ from i ˨˩ su ˨˩
i ˨˩ su ˨˩ —M2→ su ˨˩ i ˥˩

b. underlying form - -> word game - - tone sandhi - -> surface form
(U.F.) (W.G.) (T.S.) (S.F.)
i ˨˩ su ˨˩ - - -> su ˨˩ i ˨˩ - - - - - - - -> su ˥˩ i ˨˩

c. U.F. - - T.S. - -> S.F. - - -> W.G.
i ˨˩ su ˨˩ - - - - - -> i ˥˩ su ˨˩ -> su ˨˩ i ˥˩

This derivation does not explain the output of M2 in (38a). We have to assume that tone sandhi has applied before and after the word game transformation changing the tone sequence from high-high falling to mid-high falling by tone sandhi, then, to high falling-mid by syllable interchanging and finally to high-mid by a new application of the tone sandhi rule.

(38). a. 'flower' hue ˧ lui ˥˩ from hue ˥ lui ˥˩
hue ˧ lui ˥˩ —M2→ lui ˥ hue ˧

b. U.F. - - T.S. - -> S.F. - - - -> W.G.
hue ˥ lui ˥˩ - - -> hue ˧ lui ˥˩ - -> lui ˥˩ hue ˧

c. U.F. - - T.S. - -> S.F. - -> W.G. - - T.S. - -> S.F.

The transformation presented in (39) can only be explained by considering that word game rules apply before the tone sandhi rule since a sequence low falling-low falling, which is in fact a sequence mid-low falling prior to tone sandhi, is changed to low falling-mid by syllable interchanging and then to high falling-mid by tone sandhi.

(39) 'old house' ku ˨˩ tshu ˨˩ from ku ˧ tshu ˨˩
ku ˨˩ tshu ˨˩ —M2→ tshu ˥˩ ku ˧

U.F. - - -> W.G. - - T.S. - - -> S.F.
ku ˧ tshu ˨˩ - -> tshu ˨˩ ku ˧ - -> tshu ˥˩ ku ˧

3.6. Theoretical implications of suprasegmental representation

While most of the theoretically oriented work has sought to treat tone as a feature on segments, Leben (1973a) proposes that a new matrix containing suprasegmental information should be added to the segmental matrix. Since as I have argued, speakers of some languages at least, strive to maintain a constant tone pattern on words, this proposal appears to have some psychological validity. In addition, Bamgboṣe (1970) has shown that Yoruba speakers can invert a tone pattern for poetic purposes, which suggests that they have access to this suprasegmental matrix. Examples of such transformation are given in (40).

(40) a. 'open place' páárápōngbá ⟶ pààràpōngbà

b. 'argument' āpérémōpétē ⟶ āpèrèmōpètē

4. Conclusion

I have attempted to show that word games can be very useful for gaining insight into phonological systems. We have seen how they can provide information concerning syllable structure, glides and vowel length, as well as providing evidence concerning the underlying representation of tone. Information from such games is valuable when consistency across related languages and across speakers is obtained. This study using only one informant for each language (except for Cantonese and Taiwanese where two informants were used) cannot pretend to lead to indisputable claims but I do think that word games can provide an access to the native speaker's knowledge of his language. It is hoped that both existing and experimentally contrived word games will be further studied from this point of view.

In conclusion, I would like to mention a remark made by Hector Javkin: "Phonology has often consisted of the subject being serious and the linguist playing games. The time has come to reverse the process".

Acknowledgments

I would like to express my deep gratitude to my informants for their great patience and willingness to subject their linguistic knowledge to my investigation and experimentation. I would also like to thank Steve Baron, Bill Ewan, Larry Hyman, Hector Javkin, Boyd Michailovsky, John Ohala and the members of the tone seminar at UCLA for their valuable comments on this paper. Thanks also to Baruch Elimelech and Martine Mazaudon for the use of their data. Needless to say, I alone am responsible for the content of the paper.

Footnotes

1. This is an expanded version of a paper presented at the Linguistic Society of America meeting at San Diego, December 1973. The first part of this paper was read at the 1973 Summer Linguistics Conference at Santa Cruz and published in *Studies in African Linguistics* vol. 4, 3, pp. 227-236.

2. The experimental work involved in this study was done at the Phonology Laboratory, University of California, Berkeley and supported by an NSF grant (GS-2286).

3. Bakwiri belongs to the Duala group (Guthrie's A-22). In this study the sign ——→ represents the transformation operated by the word game.

4. The reason why the vowel /i/ is lengthened will be discussed later.

5. The nasalization and denasalization process is explained in greater detail in Hombert (1973).

6. Exactly the same kind of behavior for tone and length patterns is found in a word game played by young Sanga speakers as described by Coupez (1969).

'grandfather' nkàɔ́mbò ——→ mbɔ̀ɔ́nkà

7. According to J. Gandour (personal communication) dialect specific and even speaker specific variations occur with respect to treatment of tone in the Thai word game.

8. This is not the situation found in all Kru dialects (W. Welmers, personal communication).

Bibliography

Bamgboṣe, A. 1970. 'Word play in Yoruba poetry' *International Journal of American Linguistics* 36,2:110-116.

Burling,R. 1970. *Man's many voices* New York: Rinehart and Winston.

Chao, Y. R. 1931. 'Eight varieties of secret language based on the principle of Fanch'ieh' *Academia Sinica* 11,2:310-354.

Conklin, H.C. 1956. 'Tagalog speech disguise' *Language* 32:136-139.

Conklin, H.C. 1959. 'Linguistic Play in its Cultural Context' *Language* 35:631-636.

Coupez, A. 1969. 'Une leçon de linguistique' *Africa-Tervuren* XV,2:1-5.

Edmonson, T. and J. T. Bender-Samuel. 1966. 'Tone patterns of Etung' *Journal of African Languages* 5:1-6.

Elimelech, B. 1973. 'On the reality of underlying contour tones' Paper presented at the Linguistic Socity of America meeting, Dec. San Diego.

Haas, M. R. 1957. 'Thai word games' *Journal of American Folklore* 70:173-175.

Haas, M. R. 1969. 'Burmese disguised speech' *Bulletin of the Institute of History and Philology, Academia Sinica* 39:277-285.

Hombert, J. M. 1973. 'Speaking backwards in Bakwiri' *Studies in African Linguistics* 4,3:227-236.

Leben, W. 1971a. 'The morphophonemics of tone in Hausa' Kim, Chin-Wu and Herbert Stahlke, (eds), *Papers in African Linguistics* pp.201-208. Edmonton, Alberta: Linguistic Research, Inc.

Leben, W. 1971b. 'Suprasegmental and segmental representation of tone' *Studies in African Linguistics* 2:183-200.

Leben, W. 1973a. 'The role of tone in segmental phonology' In Hyman, L. (ed.) *Consonant Types and Tone*

Leben, W. 1973b. 'Suprasegmental phonology' MIT Doctoral dissertation.

Maddieson, I. 1971. 'The inventory of features' *Tone in Generative Phonology* Research notes 3, University of Ibadan.

Mazaudon, M. 1972. 'Consonantal mutation and tonal split in six Himalayan dialects of Tibeto-Burmese' Paper presented at the International Seminar on Anthropological Linguistics in Patiala, Punjab, India. 11-14 October 1972.

McCawley, J. 1964. 'What is a tone language?' Paper presented at the Summer meeting of the Linguistic Society of America.

McCawley, J. 1970. 'Some tonal systems that come close to being pitch accent systems but don't quite make it' In papers from the *Regional Meeting of the Chicago Linguistic Society* pp.526-532. Chicago: Chicago Linguistic Society, University of Chicago.

Pike, K. 1948. *Tone Languages* Ann Arbor, Michigan.

Price, R. and S. Price. 1973. 'Secret play languages in Samkara: linguistic disguise in Carribean creole' In Kirshenblatt-Gimblett, B. (ed.) *Speech Play on Display* The Hague: Mouton.

Schachter, P. and V. Fromkin. 1968. 'Phonology of Akan' *Working Papers in Phonetics* 9. UCLA.

Sherzer, J. 1970. 'Talking backwards in Cuna: the sociological descriptions' *Southwestern Journal of Anthropology* 26:343-353.

Stewart, J. 1964. 'The typology of the Twi tone system, with comments by Paul Schachter and William Welmers' Legon University of Ghana, Institute of African Studies.

Surintramont, A. 1973. 'Some aspects of underlying syllable structure structure in Thai: evidence from Khamphuan - a Thai word game. *Studies in the Linguistic Sciences* 3,1:121-142.

Welmers, W. 1962. 'The Phonology of Kpelle' *Journal of African Languages* 1:69-93.

Woo, N. 1969. 'Prosody and phonology' MIT Doctoral dissertation. Reproduced by the *Indiana University Linguistics Club* (1972).

Contextual Factors Influencing Tone Discrimination[1]

J. M. Hombert and S. Greenberg

[Department of Linguistics

University of California, Los Angeles]

Introduction

Tone languages do not exploit the various possibilities of tonal contrasts with equal frequency. For example, falling tones have phonemic status in many tone languages while rising contours are relatively rare. Such disparities in the distribution of tonal patterns may be caused in part by articulatory constraints imposed by the mechanics of laryngeal displacement (Ohala and Ewan, 1973; Sundberg, 1973).

Additional constraints may be imposed by limitations on the auditory system's ability to encode frequency information of a waveform whose periodicity changes constantly.

Detection and recognition of a pure tone is adversely affected upon the introduction of a second tone in temporal proximity to the first. When the interfering tone *precedes* the target, the decrease in performance is significant only if the interval separating the tones is less than 60 msec (Samoilova, 1960; Elliot, 1962). However, the presence of a tone *following* the target can interfere with tonal recognition for inter-stimulus intervals up to 250 msec (Massaro, 1972). A tone *preceding* the target exerts maximal interference when its frequency is slightly below the target's. The situation is a bit more complex when the interfering tone *follows* the target. Identification of temporal order is most severely affected by a tone lying one-sixth to one-third of an octave above the pattern frequency range (Divenyi and Hirsh, 1975). In a tone recognition paradigm, however, the degree of interference appears independent of the direction of frequency change (Massaro, 1972).

In an earlier study Hombert (1975) found that the onset of a falling tone was more accurately perceived than the onset of a rising tone in synthesized vowel stimuli. The data indicated that 1) low tones are more effective at masking higher tones than the converse; and 2) forward masking is more significant than backward masking in the perception of fundamental frequency contours. In the case of a rising tone, each frequency would be masked by the immediately preceding lower frequency. This would not occur in the case of a falling tone since the fundamental frequency descends.

The experiment reported here was designed in order to clarify the role of backward versus forward masking, on the one hand, and the role of relative frequency height on the other.

More generally, the purpose of this paper is to investigate whether interference caused by a neighboring tone (a phonomenon generally referred to as "masking" in the psychoacoustic literature) is relevant for the perception of pitch corresponding to the frequency range used by the human voice in speech and to use this knowledge to acquire new insights into the evolution of tone systems in the languages of the world.

Experimental Paradigm

Eight subjects, all native speakers of American English, participated in an experiment in which they were asked to compare the fundamental frequency of two synthesized vowels in the presence of an adjacent syllable. Each of the subjects had achieved a score of 95% or better on a pretest involving the comparison of the F_0 of two neighboring vowels.

As can be seen in Figure 1, the set of stimuli[2] was composed of 1) a synthesized VCV sequence ([ini]) on which two fundamental frequency patterns were superimposed (either a low steady-state F_0 of 119 Hz on the first vowel followed by a high steady-state F_0 of 150 Hz or a high steady-state F_0 followed by a low steady-state F_0), and 2) two sets of comparison vowels: the high set (140 Hz, 145 Hz, 148 Hz, 152 Hz, 155 Hz, 160 Hz) which were compared with the high reference vowel (F_0 = 150 Hz); the low set (109 Hz, 113 Hz, 117 Hz, 121 Hz, 127 Hz, 130 Hz) which were compared with the low reference vowel (F_0 = 119 Hz).

Subjects were asked to judge whether the comparison vowel's fundamental frequency was higher or lower than that of the reference vowel. If the comparison vowel was higher (lower) they were instructed to circle the letter "H" ("L") on their answer sheet.

These stimuli were grouped into eight conditions as presented in Figure 2: 1) The comparison vowel either followed the reference vowel (conditions 1, 2, 3, 4) or preceded it (conditions 5, 6, 7, 8); 2) The comparison vowel was either adjacent to the reference vowel (conditions 1, 2, 5, 6) or non-adjacent (conditions 3, 4, 7, 8); 3) The comparison and reference vowels were either high (conditions 1, 3, 5, 7) or low (conditions 2, 4, 6, 8).

The interval separating the comparison vowel and the disyllable was 350 msec. The subject had two seconds in which to make a response.

Each condition was composed of two identical subparts. Each subpart consisted of 66 trials (10 repetitions for each of six comparison vowels + six trials at the beginning of each subpart which were excluded from the data analysis).

The experiment was run in two sessions (Session A and Session B) separated by one or two days. Subjects were divided into four groups. The order in which the conditions were presented was counterbalanced across these four groups.

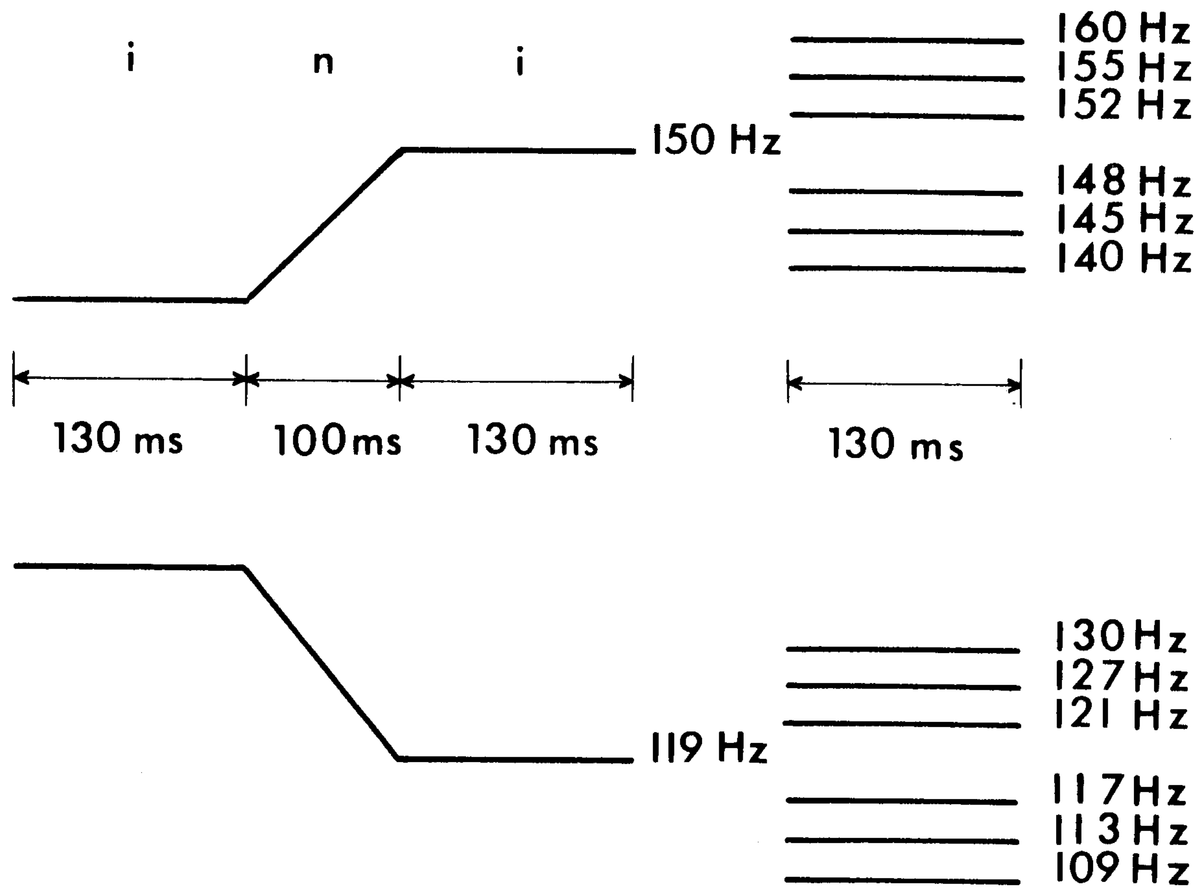


Figure 1: Stimulus configuration

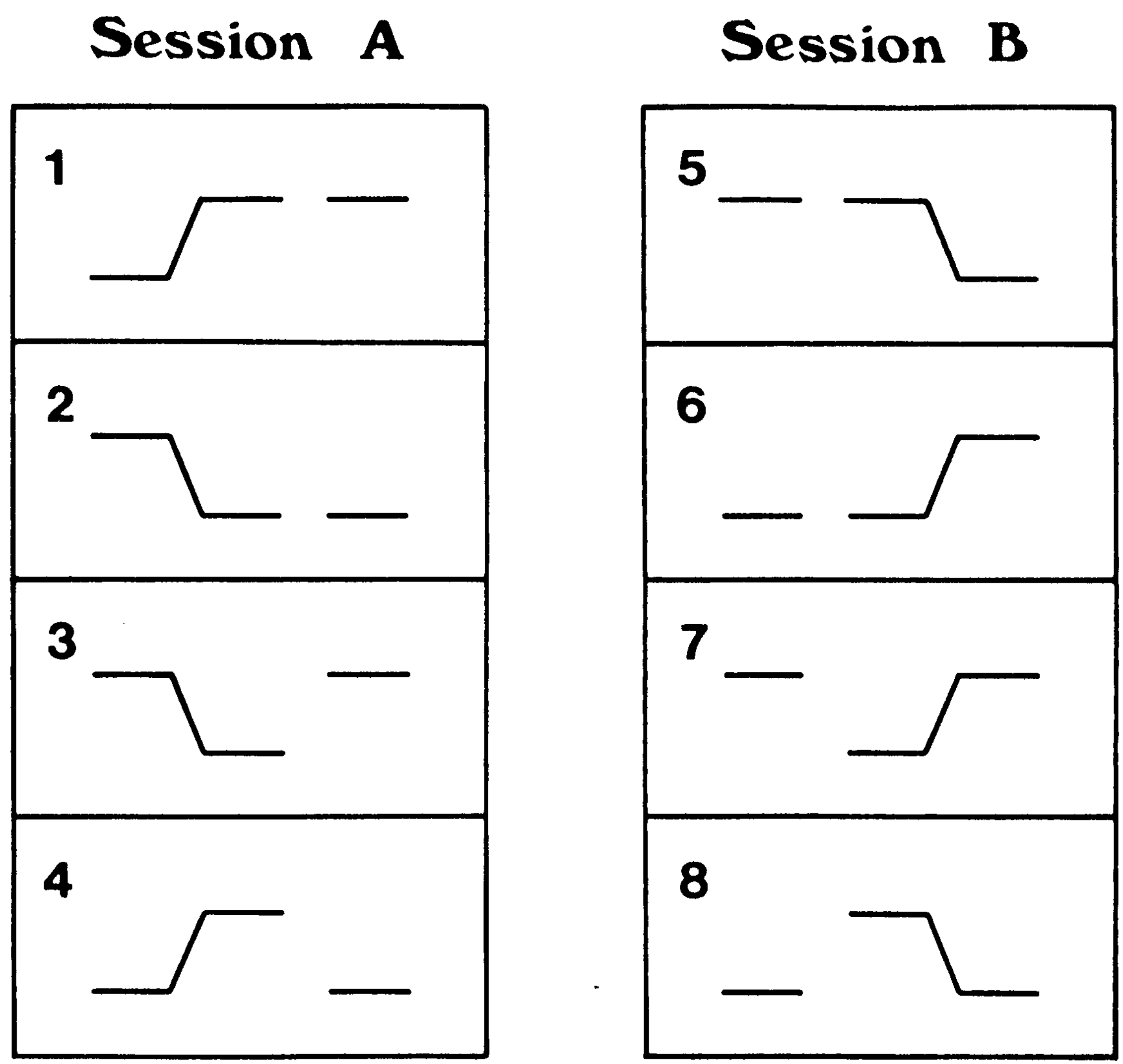


Figure 2: Experimental paradigm

Results

The results of this experiment are presented in Table 1. Our data were subjected to a four-way analysis of variance. Three parameters were within subjects: 1) Position of the comparison vowel; 2) Contiguity of the comparison and reference vowels; 3) Frequency height of the comparison and reference vowels. The fourth parameter involved the distribution of the results by group.

The d' measurements refer to a measure of the subjects' ability to judge the relative pitch of the vowels, which compensates for response bias (Green and Swets, 1966). As can be seen from the table, the position of the comparison vowel relative to the reference vowel does not affect the subjects' performance.

As expected, these results show that the comparison of two tones is easier when the comparison and reference vowels are contiguous. A more surprising result can be seen in the third row of the table. Comparison of low tones was better than comparison of high. This suggests that for our experimental paradigm, at least, low tones are perceptually more salient.

On Figure 3 performance level (the ordinate) is plotted as a function of the experimental condition (the abscissa). Subjects have been divided into two groups according to their performance. Performance values corresponding to the four best subjects (Group A) are indicated by a triangle. The performance values for the four subjects with relatively lower scores (Group B) are represented by circles. Sets of performance values (in d' units) averaged across each group of subjects are connected by a solid line.

Discussion

Analysis of the frequency distribution of errors makes it unlikely that the results are due to factors unrelated to fine frequency discrimination. For five of our eight subjects, ninety-four percent or more of the errors occurred when the fundamental frequency of the comparison vowel was either one step above or one step below that of the reference vowel.

The results are not consistent with our original hypothesis concerning the effects of backward and forward masking. A consistent pattern of interference effects is absent, both when the data is averaged and when it is analyzed by subject.

The most significant factor contributing to the observed interference involves the contiguity relationship of the comparison and reference vowels. Not surprisingly, discrimination is affected least when the compared vowels are adjacent to each other.

A second, less significant factor involves the frequency range of the compared stimuli. When the comparison involved vowels with fundamental frequencies in the lower range, discrimination is superior for all subjects in Session A and for the lower performance group in Session B.

Table 1: Experimental results

PARAMETERS		PERFORMANCE (in d′ units)	
Position of c.v.	preceding	2.79	N.S.
	following	2.87	
Contiguity of c.v. and t.v.	adjacent	2.95	$P < 0.02$
	non-adjacent	2.71	
Frequency height of c.v. and t.v.	high	2.73	$P < 0.05$
	low	2.92	

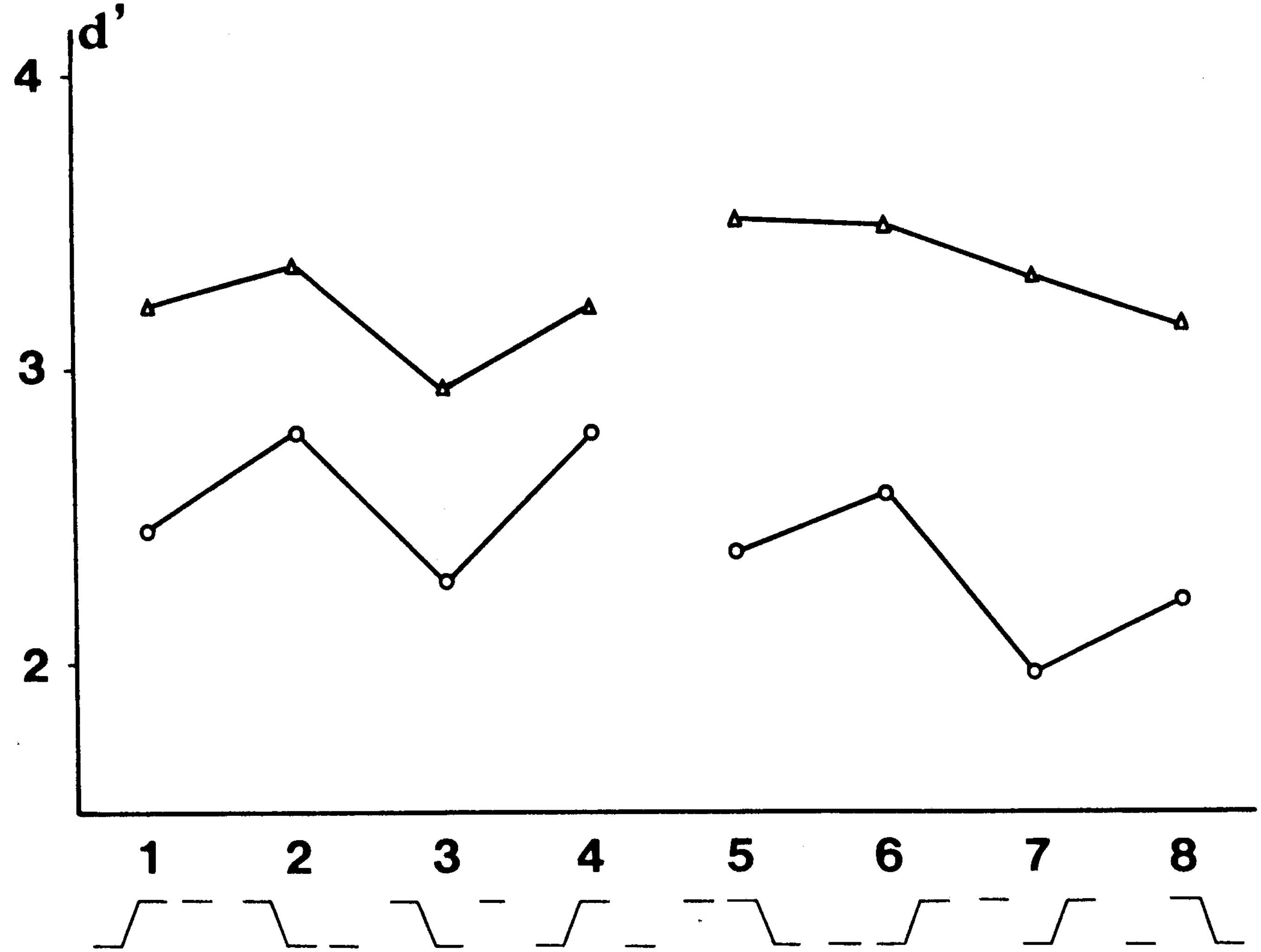


Figure 3: Performance level as a function of experimental condition

This result is somewhat surprising in view of the frequent observation of high tone prominence in many tone languages. One possible explanation could be that the violation of the normal correlation of lower intensity with low tone provides the low tone with an unaccustomed perceptual salience not found in the normal speech context. An investigation of this hypothesis is planned for the near future.

It is entirely possible that certain peculiarities of the experimental paradigm contributed to the observed effects. For instance, several subjects reported using a discrimination strategy which ignored the reference vowel in an attempt to base judgements solely upon the absolute frequency of the comparison vowel. This sort of strategy could account for the uniformly high performance level achieved by the superior performance group throughout the conditions of Session B. Randomly varying the absolute frequency of the target vowel among several possibilities within a restricted range should render this strategy ineffective. In addition, it may be advisable to increase the interval separating the comparison and reference stimuli in order to minimize possible secondary interference effects.

Conclusion

This study did not confirm our hypothesis concerning the relative importance of backward and forward interference of linguistic tones. In fact, our data appear to indicate that the findings concerning backward and forward masking are not directly relevant for stimuli having a fundamental frequency corresponding to the human voice range.

As expected, it was found that the comparison of fundamental frequency is more accurate when the two vowels are contiguous. As well, our data suggest that the perception of low tones is more accurate than the perception of relatively higher tones. This finding points indirectly to the role of amplitude information involved in the perception of tones in tone languages.

Footnotes

1. This paper was presented at the 90th Meeting of the Acoustical Society of America, San Francisco, California, November 3-7, 1975.

2. The fundamental frequencies of the comparison vowels were not in equal linear intervals because of the manner in which the OVE III synthesizer sets F_0. The OVE III generates a limited number of F_0 values. Specification of a specific frequency value provides only the synthesizer's closest approximation. We were aiming at equal spacing between our comparison vowels since pitch perception is approximately linear below 1000 Hz (Stevens et al., 1937).

Acknowledgements

The research reported in this paper was supported through NSF and USPHS grants to the UCLA Phonetics Laboratory. The stimuli were generated on the Haskins Laboratories OVE III Synthesizer. We would like to thank Terry Halwes and Cecilia Dewey for their assistance on the synthesizer. The stimuli were subsequently digitized on the UCLA Phonetics Laboratory PDP-12 Computer.

The authors thank the members of the UCLA Tone Seminar for their valuable comments. Special thanks are owed to Peter Ladefoged and Louis Goldstein for their criticism and discussion.

References

Divenyi, P. and Hirsh, I. 1975. 'The effect of blanking on the identification of temporal order in three-tone sequences' *Perception and Psychophysics* 17:246-252.

Elliot, L. L. 1962. 'Backward and forward masking of probe tones of different frequencies' *Journal of the Acoustical Society of America* 34:1116-1117.

Green, D. M. and Swets, J. A. 1966. '*Signal Detection Theory and Psychophysics*' John Wiley and Sons: New York.

Hombert, J. M. 1975. 'Perception of contour tones: an experimental investigation' *Procedings of the First Annual Meeting of the Berkeley Linguistics Society* pp. 221-232.

Massaro, D. W. 1972. 'Preperceptual images, processing time, and perceptual units in auditory perception' *Psychological Review* 79:124-145.

Ohala, J. and Ewan, W. 1973. 'Speed of pitch change' *Journal of the Acoustical Society of America* 53:345(A).

Samoilova, I. K. 1960. 'Masking of short tone signals as a function of the time interval between masked and masking sounds. *Biophysics* 4(5):44-52.

Stevens S. S., Volkmann J. and Newman, E. B. 1937. 'A scale for the measurement of the psychological magnitude pitch' *Journal of the Acoustical Society of America* 8:185-190.

Sundberg, J. 1973. 'Data on maximum speed of pitch changes' *Speech Transmission Laboratory Quarterly Progress and Status Report* (Stockholm) 4:39-47.

ON SOME CONTROVERSIAL QUESTIONS IN THE STUDY OF CONSONANT TYPES AND TONE

Larry M. Hyman
University of Southern California*

1. The recently published volume The Tone Tome: Studies on Tone from the UCLA Tone Project (Working Papers in Phonetics 27, September 1974) contains a number of papers dealing with the interaction between consonant types and tone. One of these, Ian Maddieson's "A note on tone and consonants" addresses itself to the following generalization reported by Hyman (1973) and Hyman and Schuh (1974):

(1) consonants affect tone, but tone does not affect consonants

The aim of his note is to demonstrate that "there are a number of cases which show that it [the generalization above] is false" (p.18). Maddieson claims, first, that it is just as natural for tones to affect consonants as it is for consonants to affect tone; and second, that there is no simple one-to-one correlation in this interaction because of the multiple phonetic factors involved.

In this reply I would like to focus on both of these positions. In particular, I would like to show that Maddieson's apparent counterexamples are not counterexamples at all. In so doing I hope to draw attention to what I consider to be the important questions in the study of this phenomenon.

2. Before entering into details, two general remarks are in order. Maddieson states: "Although they say they are giving an inventory of natural rules in their article, Hyman and Schuh do not explain why they think it is natural that 'consonants affect tone but tone does not affect consonants'" (p.21, his emphasis). The linguistic literature abounds with examples where a contrast in basic consonant type (e.g. voiced vs. voiceless etc.) plays a role in the tone system of a language (synchronically and/or diachronically). Many tonal contrasts owe their existence to earlier consonant contrasts (see, for example, Matisoff 1973 for a general survey of "tonogenesis" in Southeast Asia); in African languages, the subject of Schuh's and my work, tone rules often make explicit reference to the voicing state of consonants; finally, it is interesting to note that Pike (1974), whose expertise in Mexican tone languages is unequalled, concurs with the Hyman and Schuh generalization.

While such historical and synchronic interactions are numerous (I needn't give examples here), there are occasional reported phenomena where, it would appear, the identity of a tone has affected in some way the voicing state of an adjacent consonant. After conducting an extensive search of the literature, Maddieson isolates six apparent counterexamples to (1), i.e. six cases where tones appear to affect consonants (I shall turn to these below). Maddieson fails to mention, however, that there is an unmistakable statistical bias revealed by the world's languages. Although I have made no count, I know that the bias in favor of (1) holds for African languages, and specialists have expressed the same view to me (in personal communications) about Oriental and Amerindian languages. Even admitting the validity of all of Maddieson's counterexamples, it is still a fact that consonants affecting tone is vastly more frequent than is the contrary. This in itself requires an explanation. Unfortunately, Maddieson accepts each apparent counterexample at face value and does not explore the possibility of reinterpreting them in accordance with the Hyman and Schuh generalization. While Schuh and I were not aware of several of his examples when we arrived at our conclusion, I am not as much impressed by the evidence Maddieson raises against us, as I am tempted to provide alternative explanations.

Maddieson appears to be annoyed that our generalization is too "hasty": "If the <u>belief</u> that tones do not affect consonants is repeated sufficiently often, researchers will not be on the lookout for such phenomena and the <u>myth</u> will become further entrenched" (p.21, my emphasis). As just mentioned, Hyman and Schuh's "belief" (others might have called it a "claim") grew from a statistical bias which Maddieson fails to note. In looking at African languages, the interaction between consonant types and tone appeared to be unidirectional. Thus, a typical case reported in the above works is found in Nupe, where a L-H sequence becomes L-$\widehat{LH}$ (i.e. low followed by rising tone) only if the intervening consonant is voiced:

(2)	/èpá/	→	[èpá]	'is peeling'
	/èbá/	→	[èbǎ]	'is sour'
	/èwá/	→	[èwǎ]	'wants'

Maddieson is an Africanist and knows that this is common--and that the reverse, where a consonant change is in some way conditioned by a tone, is almost unattested in Africa (Ellington 1970 presents the only possible counterexample I am aware of). Thus, Hyman and Schuh's reason for coming to the conclusion in (1) should be no mystery to Maddieson.

3. Let us now turn to Maddieson's counterexamples. While I may appear to be unreasonable at first, and it may seem that I am trying to salvage (1) at any cost, the motivation for my desire to reinterpret or dismiss Maddieson's counterexamples will become apparent.

a. Maddieson mentions, citing Schachter and Fromkin (1968), that "the pitch of voiced consonants is frequently determined by the tone of neighbouring tone-bearing segments" (p.18). This is not a counterexample, however, since the "copying" of pitch from a vowel onto a neighboring consonant does not alter the consonant type in any way other than in pitch. Since what we are interested in is cases where a tone possibly affects the *seg-mental* characteristics of a consonant, this example is irrelevant. Only if some dialect of Akan devoiced [b] when followed by high tone, for instance, would Hyman and Schuh have to yield on this point.

b. The Ohũhũ dialect of Igbo has glottal fricatives which "appear to be partially voiced before low tones" (p.18). I was unable to obtain a copy of this study to ascertain what "appears to be" means (e.g. is there instrumental data?). Let us, however, file this example with the tentative hypothesis that if the above observation is confirmed, maybe glottal consonants can be affected by tone.

c. The Jingpho case has been cited to me by several people. Let us assume the analysis which best supports Maddieson's position, namely, "...words with low tone vowels and final voiceless stops voice the final stops in a process of gemination..." (pp. 18-19). However, one fact which is not mentioned by Maddieson is that the low tone is accompanied by *breathiness* in Jingpho. Since we know that glottal states can affect other glottal states, it is possible that the voicing phenomenon cited by Maddieson is conditioned by the breathiness of this tone, rather than by its pitch.

d. The fourth possible counterexample comes from Siamese. Gandour (1974) has shown that *ph* in Siamese varies in pronunciation between [ph] and [pɦ], i.e. aspirated vs. breathy. Before a higher tone the breathy variant is more likely; before a tone with a lower onset the aspirated variant is expected. According to Maddieson, "this is an example of a rule by which an earlier onset of voicing is conditioned by occurrence of high pitch" (p. 20). However, Gandour mentions (p.113) that these data are part of a historical change from *bɦ to [ph], i.e. from breathy voice to voiceless aspirated. Thus, rather than conditioning, it would appear that certain tones *retard* this sound change. Why should this be? The answer, I believe, is as follows. Assuming proto *bɦ, it is clear that the lowering effect of *bɦ will be greater before a H tone than before a L tone, where the lowering effect of *bɦ will be less noticeable. The changes in pitch are given in (3):

(3) Proto Target

a. *bɦá [⌐] > phá [~]

b. *bɦà [_] > phà [_]

In (a) part of the acoustic signal of *bɦ is a depressing pitch effect on a following vowel, especially if the tone starts high. The change to [ph], however, would require that this acoustic cue, which establishes a perceptual strategy, be abandoned and a new one be put in its place. It is the establishment of this pre-existing strategy which inhibits the change. In (b), on the other hand, where we start with breathiness having much less of a depressing effect (and therefore with pitch being of lesser importance in the perception of bɦ than in (a)), tone does not provide a retard in the change. In just this way, then, I would like to hypothesize tone to interact with consonants. It is important to emphasize that it is not the tone/pitch which is conditioning the change. The reason for the set of facts presented by Gandour is a perceptual one and not an articulatory one. Since consonant types and tone have been argued to provide support for various articulatory features, this is an important point, whose significance will be returned to below.

e. As a fifth counterexample, Maddieson cites certain cases in the history of Chinese where consonantal changes appear to be conditioned by tones. For instance, it is well-known that Middle Chinese *b (representing the class of voiced stops) develops into [ph] before Proto-Tone I, but becomes [p] before Proto-Tones II, III and IV in many of the present day dialects. It is certainly true that these facts are often presented in such a fashion that one might conclude that they constitute a counterexample to the Hyman and Schuh generalization. However, as carefully explained to me by Steve Baron and Martine Mazaudon, Pulleyblank (whom Maddieson cites) does not view these reflexes of *b as being tonally conditioned. Rather, the difference between [ph] and [p] is due to a process of "glottal dissimilation" similar to that reported for Lahu by Matisoff (1970). While there are some questions concerning the relative chronology of the changes involved, the following accurately accounts for the different reflexes of original voiced stops:

(4) a.	*ba	>	pha			>	pha		
b.	*baʔ	>	phaʔ			>	paʔ	>	TONES
c.	*bah	>	phah			>	pah		
d.	*bat	>	phat	>	phaH	>	paH		

In the leftmost column proto forms are reconstructed with different finals corresponding to the four developing proto tones. As seen in (a), Tone I derives from an open syllable; in (b) Tone II is traced back to a final glottal stop which then probably introduced creakiness on the preceding vowel; in (c) a final *h is reconstructed which probably introduced breathiness on the preceding vowel; finally, in (d) a nonnasal stop is reconstructed. (I am ignoring the possibility of final sonorants, which may have combined with one or more of the above--though not exerting a tonal effect.) In the second column historical *b becomes [ph] in all four cases. In the third column, final stops weaken to different glottalic finals (represented by an H) and as a result, forms which were in this class were redistributed into the other classes. Finally, in the last stage of the development, [ph] is "dissimilated" to [p] in those cases where a glottal consonant later appeared in the same syllable. (Equivalently, we can speak of creakiness and breathiness conditioning the dissimilation.) This process of glottal dissimilation can also be seen as the cause of the affricate/fricative split which, as pointed out by Maddieson, appears on the surface to be conditioned by tone.

f. The sixth example comes from "many Scandinavian dialects" (p.21), where a glottal stop (e.g. Danish stød) has come to be associated with a certain accent. This is not a counterexample because: 1) tone is not affecting a pre-existent consonant; 2) the glottal stop may be an accentual phenomenon rather than a pitch phenomenon; and 3) there is no agreement that this is the correct history of stød in the first place (see Gårding 1973 for a discussion of the issues involved). Point (2) is of interest. In formulating generalizations about tone rules, Hyman and Schuh made it clear that these generalizations were not intended to cover pitch-accent languages. This restriction is legitimate because of the fundamental difference between accent and tone; also Hyman and Schuh intended to isolate one aspect of tone, namely, pitch. Where pitch has an accentual function, it may be the latter that conditions a consonant change (as in stress-accent languages). This would also apply to any attempt to use examples such as Verner's Law as a counterexample (as, for example, in Anderson 1974). In order for a spirant to become voiced, the accent could not directly precede it. In slightly different terms, we could say that the weakening process was blocked where a consonant was fortified by an accent. In African languages, weakening processes are not blocked by the presence of H tone. Thus, I conclude that it must be the accentual property of the pitch prominence which is responsible rather than the pitch itself.

4. In summary, we have seen that most of Maddieson's apparent counterexamples can be reanalyzed, or are not counterexamples at all. An important aspect of the Hyman and Schuh generalization has, however, become clearer in this dispute. Namely, it is necessary to restate this generalization as in (5):

(5) consonants affect pitch, but pitch does not affect consonants

Since we do not want the generalization to cover tones which have breathiness or creakiness associated with them, or to cover pitch-accent, where accent is combined with a pitch realization, (5) seems preferable to (1). I am, however, leaving open two possibilities. First, since I could not counter the Igbo case in section 3b, the possibility is still left open that pitch may affect glottal consonants. Since pitch changes are themselves the product of glottal mechanisms, this should not be surprising. It should be noted, though, that even if such cases are occasionally found, they are rare. The second possibility is that pitch may interact with consonants for perceptual reasons, as seen in section 3d. In Siamese I argued that high pitch retards a change from breathy voice to aspiration. The reasons for maintaining (5) in the face of such phenomena will be discussed in section 6 below.

5. Maddieson's final criticism is leveled against Hyman and Schuh's hierarchy of consonant types in terms of greatest vs. least raising/lowering effects. He finds this unacceptable because the hierarchy "makes no allowance for diversity" (p.23). Maddieson thus not only does not accept the unidirectionality explicitly incorporated into the generalization in (5), but also rejects any attempt to find consistent pitch properties associated with specific consonant types. The strictness of the Hyman and Schuh hierarchy is, however, a desirable thing--if some language has voiced obstruents raising a tone synchronically, and voiceless obstruents lowering pitch, Hyman and Schuh would claim that this is a strange state of affairs. That is, just as there are crazy rules of segmental phonology, languages sometimes get crazy tone rules (and Hyman and Schuh are not responsible for these). But Maddieson does not appear to have this in mind at all: "...because there are a number of factors involved in determining the rate of vocal cord vibration, it does not seem possible to make simple predictions about unique correlations or the direction of effect. For example, a downward movement of the larynx may slacken the vocal cords (which lowers the frequency of vibration) but would tend to increase airflow through the glottis (which raises the frequency of vibration)" (p.23).

First of all, although Maddieson's example seems on the surface to be reasonable, it should be noted that there is no evidence that slackened vocal cords are involved in associating voiced sounds with low tone. More serious, however, is the idea that the same phonetic event can either raise or lower pitch. What this means is not clear. It could mean that there is in the above example a rise and then a fall in pitch, and that either of these can be picked up and phonologized. Or it could mean that for any given token of a systematic phonetic unit there can be either a rise or a fall, depending on how these multiple factors interact. The latter makes more sense. However, if one time a speaker produces a [b] and the pitch on the following vowel is raised, and then produces another [b] and the pitch on the following vowel is lowered, then the two [b] sounds are simply not the same phonetic entity. It may be that one kind of [b] (e.g. a more fully voiced one) may gradually win out and become characteristic of the language, but how can it be possible that the same physiological gestures produce two different acoustic tendencies? Thus, this criticism against the hierarchy is unfounded. Gandour's (1974) finding that unaspirated consonants raise pitch more than aspirated consonants is interesting, and it may turn out that a modification of the hierarchy is called for. But, it should be noted, Hyman and Schuh were careful to point out that the hierarchy was both tentative and incomplete (because a lot of consonant types were left out). Hopefully with more phonetic and phonological research we will learn more about which consonants have the most effect on pitch and in which direction.

6. The study of consonant types and tone has thus revealed important generalizations about the directionality of interaction and the effect of specific consonant types. The reformulated generalization in (5) not only accounts for intrinsic variations, interference of consonant types in tone rules, and cases of tone splits and tonogenesis, but also predicts that "tonoexodus", i.e. the loss of tones (Lea 1973), will not take place in as (6):

(6)	*pá	>	pa		*bá	>	pa
	*pà	>	ba		*bà	>	ba

In (6), there is a proto tone distinction between H and L which is reinterpreted as a voicing distinction of the initial consonant. In these (unattested) changes, a voiceless consonant becomes voiced before a L tone and a voiced consonant becomes voiceless before a H tone. However, I have already drawn attention to the generalization that the pitch of a vowel will not in itself change the voicing quality of an adjacent consonant (if accompanied by an accent or by some glottal state such as creakiness or breathiness, either of these may effect changes on the neighboring consonant). Thus, although historical tone change is cyclic, with languages acquiring tones and then losing them, the two processes (tonogenesis and tonoexodus) do not owe their existence to the same phonetic facts.

It has been important to show that Maddieson's counterexamples are not valid. Since we can now reconfirm our "belief" in (5), the next step is to incorporate this generalization into a model of the larynx. Phoneticians cite three factors which play a role in pitch production: tenseness of the vocal cords, rate of airflow through the glottis, and the raising/lowering of the larynx. These different factors must now be evaluated with respect to phonological universals (as well as intrinsic variations of pitch as a function of consonant types). It would seem to me that of these three possibilities, rate of airflow best accounts for the nonoccurrence of the changes in (6). If either vocal cord tension or larynx height were basic in the correlation of consonant types and tone, we would expect that anticipatory assimilations such as would be those in (6) would take place. Under the rate of airflow hypothesis, on the other hand, the prediction would be made that a high (or low) rate of airflow could not affect a preceding segment. An appropriate analogy would be vowel lengthening before voiced consonants. Although this is very general process, we do not expect to find the reverse, i.e. consonants becoming voiced after long vowels.

I cannot at the present time explain why *àp does not become [àb], and why *áb does not become [áp]. However, there is reason to believe that rate of airflow does play a role in such potential cases as well. Ewe data (Stahlke 1971) as well as data from English (Lea 1973) suggest that both voiceless and voiced obstruents lower the pitch of a preceding vowel (see also Hyman 1973, Hyman and Schuh 1974). Thus, Stahlke analyzes Ewe as having a rule which lowers mid tone to L before /p/ and /b/, but not, for instance, before /w/, which is a sonorant.

Since vocal cord tension and larynx height cannot account for the lowering of pitch before voiceless obstruents, we turn to aerodynamic factors for an explanation. An oral obstruction causes a decreases in the rate of airflow across the glottis, which in turn causes a decrease in pitch. While the interaction between voicing and pitch seems to be adequately accounted for in terms of rate of airflow, it does seem to be the case, however, that the intrinsic pitch effects of glottal and glottalic consonants are due at least in part to vocal cord tension. Thus, the pitch raising effect of a syllable-final glottal stop and the pitch lowering effect of a syllable-final /h/, both of which are widely attested in the world's languages (see, for example, Matisoff 1973), may best be seen in this light.

What is clear, however, is that we must continue to look for phonological generalizations--which in turn should be used to evaluate alternative phonetic theories. In isolating pitch from breathiness/creakiness, and from accent, and in distinguishing between articulatory vs. perceptual interactions between consonant types and tone, I hope to have provided a number of issues whose significance can now be fully appreciated.

*Research on this paper was supported in part by a Postdoctoral Fellowship from the Miller Institute for Basic Research in Science, University of California, Berkeley. I would like to thank Jean-Marie Hombert, Steve Baron, Martine Mazaudon, Boyd Michailovsky and Graham Thurgood for discussing various aspects of this paper with me.

REFERENCES

Anderson, Stephen R. 1974. The Organization of Phonology. New York: Academic Press.

Ellington, John. 1970-1. "Tone and segmental sound change in a Bantu language". African Language Review 9.

Gandour, Jack. 1974. "Consonant types and tone in Siamese". In The Tone Tome: Studies on Tone from the UCLA Tone Project. Working Papers in Phonetics 27, 91-117. Also in Journal of Phonetics 2.

Gårding, Eva. 1973. The Scandinavian Word Accents. Working Papers #8, Phonetics Laboratory, University of Lund.

Hyman, Larry M. 1973. "The role of consonant types in natural tonal assimilations". In L. Hyman (ed.), Consonant Types and Tone, 153-179. Southern California Occasional Papers in Linguistics No. 1. University of Southern California.

Hyman, Larry M. and Russell G. Schuh. 1974. "Universals of tone rules: evidence from West Africa". Linguistic Inquiry 5.81-115.

Lea, Wayne A. 1973. "Segmental and suprasegmental influences on fundamental frequency contours". In L. Hyman (ed.), Consonant Types and Tone, 15-70. Southern California Occasional Papers in Linguistics No. 1. University of Southern California.

Maddieson, Ian. 1974. "A note on tone and consonants". In The Tone Tome: Studies on Tone from the UCLA Tone Project. Working Papers in Phonetics 27, 18-27.

Matisoff, James A. 1970. "Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case study". JAOS 90.13-44.

Matisoff, James A. 1973. "Tonogenesis in Southeast Asia". In L. Hyman (ed.), Consonant Types and Tone, 71-95. Southern California Occasional Papers in Linguistics No.1, University of Southern California.

Pike, Eunice. 1974. "A multiple stress system vs. a tone system". IJAL 40.169-175.

Schachter, Paul and Victoria Fromkin. 1968. A Phonology of Akan: Akwapem, Asante and Fante. Working Papers in Phonetics 9. University of California, Los Angeles.

Stahlke, Herbert. 1971. Topics in Ewe Phonology. Unpublished Doctoral Dissertation, University of California, Los Angeles.

Universal Rules of Tone Realization[1]

Carl R. La Velle

[Department of Linguistics, University of
California, Los Angeles]

1. Introduction[2]

According to Schachter and Fromkin (1968), Peters (1973) and others, downdrift is an intonation phenomenon which lowers the fundamental frequency of tones throughout a phrase. This is accounted for, formally, by introducing pitch assignment rules into the phonological component. At a relatively abstract level, the pitch assignment rules introduce pitch specifications in terms of numerical notation. Thus, these rules derive abstract scalar features of FO.

In the first section of this paper I shall briefly review one of these proposed solutions, and offer a simpler alternative. Primarily, this paper is concerned with the physical realization of tonal phenomena, and, in this respect, the latter part of this presentation is aimed at providing universal rules of FO assignment for tone systems.

2. A current analysis

In their treatment of downdrift, Schachter and Fromkin account for the following data:

(1)	a.	mé+ɔ̀+bó	underlying form
	b.	mé+$\grave{\emptyset}$+bó	by vowel deletion
	c.	m$\overset{5}{e}$+$\overset{2}{\emptyset}$+b$\overset{4}{o}$	by pitch assignment
	d.	m$\overset{5}{e}$+b$\overset{4}{o}$	by segment deltion

Line (b) is from (a) by vowel deletion. This rule actually deletes all features except [+Segment, Low Tone]. Line (c) is from (b) by pitch assignment. Line (d) is from (c) by a segment deletion rule.

This solution has two drawbacks. First, it requires an ad hoc unit with no features other than tone. Secondly, it suggests that pitch assignment rules are ordered relatively early in the phonological component. These shortcomings can be eliminated by positing the following rule:

(2) H → D / L ___

It is true that high (H), downstep (D), and low (L) are only cover symbols for composites of tone features. One can then assume that the distinctive feature matrices of these tones are the following:

(3)

	H	L	D
HI	+	-	+
MID	-	-	+

Rule (2) can be stated formally as:

(4) [+HI] → [+MID] / [-HI] ___

One could of course also introduce a special tonal feature [+downstep] to account for this interesting phenomenon.

3. Rules of F0 realization

The output of a grammar would then be stipulated only in terms of phonological features rather than relative F0 values. If it can be shown that there are universal rules or equations by which these linguistic tones can be realized as physical signals specified as to fundamental frequency in hertz, then an important claim is being made. That is, this suggests that grammatically, only tones, not F0 values are represented as both input and output of the grammar and that the physical realization of these tones can be determined by universal phonetic processes. Such universal rules can also be conceived of as a set of instructions to a speech synthesizer. To the extent that the synthesized speech does not simulate human speech we know that the rules are incorrect.

As a first approximation I will utilize two variables. One variable 'p' represents the set of F0 values assignable to tones. If in a two tone language, 'p_1' is the set of frequencies assignable to phrase initial high tones, then 'p_2' is the set assignable to phrase initial low tones. The other variable 'n' represents the set of F0 intervals assignable to a phrase. The values which replace the variables are determined by the individual speaker.

For a non-downdrift language with low tones, the tones can be distinguished in the grammar by one distinctive feature, say +H and -H. Then the F0 assignment rules may be stated as follows:

(5) a.

$$\text{Tone} \rightarrow \left\{ \begin{array}{lll} p1 & / & \text{PhB} \begin{bmatrix} \underline{\quad} \\ +H \end{bmatrix} \\ p2 & / & \text{PhB} \begin{bmatrix} \underline{\quad} \\ -H \end{bmatrix} \end{array} \right\}$$

PhB = phrase boundary

b.

$$\text{Tone} \rightarrow \begin{bmatrix} p \\ +n \end{bmatrix} / \begin{bmatrix} -H \\ p \end{bmatrix} \begin{bmatrix} \underline{\quad} \\ +H \end{bmatrix}$$

c.

$$\text{Tone} \rightarrow \begin{bmatrix} p \\ -n \end{bmatrix} / \begin{bmatrix} +H \\ p \end{bmatrix} \begin{bmatrix} \underline{\quad} \\ -H \end{bmatrix}$$

d.

$$\text{Tone} \rightarrow p \; / \begin{bmatrix} \beta H \\ p \end{bmatrix} \begin{bmatrix} \underline{\quad} \\ \beta H \end{bmatrix}$$

Rule (5a) provides for arbitrary FO assignment to initial tones. Rule (5b) states that the FO of high is realized as the FO of the preceding low plus the relative FO interval. Rule (5c) states that low tone is realized as the FO of the preceding high minus the relative FO interval. Rule (5d) assigns identical FO to sequences of like tones.

The sentences generated by the grammar will therefore be specified only as + or - high for each tone bearing unit, and the FO assignment rules will allow for the specification of FO in terms of hertz as in (6):

(6)

	H	L	H	L	L		
a.	130Hz					by 5a	(spk.det. p_1=130Hz)
b.		115Hz				by 5c	(spk.det. n=15Hz)
c.			130Hz			by 5b	
d.				115Hz		by 5c	
e.					115Hz	by 5d	

Throughout this paper, I will continue to employ variables in the formulation of the rules. Each variable, which represents a set of actual FO values, can be roughly viewed as having some correlate in the speech mechanism. The actual physical value of the first occurrence of a variable in each phrase is principally determined by the individual speaker. In converting the phrase initial occurrence of a variable into physical reality, the speaker is constrained by certain extra-linguistic considerations. In all derivations, I assume that, once selected, the value for each variable remains constant throughout the phrasal pitch mapping operation. Note that these FO assignment or realization rules are not ordered. It will be further shown that ordering is unnecessary for pitch realization rules for downdrift languages as well, since, as we shall see, downdrift is viewed as an assimilatory process, whereby the F0 of one segment is dependent on that of the preceding or contextual frequencies. The F0 rules can only apply to a segment when the necessary F0 has been assigned to an immediately preceding segment.

3.1 Pitch realization rules in Hausa

Let us now consider a language like Hausa, in which downdrift occurs but no downstep. Hausa has two contrasting tones which can again be specified in feature terms as +H and -H.

(7) a.
$$\text{Tone} \rightarrow \begin{Bmatrix} P1 & / & \text{PhB} \begin{bmatrix} \_\_ \\ +H \end{bmatrix} \\ P2 & / & \text{PhB} \begin{bmatrix} \_\_ \\ -H \end{bmatrix} \end{Bmatrix}$$

b.
$$\text{Tone} \rightarrow \begin{bmatrix} p \\ \\ -n \end{bmatrix} \;/\; \begin{bmatrix} +H \\ \\ p \end{bmatrix} \begin{bmatrix} \_\_ \\ \\ -H \end{bmatrix}$$

c.
$$\text{Tone} \rightarrow \begin{bmatrix} p \\ \left(\begin{bmatrix} +n \\ -z \end{bmatrix} \right) \end{bmatrix} \;/\; \begin{bmatrix} -H \\ \\ p \end{bmatrix} \begin{bmatrix} \_\_ \\ +H \end{bmatrix} \quad \begin{bmatrix} \pm H \end{bmatrix}$$

d.
$$\text{Tone} \rightarrow p \;/\; \begin{bmatrix} -H \\ \\ p \end{bmatrix} \begin{bmatrix} \_\_ \\ \\ +H \end{bmatrix} \quad \text{PhB}$$

e.
$$\text{Tone} \rightarrow \begin{bmatrix} p \\ \\ -x \end{bmatrix} \;/\; \begin{bmatrix} \beta H \\ \\ p \end{bmatrix} \begin{bmatrix} \_\_ \\ \\ \beta H \end{bmatrix}$$

Rule (7a) provides for assignment of F0 to phrase initial tones. (7b) introduces a F0 interval between high-low sequences. (7c) and (d) are formulated to account for the behavior of high tone after low. Rule (7c) allows for an optional F0 interval between low-high sequences when the high is followed by another tone. When the option is not chosen the high tone is lowered to the frequency level of the immediately preceding low tone. Rule (7d) states that a high tone assimilates completely to an immediately preceding low tone when the high tone occurs at phrase boundary. The process by which a high tone is totally assimilated to the level of a preceding low is termed 'high destruction' by Meyers (1974). In (7c) the variable 'z' is included to account for the degree of assimilatory attraction a low tone exerts on a following high tone. Its realization, which is some fractional value of 'n', is determined by the speaker. Rule (7e) provides for what Meyers calls 'intonational downdrift'. This phenomenon refers to the slight degree of lowering each member of a sequence of like tones undergoes. The variable 'x', whose value is determined by the speaker, provides for this process. The operation of these rules are illustrated in the following sample derivation:

(8)	H	H	L	H	L	H	
a.	130hz						by 7a (spk.det. p_1=130hz)
b.		125hz					by 7e (spk.det. x=5hz)
c.			110hz				by 7b (spk.det. n=15hz)
d.				120hz			by 7c (spk.det. z=5hz)
e.					105hz		by 7b
f.						105hz	by 7d

3.2 FO realization rules in languages with downstep

There are languages like Akan, Igbo, etc., where due to the deletion of underlying phonemic low tones, and given the mechanism of downdrift, a three way phonetic contrast is manifest after +H. In some of these languages, in the course of historical change, the downstep produced by downdrift and low tone vowel deletion has become 'phonologized' (cf. Hyman). However, in these languages some occurrences of downstep may still be predictable given the phonological rules of the languae, while others must be underlying.

Let us consider a language in which all occurrences of downstep result from vowel deletion rules. Given a phonetic tonal sequence like

(9) H L H L H,

and if downdrift is seen as an assimilatory process, after everylow a high will be lowered by that preceding low. Using the downdrift rule posited earlier (given under (2) and (4)), the string given in (10) would result.

(10) H L D L D

If, in addition, the final low tone vowel is deleted we would derive (11)

(11)	H	L	D	D
HI	+	-	+	+
MID	-	-	+	+

The FO realization rules for such a language can now be given as follows:

(12) a.

$$\text{Tone} \rightarrow \left\{ \begin{array}{lll} p_1 & / & \text{PhB} \begin{bmatrix} \overline{} \\ +H \end{bmatrix} \\ p_2 & / & \text{PhB} \begin{bmatrix} \overline{} \\ -H \end{bmatrix} \end{array} \right\}$$

b.

$$\text{Tone} \rightarrow \begin{bmatrix} p \\ \\ -n \end{bmatrix} \; / \; \begin{bmatrix} +H \\ \\ p \end{bmatrix} \begin{bmatrix} \underline{} \\ \\ -H \end{bmatrix}$$

c.

$$\text{Tone} \rightarrow \begin{bmatrix} p \\ +n \\ -z \end{bmatrix} \; / \; \begin{bmatrix} -H \\ \\ p \end{bmatrix} \begin{bmatrix} \underline{} \\ +H \\ +M \end{bmatrix}$$

d.

$$\text{Tone} \rightarrow p \; / \; \begin{bmatrix} \alpha H \\ \beta M \\ p \end{bmatrix} \begin{bmatrix} \underline{} \\ \alpha H \\ \beta M \end{bmatrix}$$

e.

$$\text{Tone} \rightarrow \begin{bmatrix} p \\ \\ -z \end{bmatrix} \; / \; \begin{bmatrix} +H \\ \\ p \end{bmatrix} \begin{bmatrix} \underline{} \\ \\ +M \end{bmatrix}$$

Rule (12a) provides for an interpretation of phrase initial tones with a different starting F0 for high and low tones. Rule (12b) introduces a F0 interval between high-low sequences. Rule (12c) interprets downstep tone when preceded by a low tone. As in rule (7c), the variable 'z' in (12c) functions to compress the F0 interval between low-high sequences. (12d) assignes identical F0 to a sequence of contiguous high or low tones. (12e) lowers a downstep tone or high tone by 'z' hertz. These rules are illustrated in (13).

(13)	H	L	H	L	H	
a.			D		D	by 2
b.				Ø		by vowel deletion
c.	$\begin{bmatrix} +H \\ -M \end{bmatrix}$	$\begin{bmatrix} -H \\ -M \end{bmatrix}$	$\begin{bmatrix} +H \\ +M \end{bmatrix}$		$\begin{bmatrix} +H \\ +M \end{bmatrix}$	input to the pitch rules
d.	130hz					by 12a (spk.det. p_1=130hz)
e.		115hz				by 12b (spk.det. n=15hz)
f.			125hz			by 12c (spk.det. z=5hz)
g.					120hz	by 12e

4. Contours and downdrift in Yoruba

In this paper I have not yet considered rules of contour F0 realization. In consideration of this surface phenomenon, I shall turn to Yoruba, a West African Kwa language with three level tones. In this language phonemic low after high is always realized as a phonetic fall from the level of the preceding high, and phonemic high after low is always a phonetic rise from the level of the preceding low.

(14) a. L H → L $\$_{LH}\$$

b. H L → H $\$_{HL}\$$ $\$$ = syllable boundary

As shown in (14), a rule of tone spread converts low-high and high-low tonal sequences to low low-high and high high-low respectively. This well motivated process is considered the phonological source of the Yoruba contours.

It is specifically claimed that in respect to downdrift, the end points of rising and falling contours function as high and low tones respectively. In support of this claim, it is pointed out that the end points of the rise and fall are always lower than their respective level counterparts which precede. Such observations are in accord with the downdrift phenomenon. Courtenay (1971) claims that Yoruba utterances also reflect the operation of downdrift. Assuming the correctness of Courtenay's observation, let us examine the following Yoruba data:

(15)[3]	a. ɛ̀búté (1 2)	1. 100 → 120Hz	2. 143 → 141Hz
	b. ɑ̀kpótí (1 2)	1. 103 → 124"	2. 141 → 140"
	c. ìšɛ́dʒú (1 2)	1. 98 → 120"	2. 140 → 143"
	d. ...wɑ̀ níbí... (1 2)	1. 89 → 105"	2. 137 → 140"

In (15) there are two columns to the right of the phonetic transcriptions. Each column contains four rows of figures corresponding to the indicated phonetic segment, and each phonetic segment is represented by two F0 values. The figure to the left of the arrow represents the approximate starting F0 of the indicated tone bearing segment, and the figure to the right of the arrow represents its approximate final F0. If, in (15), the end point of the Yoruba rise is functioning as a high tone physically, then the following level high tone violates an important phonetic consequence of downdrift.

(16) Once a tone has been lowered in an utterance, a following identical tone cannot occur higher than a similar preceding tone.

If it is agreed that downdrift is the cause of the lowered rise, then I believe that the only reasonable alternative is to posit 'upstep' to account for the following high tone. In my opinions, such an alternative is poorly motivated.

The evidence in (17), which is the phonetic inverse of (15), is interesting in another respect. In these data notice that the fall also fails to reach the frequency level of the following tone which is low.

(17)	a.	...fṹ ìyɑ̀wó	1.	120	→	90hz	2.	74	→	74hz
	b.	...ɑ̀ɑ́rɔ̀ ɔ̀lɑ	1.	142	→	100"	2.	80	→	83"
	c.	...iléìkɑ̀wé	1.	168	→	147"	2.	95	→	87"
	d.	...lɑ́sìkò yɛ̃	1.	140	→	127"	2.	98	→	90"
	e.	...iléìkɑ̀wé	1.	160	→	130"	2.	70	→	71"
	f.	kílèdèbɑ̀dɑ̃̀	1.	124	→	102"	2.	80	→	70"

Ohala and Ewan (1973) report that for a given pitch interval a subject can execute a fall faster than a rise. They suggest that more physiological effort is required for the production of rising F0 as compared to falling F0.

5. Realization rules for contour pitch

One expectation which follows from the proposal by Ohala and Ewan, is that at least the end point of a fall should easily reach the level of a following low tone, if the fall functions phonetically as a low tone. In my opinion then, to explain a physical fall as a function of two physical levels becomes quite unacceptable upon consideration of the data in (17). A plausible counter-proposal which adequately accounts for the behavioral similarity of rising and falling is:

(18) a. Physical contours in Yoruba are identical to similar contours which are underlying and segmental in other languages.

b. Physical contours tend to assimilate to their points of origin.

If part (a) of this proposal is valid then its acceptance implies the exclusion of pitch intervals on minimal units at the physical level in Yoruba and similar languages. Furthermore, if the domain of tonal features is the segment, then it is likely that this natural physical constraint pervades the entire phonology. Under this proposal, (19) shows how Yoruba can introduce contours:

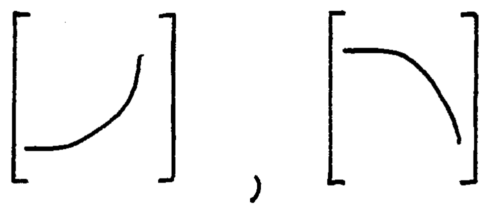
(19)[4,5] a. $L\ H \rightarrow L\ \widehat{LH}$ L LR or L R

b. $H\ L \rightarrow H\ \widehat{HL}$ H HF or H F

If the full implications of the hypothesis in (18) are pursued, I expect that they will at least shed some light on the questions concerning the theory of tonal representation. I will not pursue those consequences in this paper.

According to my observations of instrumental data, the FO shapes during contour tones is as shown in (21).

(21)



The significance of these curves is that the begin point carries a sustained FO, but not the end points. The generalization stated in (18b) is missed by those linguists who insist on maintaining that the physical endpoints of Yoruba contours function as level tones.

6. Conclusion

In this paper I first considered an analysis of a tonal problem which seems questionable in that it requires the maintenance of an ad hoc tone bearing unit, and pitch assignment rules early in the phonological component. An alternative was proposed which resulted in a simpler description.

In the latter part of this paper an attempt was made to formulate rules which describe or provide for the conversion of abstract features into physical reality. As such these rules can be conceived as rules of speech synthesis. It is assumed that if the rules are stated correctly, they can apply to all such languages universally and need not be stated for each language. All that need be stated in the grammar of each language is the type of tone language it is.

Finally, in the attempt to account for Yoruba contours, I offered a more explicit definition which may lead to a greater understanding of this phenomenon at both the abstract and phy ical levels of representation. One should note that my characterization of contours precludes the necessity to prevent such phenomena from appearing in either phonological rules or in underlying representations on the grounds that such representations violate theoretical constraints.

Footnotes

1. This paper was presented at the Sixth Annual Conference on African Linguistics at The Ohio State University, April 11-13, 1975.

2. I am greatly indebted to Vicki Fromkin for her guidance and constructive criticisms throughout the development of this paper. I would also like to thank Peter Ladefoged and Paul Schachter for their patience.

3. The notation '...', refers to the fact that the data was taken from a larger utterance.

4. The symbol '⌒' means that both tones occur on a single segment.

5. $\widehat{LH}$ → LR or R[_/] and $\widehat{HL}$ → HF or F[¯\ ] may be universal generalizations about tone languages. However, it is an empirical question because it is quite conceivable that there are languages which have the rules: $\widehat{LH}$ → RH[/¯] and $\widehat{HL}$ → FL[_].

References

Courtenay, K. R. 1971. 'Yoruba: a 'terraced-level' language with three tonemes' *Studies in African Linguistics* 2:239-255.

Hyman, L. M. 1972. 'Nasals and nasalization in Kwa' *Studies in African Linguistics* 3:167-206.

Meyers, Laura. 1974. 'Tone patterns in Hausa: a re-analysis of Hausa downdrift' *Working Papers in Phonetics* No. 27. Los Angeles: University of California Press.

Ohala, J. and W. Ewan. 1973. 'Speed of pitch change' (Abstract) *Journal of the Acoustical Society of America* 53:345.

Peters, A. M. 1973. 'A new formalization of downdrift' *Studies in African Linguistics* 4:139-153.

Schachter, P., and V. Fromkin. 1968. 'A phonology of Akan' *Working Papers in Phonetics* No. 9. Los Angeles: University of California Press.

The Acquisition of Tone in Mandarin-Speaking Children*

Charles N. Li
UCSB

Sandra A. Thompson
UCLA

1. Background

In spite of the extensive literature now available on children's acquisition of phonology, there exists very little data on the acquisition of tones by children learning a language with phonemic tone distinctions.

There is a substantial body of literature on the acquisition of suprasegmental aspects of language, i.e., stress, intonation, duration, pitch, etc., which suggests that these aspects of language are acquired relatively early. For example, there are a number of studies suggesting that distinctive intonation patterns can be detected in the young infant's early vocal behavior. Kaplan (1970) shows that normal falling and rising intonation contours in English can be discriminated by eight-month old children. Lenneberg (1967) cites complementary evidence suggesting that an infant's produced intonation patterns become distinct at eight months. The early use of prosodic features is also reported by Weir (1966) and Atkinson-King (1973) (also see references cited in these two studies). Crystal (1970) cites studies demonstrating that children from 7-10 months have "primitive lexical items" with a "suprasegmental character" which is more stable and more readily elicited than the segmental character (p. 80). Kaplan and Kaplan (1970) report that most of the literature suggests that at about 5-6 months, the child is primarily processing suprasegmental aspects of speech, while Crsytal (1973) states that most observers conclude that 6-7 months is the most likely period for the emergence of such features in production. From these studies, Bever, Fodor, and Weksel (1965) conclude that "it is widely accepted in the literature that the child effectively masters the intonation pattern of his language before he has learned any words at all" (p. 275).

As far as the acquisition of tone is concerned, however, much less is known and the evidence is much less clear about early acquisition. Chao (1951), in his pioneering study describing the Mandarin of his granddaughter Canta during her 28th month, claims that

> "Canta acquired tones very early, as most Chinese children do. Isolated tones of stressed syllables are practically the same as in standard Mandarin". (p. 32)

However, he does not say how early tones are acquired, nor does he give any specific data. He does mention (p. 27) that records from the time Canta was about 18 months old showed that she had

> "one instead of two rising tones, so that [hɛ] high-rising was either shye 'shoe' or shoei 'water', which are now distinct in her speech".

Interestingly enough, however, though Canta may have xié 'shoe' and shuǐ 'water' with the correct rising and dipping tones respectively, in the Cantian lexicon which he provides, out of 51 rising tone words, 15 of them are transcribed as dipping tone words in Canta's speech, though only one dipping tone word is transcribed as a rising tone word. This suggests that there is still confusion for Canta between rising and dipping tones, a finding which is strongly supported by our data. But it also suggests that the errors are overwhelmingly in the direction of substituting dipping tones for rising tones, a finding which is not supported by our data. Chao also reports that the "tone sandhi" rule of Mandarin which raises a dipping tone to a rising tone before another dipping tone is only beginning to be learned by Canta.

E. Pike (1949) reports that Mixteco-speaking children learn lexical tones as they learn the segmental items in their language and that their first words show correct lexical pitch. No systematic data are cited to support her impression; in fact, the data we have collected suggest that it is not universally valid. Kirk (1973) notes that in the speech of four Gã children between 2 and 2-1/2 years of age, 13% of their distortions in imitation are tonal. Apronti (1969) shows that one 2-year-old Dangme child makes no tonal mistakes and can correctly apply tone sandhi rules.

One recent study of stress (Atkinson-King [1973]) suggests that, although stress in individual lexical items may be learned as early as 2 (i.e., children do not typically produce [teybə́l] for [téybəl]), the use of stress patterns to distinguish between grammatical patterns such as nominal compounds (like hót dog) and adjective-noun phrases (like hot dóg), may not be acquired until much later, somewhere between the ages of 7 and 12 years of age.

These studies suggest, then, that tone acquisition takes place relatively early. In order to satisfactorily study this hypothesis, however, it is clear that what was needed was an in-depth longitudinal study with a number of children. We carried out such a study in Taiwan in 1974-75 which provides for the first time answers to the following questions:

a. In what ways do the child's utterances deviate from the adult's with respect to tone during the acquisition process?

b. What is the duration of the tone acquisition process?

c. What is the chronological relationship between the acquisition of tones and the acquisition of segmental units?

d. What explanations can be offered for the patterns of errors which are found and for the relative speed with which tone seems to be acquired?

And, specific to Mandarin, we might include:

e. What is the nature of the rising-dipping tone "confusion" mentioned by Chao?

f. What are the facts about the acquisition of the "tone sandhi" rule?

g. What is involved in the acquisition of the "neutral tone" in Mandarin?

The data we have came from 17 children of Mandarin-speaking families in Taipei, Taiwan. The children ranged in age between 18 and 36 months. Longitudinal data over a period of about 7 months were collected for 10 of these children, and more sporadic data were collected for the other 7. We visited each child roughly every 3 weeks. The procedure was to engage the child in looking at picture books. The adults on hand would point to the picture, ask the child what it was, and try to elicit a naming response from the child. The importance of using the pictures lies in the control they provide on the possible responses the child could have been giving. After each child utterances the adult would repeat what the child had said. If the child did not respond, the adult would often supply the answer, which the child might then imitate. All sessions were taped, using a Uher-134 cassette recorder and an AKG-109 microphone, and were later transcribed by the two of us. A total of about 80 hours of data were collected.

Before reporting our results, we would like to briefly present the Mandarin tone system.

2. The Mandarin tone system

There are four tone phonemes: high level, rising, low, and falling. There is, in addition, a "neutral" tone which occurs on unstressed syllables, suffixes, particles, and the second syllables of many disyllabic words.

Chao (1930) provided a way of representing these tones by dividing the pitch range into five levels, numbered 1 to 5, where 1 corresponds to low, 3 to mid, and 5 to high. A tone can then be represented by sequences of numbers. He also provided a graphic representation by placing a tone contour line corresponding to the pitch range of the speaking voice. These two representations can be displayed together as follows:

(1)	high tone	˥	55
(2)	rising tone	˧˥	35
(3)	dipping tone	˨˩˦	214
(4)	falling tone	˥˩	51

This traditional understanding of Mandarin citation from tones is confirmed by acoustic measurements (see Brotzman [1964], Wang, and Li [1967], and Dreher and Lee [1966]).

The "neutral" tone occures when a syllable is unstressed. It is predictable and depends for its phonetic pitch primarily on the tone of the preceding syllable. The relative pitch of the neutral tone after each of the four citation forms in Mandarin can be illustrated as follows (after Chao [1968]):

half-low	˨	after high tone
middle	˧	after rising tone
half-high	˦	after dipping tone
low	˩	after falling tone

As mentioned above, certain suffixes and particles always carry the neutral tone; the second syllables of many compound words are also in the neutral tone. Other syllables may be reduced in stress during speech and acquire a neutral tone as a result of their position in a given syntactic construction (see Woo [1969] and Cheng [1971] for discussion).

In the _pinyin_ romanization system, the Mandarin tones are represented by diacritics as follows:

high:	_bā_	'eight'
rising:	_bá_	'to pull'
low:	_bǎ_	object marker
falling:	_bà_	'a harrow'
neutral:	_ba_	attitude particle

There are basically two tone sandhi rules, both of which affect the dipping tone. As illustrated in (1), the first raises the dipping tone to a rising tone before another dipping tone. The rising tone which results from the application of this rule has been shown experimentally to be indistinguishable from an inherent rising tone (see Wang and Li [1967]). The other tone sandhi rule changes the dipping tone to a low level tone before a high, rising, or falling tone. Notice that the effect if (1b) is essentially to shorten and simplify the complex dipping tone by removing the fall at the beginning and the rise at the end of the vocalic segment.

(1) Tone sandhi rules:

a. ˨˩˦ → ˧˥ before ˨˩˦

b. ˨˩˦ → ˩ before { ˥ , ˧˥ , ˥˩ }

Significant for the interpretation of our results is the fact that the only environment in which the dipping tone has its full dipping properties is in isolation or utterance-final position.

3. Results

To provide a framework within which to discuss our results, let us delineate four stages in the tone acquisition process.

(2) Stages of tone acquisition:

<u>Stage 1</u>: The child's vocabulary is small. High and falling tones predominate irrespective of the tone of the adult form.

<u>Stage 2:</u> The child is still at the one-word stage, but he has a larger vocabulary. The correct 4-way adult tone contrast has appeared, but sometimes there is confusion between rising and dipping tone words.

<u>Stage 3</u>: The child is at the 2-3 word stage. Some rising and dipping tone errors remain. Tone sandhi is beginning to be acquired.

<u>Stage 4</u>: Longer sentences are being produced. Rising and dipping tone errors are practically non-existent.

The results of our study may be summarized as in (3):

(3) Results:

a. The correct tone system is acquired relatively quickly and is mastered well in advance of the segmental system.

b. The high and falling tones are acquired earlier and more easily than the rising and dipping tones.

c. Confusion persists in the form of substitution errors throughout stages 2 and 3 between the rising and dipping tones.

d. The tone sandhi rules are learned, with infrequent errors, as soon as the child begins to produce his own multi-word utterances.

e. The neutral tone is often misinterpreted as a full tone.

We would now like to discuss each of these results in greater detail.

a. Tones vs. segmentals

Our data provide rich support for the claim that the tonal system is mastered well in advance of the segmental system. Examples abound from all stages in which the child produces a phrase or word with the appropriate tone pattern which is segmentally very far from the adult form and often downright incoherent. (4) gives some typical examples from each of the four stages.

(4)

child		adult		
Stage 1				
dədə	[¯ ¯]	čəčə	[¯ ¯]	'car'
fɔdi	[¯ ¯]	feyǰi	[¯ ¯]	'plane'
Stage 2				
ipi	[¯ ˇ]	čyɛnbi	[¯ ˇ]	'pen'
dziyu	[¯ ʹ]	ǰünrən	[¯ ʹ]	'soldier'

Stage 3

dadaydu	[_ ′ \]	šyaobaytu	[_ ′ \]	'rabbit'
suya	[′ ˅]	šoubyao	[′ ˅]	'watch'

Stage 4

yaba day dəyi	[_ ¯ \ \ ˅]	laba dzaydžəli	[_ ¯ \ \ ˅]	'the horn is here'
dzo wotə	[\ _ ¯]	dzwɔ hwɔčə	[\ _ ¯]	'ride a train'

Recall that the children producing these utterances are responding to pictures; we can thus be quite confident in assigning them the meanings that we have. The extent to which our subjects produced utterances which were tonally perfect but segmentally inaccurate or even incomprehensible provides support for the claim discussed by Van Lancker (1975) that features of pitch are stored and processed in the brain differently from segmental phonological features. We base our claim that the correct tone system is acquired relatively early on the fact that the child is still at the one-word stage when he has the adult tone system almost completely under control.

b. Rising and dipping tones

Now we want to look more closely at the data which support the second and third claims in (3), which involve the contrast between the high and falling tones, on the one hand, and the rising and dipping tones on the other. Together these two claims state that the rising and dipping tones cause trouble for the child.

First, let us look at what happens before the adult 4-way contrast has appeared. One Stage 1 child represented in (5a) produced only high and falling tones. (5a) represents data from a single taping session.

(5) a.

child A		adult		
u	[\]	ü	[′]	'fish'
da	[\]	dan	[\]	'egg'

didi	[\ .]	didi	[\ .]	'brother'
dei	[¯]	č̣ə	[¯]	'car'
wa	[¯]	hua	[¯]	'flower'
ya	[¯]	ya	[¯]	'duck'
æ	[¯]	deŋ	[¯]	'light'

Utternaces to which we couldn't assign a meaning:

falling	48 tokens
high	12 tokens
falling-high } high-falling	22 tokens

As can be seen from the tones in the display in (5a), the child's utterances on this occasion were not random. Words with high and falling tones were pronounced correctly, and many syllables to which we could not assign a meaning occured with high or falling tones or combinations thereof. No syllables were uttered with rising or dipping tones. Further, when the child was presented with a rising or dipping tone word, she would refuse to repeat it, with the single exception of the first word in (5a), to which she consistently gave a falling tone. It is important to note that the four tones are distributed fairly evenly among the words which the child is likely to know. Thus, we can discount frequency as a factor in explaining the skewed distribution of tones in this child's speech. The fact seems to be that the rising and dipping tones are simply somehow more "difficult". We will return to this point.

Another Stage 1 child, shown in (5b) below, had a slightly different strategy. This child produced very few contour tones of any kind. With a few exceptions, all tones were re-interpreted as level tones in his speech.

(5) b.

child B		adult		
u	[¯]	ǰu	[¯]	'pig'
ya	[-] (mid)	yang	[´]	'sheep'

wu [¯]	ü [´]	'fish'
ya [_]	nyao [ˇ]	'bird'
ma [_]	ma [ˇ]	'horse'
u [_]	tu [\]	'rabbit'
ša [_]	šyaŋ [\]	'elephant'
?i?ə [_ ¯]	čičə [\ ¯]	'car'
?ə?i [_ ¯]	šyaoji [_ ¯]	'little chicken'
?i?o [_ ¯]	ǰəntou [_´]	'pillow'
?o?i [_ ¯]	guoči [´ ´]	'flag'
?i?u [¯ _]	byɛnfu [¯ .]	'bat'
?i?u [¯ _]	ifu [¯ .]	'clothes'
?u? [¯ _]	kuzi [\ .]	'trousers'
?ə?ə [¯ _]	maozi [\ .]	'hat'

Notice that with this strategy, this child never missed a high tone. Accordingly, on the basis of the data from these two Stage 1 children, we can set up a hierarchy of ease of learning for the four tones:

(6) 1. high

2. falling

3. rising and dipping

Now, let us look at the types of errors made by all the Stage 2 and 3 children, who we claimed have acquired the adult 4-way contrast, as indicated in (2). The striking fact about Stage 2 and 3 children is that they sometimes confuse rising and dipping tone words. Specifically, what the data show is that at Stages 2 and 3, the children produce high and falling tones correctly all the time, whereas some of the time a dipping tone word is pronounced with a rising tone or a rising tone word is pronounced with a dipping tone. The confusion thus takes the form of substitution of rising for dipping tones and vice versa.

The table in (7) shows the percentage of substitution errors made under the two conditions, utterance-final and before another syllable. The percentages were arrived at by averaging the percentage of errors for the rising and dipping tone syllables for each child bu not counting repetitions within one taping session.

(7)

Adult Tone / Environment	Rising	Dipping
a. ________# (utterance-final)	21%	28%
b. ________X (non-utterance-final)	17%	5%

Average of percentages of errors made by Stage 2 and 3 children for adult rising and dipping tones.

What this table shows is that in utterance-final position (7a), 21% of the rising tone syllables were pronounced with a dipping tone, while 28% of the dipping tone syllables were pronounced with a rising tone. In both these cases, all other responses were correct, i.e. there were no substitutions involving high or falling tones.

In (7b) it can be seen that in non-utterance-final position, substitution errors occured on 17% of the rising tone syllables, but such errors occured on only 5% of the dipping tone syllables in non-utterance-final position.

The striking fact brought out in (7), of course, is that only 5% of the dipping tones in non-utterance-final position were missed by our subjects.

To explain this fact, we return to the tone sandhi phenomena mentioned earlier. Looking again at (1b), recall that the dipping tone changes into a low level tone before any tone except another dipping tone. That is, the rising component disappears, in such environments, and with it the similarity between the rising and dipping tones disappears. Hence, the child is less likely to confuse the two, and the total number of mistakes he makes is much smaller than in utterance-final position. Consequently, there are many fewer dipping tone errors than rising tone errors in his environment because of this fact.

With respect to the data we have been discussing, two points should be made clear. First, it is not in general the case that a given child would have two parts to his lexicon: those words which were always pronounced correctly and those which were always pronounced incorrectly. Rather, as would be expected, it was the case that a given word would be pronounced correctly most of the time and incorrectly once in awhile.

The other point to make clear is that there seems to be no evidence whatever in our data for a stage of tone acquisition where a child has only one rising tone which he uses for all rising and dipping tone words. That is, our data provide no evidence for a stage at which only three-way contrast is being made: high, falling, and one generalized rising tone. Instead,we find that if the children have rising or dipping tones at all, they have them correct at least 70-80% of the time. There seems to be nothing gradual about this transition, which in fact we observed to have occured within the space of less than a month in two of our subjects. The other subjects were all beyond this point when we began our study.

The data we have just examined, then, show clearly that the rising and dipping tones of Mandarin present more problems to the language learner than do the high and falling tones. The rising and the dipping tones seem to be more difficult for the child to produce, they appear later in his speech, and are the only tones confused with each other throughout much of the early acquisition period.

In seeking an explanation for these facts, we note once again that the rising and dipping tones are alike in one important respect: they share, at leat in utterance-final position, a rising end component. It seems clear that it is this rising end component which is involved in the confusion of these two tones. This confusion can be found in data from perception studies as well. For example, Kiriloff (1969) reports an experiment in which adult learners of Mandarin as a second language made considerably more errors in perceiving the rising and dipping tones than the high and falling tones, and that they showed a marked tendency to confuse the rising and dipping tones in identification tasks.

Klatt (1973) reports an experiment in which native speakers of Mandarin were asked to identify tones synthetically produced within a frequency range which was systematically reduced. When the frequency range was reduced down to 2 hz, the high and falling tones continued to be identified nearly perfectly, but the rising and dipping tones were frequently confused.

Thus, our tone acquisition data showing confusion between the rising and dipping tones squares with the results of these perception studies.

Given the rising end component which these two tones share, we can propose two hypotheses to account for the facts. The first is that the difficulty and confusability of the rising and dipping tones are due to the simple fact that these two tones are similar to each other and different from the high and falling tones. We will call this the Similarity Hypothesis.

The other hypothesis is that rising tones are somehow more difficult than high and falling tones.

Although the data concerning the confusion between the rising and dipping tones can be explained by either hypothesis, there is other evidence to favor the Difficulty Hypothesis.

Recall first our Stage 1 child who refused to repeat any rising or dipping tone words. The Similarity Hypothesis would predict that she would respond to such forms, but that her responses would be randonly distributed between the rising and dipping tones, or that they would be forms with tones in between the rising and dipping tones. The fact that she would not respond at all suggests that these tones were too difficult for her to produce at this stage.

Then there is the growing body of independent facts which suggest that rising tones are more difficult than level and falling tones. First, there are three phonological facts brought up by Ohala (1973) which could be accounted for in terms of the Difficulty Hypothesis. The first is the often-noted fact that the incidence of falling tones is far greater than the incidence of rising tones in the languages of Asia and Africa. The second linguistic fact is pointed out by Hyman (1973a) and Hyman and Schuh (1974): in African languages, a low-high (i.e., rising) sequence is vulnerable to assimilations (either anticipatory or perseverative), but a high-low (i.e., falling) sequence is not. In other words, languages tend to try to reduce the number of occurrences of low-high sequences, involving rising pitch, while being content, as it were, to leave alone high-low sequences, involving falling pitch. Finally. the well-known "downdrift" in African languages is very rarely matched by analogous raising or "updrift" phenomenon.

These phonological facts support the hypothesis that rising tones are more difficult, but do not support the Similarity Hypothesis. Furthermore, there is both production and perception evidence supporting the Difficulty Hypothesis as well. On the production side, Ohala and Ewan (1973) and Sundberg (1973) have shown that subjects can execute a falling pitch faster than a rising pitch. Ohala and Ewan speculate that the greater time taken to raise pitch than to lower pitch may reflect the greater physiological effort required to raise pitch.

On the perceptual side, Hombert (1975) indicates that falling tone patterns are perceived more accurately than are rising tone patterns, suggesting that rising tones are more difficult perceptually than falling tones. We suggest that our tone acquisition data, then, when taken together with the phonological facts, the production data, and the perceptual data, constitute another important type of evidence which strongly suggests the correctness of the claim that rising tones are more difficult than falling tones.

c. Tone sandhi

Chao (1951) reports that for Canta at 28 months, "the raising of a 3d-Tone dipping word before another 3d Tone i.e., our (1a) is...only beginning to be learned" (p.33). What we found is that this rule does not appear until Stage 4, for the simple reason that the child cannot be said to be actively applying any rule until he is able to make up his own multi-word predications. Thus, we do not consider the child who correctly produces šyaonyao for the adult šyao nyao (literally 'small bird') to have applied a sandhi rule, since we have no evidence that he perceives this form as anything except a rising-dipping word. In fact, we have evidence showing that, indeed , the child does take šyao nyao 'birdie' to be the word for bird and not a phrase consisting of šyao 'small' and nyao 'bird'. This evidence is the existence of such spontaneous forms as:

da šyao nyao [\ ′ ˇ]

big small bird

in which the expression šyao nyao could only be modified by dà 'big' if it means 'birdie', but not if it means 'small bird'.

That this tone sandhi rule gives the children some trouble at Stage 4 can be seen in three types of data.

(a) mistakes. Sometimes the child simply fails to apply the rule:

hay you šyao ü [ˊ _ _ ˊ]

for which should have been realized as

[ˊ ˊ _ ˊ] 'there are more small fish'

The child's utterance contains the impermissable sequence of two dipping tones in a row. Another type of mistake is exemplified by the following:

hwey yao ni [ˋ _ _]

for /ˋ ˇ ˇ/→[ˋ ˊ ˇ] 'will bite you'

This child produced both the low tone words with low level pitches. Our interpretation is that she knew something should happen with those two words, but she wasn't sure what it was.

(b) self correction. Chao (1951) gives an example from Canta's speech in which she first says incorrectly:

byao you [ˊ ˇ]

for /ˇ ˇ/→[ˊ ˇ] 'has the watch'

and then, apparently realizing that she had pronounced a non-canonical form, gives the correct sandhied form:

byao you [ˊ ˇ]

(c) hesitations. The same child who gave up and produced yao ni 'bite you' with two low level tones was later talking about an animal biting people, and correctly said:

hwey yao ren [\ _ ˊ]

for /\ ˇ ˊ/→[\ _ ˊ]will bite people'

Then she wanted to say that it would bite her sister, which requires the application of the tone sandhi rule. Remembering that she had just said yǎo 'bite' with a dipping tone, she said:

hwey yao ... yao ... ǰyɛǰyɛ [\ ˊ ... ˊ ... _ ˉ]

for /\ ˇ ˇ ˉ/→[\ ˊ _ ˉ]

Her yao 'bite' is correctly sandhied in this utterance, but her hesitation and repetition suggest that the rule has not been completely internalized.

As for the tone sandhi rule (1b), which modifies a dipping tone to a low level tone, Chao has this to say:

> Like that of most Mandarin-speaking children, Canta's modification of 3d Tone dipping into "half 3d" low level ... comes sooner than the modification of 3d Tone into 2d Tone rising before another 3d Tone. (p. 33)

Again, Chao does not provide data for his claim about "most Mandarin-speaking children"; however, here, our data support his assertion: we did not observe any instance of a child failing to apply or having difficulty with this particular rule.

d. Neutral tone

As mentioned above in Section 2, the neutral tone occurs in adult Mandarin on unstressed syllables. Words with unstressed syllables in the child's vocabulary fall into four groups:

(1) Demonstratives and numerals ending with the classifier -ge:

zhèige	'this'
nèige	'that'
yíge	'one, a'

(2) Words ending with the particle -de functioning as:

a.	nominalizer	dàde	'big one'
		xīnde	'new one'
b.	possessive	wǒde	'mine'
		nǐde	'yours'

(3) Words with reduplicated syllables:

bàba	'daddy'
gēge	'older brother'

(4) Nouns ending in the syllable -zi [dzɨ]:

tùzi	'rabbit'

bí<u>zi</u>	'nose'
kù<u>zi</u>	'trousers'
xié<u>zi</u>	'shoes'

Most of the errors involved words of Group 4, those ending in -zi. The most common single tone mistake made by the children was to substitute a dipping tone for the neutral tone in the [dzɨ] syllable of words of Group 4.[1] Looking at Table 1, one can see the total number of neutral tone syllables for each child from whom we had neutral tone data broken down into totals and errors first for Groups 1,2, and 3, and then for Group 4. What Table 1 shows is that most of the errors occurred in Group 4, and most of these involved the substitution of a dipping tone for a [dzɨ] syllable.

There are two facts that call for an explanation in these data. First, why is it that more substitution errors occur on [-dzɨ] syllables than on other neutral tone syllables? Second, why is it the dipping tone that is most often substituted?

The answer to the seond question has been mentioned in Footnote 1: the pitch of the neutral tone is low relative to all the full tones (except dipping); therefore, it is entirely natural that dipping tone, the lowest in the child's inventory, should be the one chosen to replace the neutral tone.

But why does the substitution occur in the first place, and why so much more with [dzɨ] words than with other neutral tone words? Recall that Atkinson-King, in her study of the acquisition of stress (1973), suggested that children do not typically produce stressed for unstressed syllables in words such as table, Daddy, etc. Yet, since full tones occur only on stressed syllables, this is exactly what our Mandarin-speaking subjects were doing. For an explanation, we turn to the language of the caretakers. It turns out that words in Group 4, but virtually never in Groups 1, 2, or 3, are sometimes stressed on both syllables in caretakers' speech to children. But what is so special about words in Group 4, that caretakers should sometimes stress the neutral tone [dzɨ], thus assigning it a dipping tone themselves? We suggest that the answer lies in the fact that these words, unlike those in Groups 1, 2, and 3, are all concrete nouns, that is, names of objects which are found in the child's environment. For caretakers anxious to "teach" the children, then, these are precisely the words which may be presented to the children in an exaggerated and emphasized form. In fact, we have many instances in our data of caretakers attempting to get the child to name a picture by repeating the word in an emphatic form, with stress on the [-dzɨ] syllable. We claim that this fact may explain the relatively large number of errors found in [dzɨ] words in the

Table 1

Summary of Neutral Tone Errors of Word Types

	Groups 1, 2, 3		Group 4	
	Total	Errors	Total	Errors: Dipping tone
1	12	3	25	22
2	43	8	50	24
3	24	8	73	23
4	5	3	8	3
5	0	0	7	1
6	6	2	2	0
7	38	8	53	19
8	13	2	6	0
9	19	4	7	1
10	0	0	6	1
11	0	0	5	5

children's speech compared to the errors made with other neutral tone words.

Our findings, then, do not actually countere emplify Atkinson-King's generalization that young children typically do not misproduce unstressed syllables as stressed where no grammatical function is involved: our subjects' tendency to do this probably has its roots in conflicting stress information for a certain class of words which they get from their caretakers.

Footnotes

[1]It will be recalled that the pitch of the neutral tone is determined by the tone of the preceding syllable:

half-low	.˨	after high tone
middle	·˧	after rising tone
half-high	˙˦	after dipping tone
low	.˩	after falling tone

It can be seen that the neutral tone pitch is relatively low after the high, rising, and falling tones; it is high only after the dipping tone. Correspondingly, the substitution of dipping tone for neutral tone occurred, as would be expected, only after high, rising, and falling tones, but never after dipping tones. After dipping tones, not suprisingly, the neutral tone was on occasion misinterpreted as a high tone. However, there are so few -<u>zi</u> words whose first syllable is dipping tone in the children's vocabulary that we have not included any data for this error here.

*An earlier version of this paper was presented at the Annual Meeting of the Linguistic Society of America, December 1975, and the 8th Annual Child Language Research Forum, Stanford University, April 1976. We are grateful to the following people for their help on this study: to the parents of our subjects, to Patricia Meng-Chen Li for her assistance in collecting the data, to Cecilia Lim for her help in tabulating the data, and to Dale Elliott, Vicki Fromkin, Louis Goldstein, Jean-Marie Hombert, and Russell Schuh for their advice and helpful comments. The responsibility for what appears here, of course, rests entirely with us.

References

Apronti (1969). "The Language of a 2-year-Old Dangme," Proceedings of the Eighth West African Language Congress.

Atkinson-King, K. (1973). "Children's Acquisition of Phonological Stress Contrasts," UCLA Working Papers in Phonetics, #25.

Bar-Adon, A., and W. Leopold (eds.) (1971). Child Language: A Book of Readings. Englewood Cliffs, N.J.: Prentice-Hall, Inc.

Bever, T., G. Fodor, and W. Weksel (1965). "Theoretical Notes on the Acquisition of Syntax: A Critique of 'Contextual Generalization'," Psychological Review, 72:467-482. Reprinted in Bar-Adon and Leopold (1971).

Brotzman, R.L. (1964). "Progress Report on Mandarin Tone Study," Project on Linguistic Analysis, 8, 1-35, Ohio State University.

Chao, Y.R. (1930). "A System of Tone Letters, " Le Maitre Phonetique, 45, 24-27.

_________ (1951). "The Cantian Idiolect: An Analysis of the Chinese Spoken by a Twenty-eight-months-old Child," in W.J. Fischel (ed.) Semitic and Oriental Studies, Berkeley and Los Angeles: University of California Press, 27-44. Reprinted in Ferguson and Slobin (eds.) (1973) and Bar-Adon and Leopold (1971).

_________ (1968). A Grammar of Spoken Chinese. Berkeley and Los Angeles: University of California Press.

Cheng, C. (1971). "A Synchronic Phonology of Mandarin Chinese," Project on Linguistic Analysis, 14.

Crystal, D. (1970). "Prosodic Systems and Language Acquisition," in P. Leon (ed.), Prosodic Feature Analysis. Montreal, Paris: Didier, 77-90.

_________ (1973). "Non-Segmental Phonology in Language Acquisition: A Review of the Issues," Lingua 32, 1-45.

Dreher, J. and P. Lee (1966). "Instrumental Investigation of Single and Paired Mandarin Tonemes," Douglas Advanced Research Laboratory.

Ferguson, C. A. and D.I. Slobin (eds.) (1973). Studies in Child Language Development. New York: Holt, Rinehart and Winston.

Hombert, J.M. (1974). "Production and Perception: Their Respective Roles in One Case of Tonal Development," presented at 49th Annual Meeting of the Linguistic Society of America.

_______ (1975). "The Perception of Contour Tones," Proceedings of the First Annual Meeting of the Berkeley Linguistics Society, Berkeley: Berkeley Linguistics Society.

Hyman, L. (1973a). "The Role of Consonant Types in Natural Tonal Assimilations," in Hyman (1973b).

Hyman, L. (1973b). Consonant Types and Tone, Southern California Occasional Papers in Linguistics,1.

Hyman, L., and R. Schuh (1974). "Universals of Tone Rules: Evidence from West Africa," Linguistics Inquiry 5.1:81-115.

Kaplan, E., and G. Kaplan (1970). "Is There Any Such Thing as a Prelinguistic Child?" J. Eliot (ed.) Human Development and Cognitive Processes. New York: Holt, Rinehart, and Winston.

Kiriloff, C. (1969). "On the Auditory Perception of Tones in Mandarin," Phonetics 20, 2-4

Kirk, L. (1973). "An Analysis of Speech Imitations by Ga Children," Anthropological Linguistics, 15.6, 267-75.

Klatt, D. (1973). "Discrimination of Fundamental Frequency Contours in Synthetic Speech:Duplications for Models of Pitch Preception," Journal of the Acoustical Society of America, 53.1:8016.

Lenneberg, E. (1967). Biological Foundations of Language. New York: Wiley.

Ohala, J. (1973). "The Physiology of Tone," in Hyman (1973).

Ohala, J., and W. Ewan (1973). "Speed of Pitch Change," abstract, Journal of the Acoustical Society of America, 53.345.

Pike, E. (1949). "Controlled Infant Intonation," Language Learning 2.1.

Sundberg, J. (1973). "Data on Maximum Speed of Pitch Changes," Speech Transmission Laboratory Quarterly Progress and Status Report, Royal Institute of Technology, Stockholm. 4.1973, 39-47.

Van Lancker, D. (1975). "Heterogeneity in Language and Speech: Neurolinguistic Studies," UCLA Working Papers in Phonetics, #29.

Wang, W., and K. Li (1967). "Tone 3 in Pekinese," Journal of Speech and Hearing Research 10, 629-36.

Weir, R. (1966). "Some Questions of the Child's Learning of Phonology," F. Smith and G. A . Miller (eds.) The Genesis of Language: A Psycholinguistic Approach. Cambridge, Mass.: MIT Press.

Woo, N. (1969). "Prosody and Phonology," unpublished Ph.D. dissertation, MIT. Available from Indiana University Linguistics Club.

A Further Note on Tone and Consonants

Ian Maddieson

[Phonetics Laboratory, University of

California, Los Angeles]

I am grateful for the attention given to my earlier note on tone and consonants (Maddieson, 1974a) by Larry Hyman (preceding article) and for the opportunity to include a response to his views in this issue of *Working Papers in Phonetics*, although even in this lengthy reply I have not found space to comment on every point that I would have liked to.

As Hyman rightly observes, in linguistic literature there is a predominance of examples where consonant types influence or have influenced tone rather than vice-versa. But I am less impressed by this bias in the literature than Hyman is. Unlike him, I find the bias is not general but is mainly evident in the tone languages of Asia. And I believe a bias in reporting on tone languages contributes to the bias in the published literature. Because of this, cases where consonants influence tone are overrepresented, particularly in historical phonology, and cases where tone influences consonants are underrepresented. Nevertheless some well-documented examples of consonant change resulting from tone have appeared. General physiological and psychological considerations lead to the expectation that such cases should occur with some frequency.

This paper is organised in three sections. The first discusses the question of alternative interpretations of the direction of influence. The second is a survey of the three main areal groups of tone languages, African, American and Oriental. In the course of these sections some additional cases of tonal influence on consonants are added to those discussed in the earlier article and Hyman's refutations of counter-examples to the claim that

(1) consonants affect tone, but tone does not affect consonants (Hyman and Schuh, 1974)

are in turn refuted. A final section discusses briefly the phonetic factors involved in the influence of tones on consonants, and consonants on tones.

1. Interpretation of data

One reason that there may seem to be a predominance of cases where consonants influence tone is that situations are interpreted uniquely because change is presumed to be more likely in that direction. Distributions which could provide the basis for two historical interpretations are almost universally interpreted as showing consonant influence. For instance, if two related languages show the correspondences in (2):

(2)	language A		language B
	$C_{vd}V$	:	$C_{vless}\grave{V}$
	$C_{vless}V$	:	$C_{vless}\acute{V}$

most comparative linguists would prefer to believe A is the conservative language while B has innovated a tone system. These correspondences would equally well support a reverse interpretation[1], that language A has replaced a tonal distinction by a contrast among initial consonants lacking in B. Without further evidence, the correspondence is ambiguous.

It seems that many published claims of consonant influence have failed to consider alternative analyses of the data or look for the decisive evidence that would enable a choice between analyses to be made. In many cases the claim of consonant influence is made *a priori* without evidence being presented.

Let us look at an example of this kind. Luce (1965) reported on two dialects of Riang, a language in the Palaungic division of Khmer-Nicobarese (Thomas and Headley, 1970). 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 we may compare pairs such as (3):

(3)[2]	<u>BR</u>	<u>WSR</u>	
	kók	kók	'nape of neck'
	kâk	kâk	'to bite'
	nán	ná:m	'blood'
	nûm	nûm, nûm	'urine'

Luce does not attempt to present 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. He shares the concensus among scholars studying Mon-Khmer languages and other Asian groups that such a development is simply likely to occur (largely because of the Chinese example). But 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 claims of consonant influence on tone.

Two changes are assumed to have occurred. The development of tones; and the loss of voicing in initial voiced consonants. This latter change is a stock-in-trade of historical phonology in Asia. It has been held accountable for the rise to phonological significance of 'register' distinctions as well as tone in Mon-Khmer. 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[3] reflecting earlier voicing."

If this change is indeed real it may be more explanatory to reverse the assumption that tonal contrasts are innovated in Riang and posit original tones, and voiced and voiceless stop consonant series with equal frequencies. 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 would outnumber voiced ones by 3:1 and this predominance might complete the progress towards merging the stop series. We could thus provide motivation for the lack of voicing contrasts in stops and avoid some of the more awkward assumptions about voicing contrasts and their effects in nasals and approximants.

However the evidence for any original voiced stops in 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 distincion that implied the major one. The choice of these letters is far from proof of the existence of a primary voicing distinction in the earlier language.

If indeed the correspondences of initial voicing with tone and register can be established in Mon-Khmer (and no such correspondences convincingly emerge from the wordlists I have seen) it may rather provide a good case for an evolution of voicing in obstruents from a pre-existing tone or register distinction. The important thing is that a language such as Riang should not be prejudged and counted in among those which show a consonant influence on tone.

In a few cases a language has been the focus of some explicit discussion of the choice between analyses which propose consonant influence on tones and tonal influence on consonants. One such language is

Jingpho, a Tibeto-Burman language. In my earlier paper I reviewed the evidence that favours the conclusion that tone has affected voicing in the stop consonants in Jingpho, and hence this language is a counterexample to the claim in (1) above.

Hyman hopes to dispose of the Jingpho counterexample by asserting that "low tone is accompanied by breathiness in Jingpho" and attributing any effect to the breathiness. Our main source of knowledge on Jingpho is Maran (1971). Maran is a native speaker of the language. He does not report that low tone in Jingpho is accompanied by breathiness, nor have I observed this in listening to the language myself.

In order to understand Maran's account of his language, it may be useful to clarify some aspects of his transcription, particularly the use that he makes of the symbol 'ħ'. This stands, first, for a syllable-final segment which is an 'aspirated' voiceless pharyngeal fricative. Maran (1971:173) describes this segment thus:

> "it is heavily aspirated with the tongue-root somewhat lowered and backed against the pharyngeal wall. The vocal bands [cords] are fully spread in this configuration and the friction created by the radix against the pharyngeal wall almost makes ħ a strident sound"

Because of considerations of phonological patterning Maran groups this segment with the voiced consonants in a [- raised] class. Then, in order to represent a tonal difference on syllables closed by nasal segments, Maran uses the symbol 'ħ' as a diacritical mark to distinguish low tone in pairs of words as in (4)

(4)	ŋaŋ	'wanton'	gaŋ	'pull'
	ŋaħŋ	'firm'	gaħŋ	'wide'

Such sequences of symbols an 'ħŋ' 'ħm' do not represent phonetic segment sequences, nor voiceless nasals, and 'aħ' is not a symbol for a breathy (murmured) vowel. The purpose is rather to present an abstract form which includes a [- raised] consonant from which the occurrence of low tone can be predicted.

In syllables with low tone and final stops Maran does not use 'ħ'. These are represented with the symbols 'b,d, g' as in such words as (5):

(5) gab 'shoot' sad 'kill' dig 'be content'

These are pronounced usually with voiceless final stops thus:

(6) [gàp] [sàt] [dìk]

They only appear voiced when geminated before certain suffixes. It seems to be the case that when these segments appear voiced, it is the result of the low tone environment, voiceless stops do not alternate with voiced segments after high tone. We might also note what Maran (1971:194) says about an archaic form of Jingpho that is used for ritual purposes:

> "while we have shown that synchronic [Jingpho] has a system of finals which contrast phonetically in terms of voicing, archaic [Jingpho] as preserved by its ritualized form shows only one series there, a state of affairs analogous to the other [Tibeto-Burman] languages"

The best explanation of the Jingpho facts, archaic and synchronic, is that there is, or was, a rule that voiced the final voiceless stops in certain low-tone environments. There is no evidence that this is influenced by 'breathiness'.

2. Areal survey of tone languages

The generalisation on tone and consonants (1) is supported in Hyman's reply by an appeal to the widespread nature of the evidence of a 'bias'. The bias is the predominance of reports of consonant influence on tone over tonal influence on consonants. Hyman suggests that in the languages of three areas, i) Africa, ii) the Americas, iii) the Orient, this bias exists.

I would deny that the bias exists in any significant fashion in either the languages of Africa or the Americas. And it is in the literature on Oriental languages that there seems to be a widespread failure to give serious consideration to alternative analyses. Let us examine these three areas in turn.

i) Africa

In the languages of Africa Hyman comments that "tone rules often make explicit reference to the voicing state of consonants". He illustrates this with evidence from Nupe (cf. George, 1970). A rising 'allotone' of high appears after a low tone if the intervening consonant is voiced. However, and Hyman himself has made this observation[4], what is exemplified here is a natural assimilation involving the tone levels. Such a rule is not a case of consonant influence on tone.

In Nupe, a preceding low tone induces a low beginning to the following high tone. In Yoruba, the assimilation occurs regardless of the nature of the intervening consonant, compare (7):

(7)

Nupe			Yoruba		
èdé	[- /]	'cloth'	èdá	[- /]	'creature'
ègbá	[- /]	'taxes'	ìgbá	[- /]	'locustbean tree'
èfú	[- ¯]	'honey'	èfó	[- /]	'vegetable greens'
èkó	[- ¯]	'shea-butter nut'	èkó	[- /]	'learning, studying'

In some ways the Nupe situation is more straightforward than the Yoruba one. Voiced consonants, since they have a fundamental frequency component, provide a more natural means of transmitting the low tone of the preceding vowel to the following vowel. But their role is confined to this transmission. No high tone appears as rising unless a low tone precedes. For example, the sequence high tone-voiced consonant-high tone does not produce a rising tone in the second syllable in (8) as it would if this was a consonant effect on the tones.

(8) e.g. málό [¯ ¯] 'be expensive', not [¯ ╱]

For the same reason Ngizim (Schuh, 1971) and Bole (Lukas, 1969) are not examples of consonant effects on tones[5].

I can add a further illustration from my own investigations (cf. Maddieson, Shopen and Okello, 1973). The Nilotic language Lango has a falling 'allotone' of low after a lexical high tone (the rule deriving the fall does not apply if the high tone preceding low arises from tone sandhi rules). However, if a 'strengthened' consonant intervenes between high and low tones the falling variant of low does not occur. This may be conveniently illustrated by some recent loans, shown in (9), which have two variant pronunciations, one with a strengthened consonant and a final level low tone, the other with a 'weak' consonant but with a final fall. The strengthened consonants are written with a doubled consonant letter.

(9)

	Lango variants				Swahili source	
'sheet'	cukka	[¯ _]	cuka	[¯ ╲]	shuka	'sheet'
'machete'	paŋŋa		paŋa		panga	'machete'
'cotton'	pamma		pama		pamba	'cotton'
'bed'	tanna		tana		tanda	'sleeping mat'

A prinicipal difference between weak and strong consonants is the longer duration of the strong ones. Weak /k/ is a short voiceless velar fricative. Now, in Lango, low tones following weak consonants only occur as falls when a high tone precedes. High tones after strengthened consonants are not lowered. Again there is no consonant influence here. Instead, this is a tone assimilation where the high tone preceding a low may raise the beginning of that following low. Strengthened consonants act like a pause or phrase-boundary and no assimilation occurs.

It is true, in languages such as Nupe, Ngizim, Bole and Lango, that the tone rules refer to adjacent consonant segments. But this does not constitute a demonstration of an 'effect' in the sense under discussion here, any more than a rule predicting the voiceless, alveolar and fricative properties of a consonant preceding a prevocalic stop or nasal in English (e.g., in 'step' or 'smell') demonstrates an 'effect' of those stops or nasals on the preceding consonant.

Once the cases discussed above are set aside as illustrating something other than consonant effects on tone, the number of African cases supporting the generalisation (1) is reduced to a handful. Those I have come across are listed below.

Tone levels in Ewe are partly conditioned by consonant classes, among other things (Ansre, 1961; Sprigge, 1972; Sprigge and Ford,1972; but see

Stahlke, 1971a, 1971b). In Cama, consonant classes condition tonal variants (Stewart, 1970), and this may have historical implications for Yoruba (Maddieson, 1974b). Tonal distribution in Zar (= Sayanci) also depends in part on consonant classes, in particular nonhigh tone is sorted in some environments into mid and low variants, being low after voiced non-implosive segments (Schneeberg, 1971, 1973). Bassa has tonal variants following classes of fortis and lenis consonants (Hobley, 1964).

In a group of Southern Bantu languages there is a class of 'tone-depressing' consonants (Lanham, 1958) which may lead to a following tone occurring as low, when for morphological reasons it would be expected to be high, and cause rising allotones of actual high tones. These languages are in the Nguni and Sotho groups (Doke, 1954). It seems probable that an innovation in a language ancestral to just these groups of languages introduced these effects. There may, in a historical sense, be only one 'event' of consonant influence to be recorded here[6]. Beach (1938) proposed that each of the two lower tones (a mid-low and a mid-falling tone) in the parent language of Nama and Korana (!ora) split into a rising and an unchanged variant in Nama. He believed the rising variant was due to an originally voiced initial consonant, as many of the roots concerned retain voiced initials in Korana.

This is the total extent of examples from the African continent where I know of evidence of consonant effects on tonal phonology. I cannot claim to have examined everything written about African tone languages critically, but I believe I have read enough to be able to say that consonant effects on tone are by no means "common" in the literature. Except as noted above, I have not come across them in firsthand research either.

On the other hand there seems to be a similar handful of cases where tones have determined some property of adjacent consonants. Dunstan and Igwe's (1966) observation that "glottal fricatives" (symbolized /h, ɦ, hj, ɦj/) "appear to be partially voiced before low tones" in the Ohũhũ dialect of Igbo was quoted in my earlier note. Hyman asks if there is instrumental data to confirm this observation. There is not; but linguists with the experience of Dunstan and Igwe are unlikely to have been misled on a point of this kind when relying on their ears. It should be noted that "glottal fricative" is a traditional term usually used for a segment whose articulation is not glottal at all. Prevocalic segments symbolized with /h/ are generally voiceless vowels homorganic with the following vowel. Other accounts of Igbo phonology make it clear that /h/ is a segment of this kind, and /hj/ is a voiceless onset to a nonsyllabic variant of one of the high vowels /i/ or /ɨ/. Therefore this is a case where low tone conditions an earlier onset of voicing in a voiceless segment, and is not to be dismissed by wondering if "glottal consonants can be affected by tone". I would rather be inclined to ask whether more meticulous observation might not demonstrate that there is a similar effect on the aspirated stops in this dialect.

In Guthrie's 'Group C' of Bantu languages, Ellington (1970-1) shows that Ngɔmbɛ/l/ corresponds to /t/ before high tone in Lingala, Bobangi, Mongo, Ntomba, and Bọlia; but /t/ before low tone is common to the group, see (10).

(10)		Ngɔmbɛ	elsewhere in Group C
	'thousand'	kótò	nkótò
	'scorpion'	kòló	nkòtó
	'to leak'	tàngà	tàngà
	'to count'	lángà	tángà

Ellington formulates a 'sound change rule' (11):

(11) $*t \longrightarrow l \;/\; \_\; \acute{V}$

and has over 30 examples of corresponding series which illustrate it from a relatively early layer of vocabulary. He considers and rejects the possibility that these correspondences instead show a consonant effect on tone, and considers the possibility of other causes, but concludes that:

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

The change (11) is shared by the closely related Bwela (or Dɔkɔ) language (Hulstaert, 1961), e.g. Bwela 'to count' is /lánga/. Interestingly, part of the change /t/ ⟶ /l/ is the change from voiceless to voiced. Voiced segments sometimes condition higher tone reflexes in Oriental languages. Here high tone has conditioned voicing.

Work in progress in the UCLA Phonetics Laboratory by Laura Meyers suggests that intervocalic voicing of voiceless consonants in Hausa is more noticeable in low tone environments.[7] Full instrumental documentation of this observation will be available in due course (Meyers, 1976).

Beach (1938) reports that in Korana, /g/ and perhaps other (previously) voiced segments have become devoiced only when one of the higher tones follows. This devoicing must be related to the absence of a tone-split in the higher tones in Nama (see above). In Nama the higher tones must have caused devoicing before the tone-splitting process began; Korana seems to represent an early stage in this tone-dependent devoicing.

In the Southwestern Mande languages, Dwyer (1974, cf. also Dwyer, 1973) provides the most authoritative and up-to-date account of an alternation among initial consonants. His reconstruction of these alternations involves a gemination rule, posited for Proto-Southwestern-Mande, which produces geminated voiceless consonants from voiceless consonants preceded by a low-toned nasal prefix. In Kpelle these geminates occur voiced, so an additional post-nasal voicing process must be reconstructed, whereas in the other Southwestern Mande languages the gemination rule was broadened to include high tone nasals. The result in Kpelle is shown in (12).

(12) *ń-b ⟶ ḿb
*ǹ-b ⟶ bb

Dwyer denies that the low tone is 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. It can only be as good as the reality of stress and the historicity of stress placement in Kpelle. 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. Bird proposes an assimilatory rule between glottal features, which is probably an exaggerated claim.

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 nasals. He cites such examples as (13) in Central Kpelle. By a tone-lowering and voicing rule (13i) becomes (13ii), not (13iii).

(13) i. kálòŋ 'chief' + pɔ́lɔ́ 'old'
→ ii. kálòmbɔ̀lɔ̀ 'old chief'
not iii. kálòmbbɔ̀lɔ̀

However, the non-occurrence of gemination in such an environment is equally a problem for his claim that the process occurs when syllables are unstressed since by his rule this would be an unstressed syllable. As stress is a secondary feature in Kpelle, and unstressed syllables do not undergo change in other ways (e.g. no vowel contraction or vowel deletion conditioned by unstress) it may be mistaken to accept that lack of stress caused the gemination in Kpelle, and to absolve the low tone from some role in the change.[8]

In sum, the languages of Africa provide a number of cases both of consonant effects on tone and tone effects on consonants. At this stage of our knowledge of the phonology of African languages there does not seem to be a preponderance of only one of these.

ii) the Americas

In the literature on the languages of the Americas it is very rare to find references to consonant effects on tone. Hyman notes with satisfaction that Eunice Pike (1974) "concurs with the Hyman and Schuh generalization". But Pike does not say that she concurs because she finds the evidence in the languages of Mexico which she has pre-eminently studied. She uses an illustration from Taiwanese.

Studies of American languages in which comparison and reconstruction of tones is attempted include Trager (1942) on Tiwa, Longacre (1957) on Mixtecan, Rensch (1958) on Chinantecan, Gudschinsky (1959) and Longacre (1962) on Popolocan and Mixtecan, Grimes (1969, 1971) on Quichean, Fisher (1970) on Mayan 'proper', and McLendon (1973) on Poman. In none of these studies is any consonant influence on tonal development reported.

I cannot claim to have read exhaustively in the literature on American tone languages, but from all the comparative studies that I have read the strongest evidence of consonant influence on the tones that emerges is a speculation by Olmsted (1964:56) that the origin of falling tone in Achumawi (= Pit River) may "in some cases ... be connected with the loss of final *-p". This is based on the correspondence of an Achumawi suffix /-ô·/ or /-ô/, which denotes the names of trees and plants, with Atsugewi /-o·p/ as in (14).

(14)

	Achumawi	Atsugewi
'cedar'	làˀtô·	naˀto·p
'spruce'	jɪmjɪməlô·	jɪmjɪmo·p

The correspondence of $\hat{V}$: Vp proves to be far from general (cf. also Olmsted, 1966).

Consonant influence on tone does not feature in Longacre's survey article on "Systemic Comparison and Reconstruction" of the languages of Middle America (1967). However, in an appendix to this article, Longacre does refer to an unpublished reconstruction by Donald Stewart of Proto-Otomi-Mazahua tones. Stewart reconstructed Proto-Otomi tones from three Otomi dialects and compared these with Mazahua. According to Longacre, Stewart's Proto-Otomi-Mazahua high and rising tones each have two reflexes in Mazahua; one occurs before h or a voiceless consonant, the other elsewhere.

Bartholomew (1965) states Stewart's conclusions differently from Longacre. According to her, Stewart found that Proto-Otomi-Mazahua High and Rising tones in disyllables correspond with Mazahua High when followed immediately by /ʔ/ or /h/ and another consonant. Elsewhere in disyllables the correspondences are with Low and Falling tones respectively. After reviewing an additional Mazahua dialect, Bartholomew concludes that Proto-Otomi-Mazahua Low tone also split in Mazahua, becoming High in the /-ʔCV/ and /-hCV/ environments. At a higher level of reconstruction Bartholomew also establishes that separate series of tone correspondences exist between Otomian and Pamean according to the presence or absence of a final consonant in Proto-Oto-Pamean. These cases may be examples of consonant influence on tone but could equally be examples of the influence of syllabic structure. There may be a real case of consonant influence on tone in the languages of this group.Newman and Weitlaner (1950) say in a footnote.

> "That the history of Otomian tones may be closely allied to that of other phonemes is more than a mere suspicion: the multiple stem system which shows up in Central Otomian as one of initial consonant alternation, is a system of tone alternation in the stems of Matlatzinca (Southern Otomian)."

But they do not elaborate and no linguist has pursued this possibility.

Stanley (1969) proposes to predict high tone in Navaho from the occurrence of underlying nasal and glottal stop segments. Krauss (1973) has the following comment on this proposal:

> "that Stanley associates high tone with glottalization is counter to very clear historical evidence that Navaho high tone developed in the absence of glottalization and low in its presence."

Krauss does not cite and I have not found the historical evidence referred to.

Synchronic accounts of tones affected by consonants in American languages are equally hard to come by. The examples I have come across are rather marginal. Arlene Agnew and Evelyn Pike (1957) report that, in Ocaina:

> "Palatalized consonants tend to have a slightly rising pitch glide from the consonant to a following high tone vowel and a fast falling pitch glide to a following low tone vowel:

These consonants are singled out for comment and we might say there is an effect of palatalized consonants on tone. A fuller analysis might find another explanation. Perhaps the palatal quality is due to an underlying (or historically earlier) high vowel which becomes non-syllabic when another vowel follows. The fast falling pitch might then result from a high-low tone sequence. Such an explanation would be capable of explaining the distribution of palatalization, falling tone, and vowel sequences as related. The article does not give enough information on the co-occurrence of vowels and consonants to test this theory.

In a number of Chinantecan dialects (Robbins, 1961; Skinner, 1962; Merrifield, 1963; Westley, 1971) there is some limitation of tone variants in 'ballistically' stressed syllables with a final glottal stop. In Soyaltepec Mixtec (Eunice Pike, 1956), certain enclitic morphemes with the lowest tone (tone 4) occur with a falling tone contour when they follow rising contours. This rise is simplified by the same tone change rule. However, if the enclitic morpheme consists of a vowel alone or /h/ plus a vowel the derived tone is not falling (15i) but level (15ii).

(15) i. $\text{wa}^{3}\text{ku}^{21}$ 'teach' + hnı^{4} 'we (exclusive)'

→ $\text{wa}^{3}\text{ku}^{3}\text{-hnı}^{24}$ 'we teach'

but ii. $\text{wa}^{3}\text{ku}^{21}$ + hı^{4} 'you (sg.)'

→ $\text{wa}^{3}\text{ku}^{3}\text{-hı}^{2}$ 'you (emph.) teach'

The sandhi rule to cover this set of morphemes must be sensitive to their segmental composition, but it is not obvious that the simplification of the fall in the case of (15ii) is due to the effect of /h/.

In the same article, Pike notes that the pitch of a single voiced consonant or a nasal + stop or africate cluster is:

> "conditioned by the tone of the vowel which follows them, and frequently enters into nonphonemic glides constituted of nonphonemic slur from the consonant pitch to the nuclear pitch"

These consonants are described as being "a little higher" than the lower tones 3 and 4 and "a little lower" than the higher tones 1 and 2. In other words, these consonants have a tendency to a middle pitch, and, rather than being subject to conditioning, could be regarded as themselves conditioning rising allotones of high tones and falling allotones of low tones.

Loos (1969:197) writes a rule for Capanahua that alters a rule-assigned high pitch to low in an unstressed syllable if a glottal stop and a low pitch syllabic segment follow in the same word. In Kickapoo, which is not a tone language, Voorhis (1971) observes that a vowel preceding a voiceless fricative which occurs earlier than the primary word stress has low pitch. Low pitch usually demarcates the primary stress, which is recognized by the drop in pitch following it. Such a system could develop into a system with tones if, say, fricatives were elided.

In considerable reading these are the only examples I have found where a prima facie case could be made for consonants affecting or innovating tone in the languages of the Americas; unless some of the Acoma morphophonemic rules discussed by Miller (1965) might be interpreted in this light. There are also a few cases where tones seem to affect consonants.

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 laryngealization (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 in the environment VV-(C) perhaps as a consequence of the effort required to prolong low pitch. Hints for a model of tone-induced change to glottalic consonants could be drawn from this example, but no more than hints.

There is also an observation of interest on Picuris by Trager (1971). The first member of a consonant cluster with a glottal stop as the second member is affected by low tone in a way not found with high and mid. Trager illustrates this effect with the example in (16):

(16)	'chokecherry'	phonemically /êmʔene/	(^ represents low tone
		phonetically [ʌ̂^{m}bʔɛnɛ]	and primary stress)

Of (16), she says:

> "the /m/ becomes partially denasalized and devoiced, and more tense, in anticipation of the /ʔ/"

Unfortunately no further examples are given of this phenomenon. Trager attributes the effect principally to the glottal stop, but as glottal stops unaccompanied by low tone do not have this effect it seems clear that here is another example where consonants are affected by tone.

The general impression made by extensive reading of the literature on American languages (but no first-hand experience) is that few cases are noted of any kind of interaction of consonants and tones. No doubt as more of the languages benefit from adequate attention to pitch phenomena, and reconstructions and analyses of historical processes become more ambitious, further cases will be reported. But, at this time, for most of our evidence of consonant effects on tone and tone effects on consonants we must look to the languages of Asia, not those of America or Africa.

iii) Asia

It is rather unusual to find an account of an Asian tone language that does not prove, claim or assume that consonant types have influenced the history of the tone system. Languages of this area provide some of the clearest evidences of the fact that consonants can affect tones. Asia also provides some of the clearest cases where tones can be shown to have affected consonants, for example, the Jingpho voicing of stops in low tone environments discussed above. In some other cases the available evidence is ambiguous, allowing either an interpretation that consonants affected tone or that tone affected consonants (cf. the discussion of Black Riang above). Since both processes are documented these ambiguous cases cannot be resolved by an assertion that the change is more probable in one direction than the other.

To review all the evidence for consonant effects on tone in Asian languages would be a lengthy and unnecessary process -- the best-known cases, such as the Chinese and Thai tone splits, are fairly well-documented in accessible accounts. But I would like to draw attention to a few more cases of tone effects on consonants not mentioned in my earlier note.

Another instance of low tone facilitating voicing is noted by Bhat (1969). In Tankhur Naga, a Tibeto-Burman language:

> "the three unaspirated stops become voiced intervocalically especially before a vowel in low tone"

Another case of high rising tone introducing glottal stops where none existed before is reported from 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 is a devoiced 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 this sharp rise in pitch can cause the vowel to be broken up by a glottal stop. As this involves closing the vocal cords it is not plausible to suggest a direct change h → ? . Rather, an incidental pitch rise may have been interpreted as phonological, and subsequently the overarticulation of an energetic rise resulted in vocal cord tension reaching the point where vibration was halted. Compare Vietnamese, Danish[9], and Latvian (Ekblom, 1933) for similar introductions of glottal stop segments.

An interesting case where tone seems to affect place of articulation (? stridency) is reported by Burling (1967). Discussing variations in the realisation of Atsi /v/, which in most positions is labio-dental, he notes that before /e/ and /a/ in high-toned syllables it has bilabial friction. The facts are not quite clear from his description, but it seems certain that he did observe some tone-dependent variation in the articulation of this segment.

The largest number of instances of tone effects on consonants found in the literature on languages of the Orient relate to changes in phonation type (including aspiration among these) in initial consonants. These are best known from Chinese but are found elsewhere too. For example, in the Yang-hao dialect of Miao a series of prenasalized obstruents apparently originally with 'voiced aspiration' have lost the aspiration when they occurred with original tones A and D but not with B and C (Li et al., 1972:99). In Puyi dialects of the Tai family, original voiced obstruents have modern voiceless or aspirated reflexes depending on the tones (Sarawit, 1973).

In two of the major groups of Chinese dialects, Mandarin (= Northern Chinese) and Yüeh (= Cantonese), similar effects are reported. For example, Forrest (1965:230) says of Mandarin that, while becoming voiceless,

> "old voiced occlusives ..... became aspirated if the word carried the even tone, but not otherwise"

In some Yüeh group dialects, words in the "lower" rising tone (shang-sheng) also yield voiceless aspirates (Hashimoto, 1972:32). In other tones these earlier (breathy) voiced obstruents yield voiceless unaspirated reflexes.

In Hyman's article the development in Mandarin is discussed, but is dismissed as due to "glottal dissimilation" rather than the tonal difference. Independently of the dubious nature of "glottal dissimilation" it can be shown that Hyman's account of the process in question is inadequate. His presentation is based on one version of the theory that Old Chinese had no tones and that the tones of Middle Chinese developed from characteristic final consonants. In his version distinctive endings for each tone category must be assumed to persist past Middle Chinese times into early Mandarin.

The data on this question are very complex; and the practise of referring to syllable types differentiated by features other than contrastive pitch as "tones" does nothing to simplify the problem. Thus it is traditional to refer to four categories of syllables in Middle Chinese as four "tones" despite the fact that the fourth category (the ju-sheng or "entering tone"[10]) is marked by the presence of final stops which are absent from the other three. Haudricourt (1954a) proposed, on the basis of correlations with hypothetical sources for Vietnamese tones that the other Middle Chinese categories could have earlier had their own characteristic final consonants; specifically the shang-sheng ("rising tone") having had a final glottal stop or glottalized consonant and the ch'ü-sheng ("departing tone") a final /-s/ or /-h/ (< -s).

Following Haudricourt's suggestion, Pulleyblank (1962) elaborated a hypothesis that a final -ɦ or -ð and possibly -v and the plain nasals -m, -n, -ŋ closed syllables in the later p'ing-sheng category; a final glottalized nasal, glottal stop, or cluster including a glottal stop closed the shang-sheng category; and a final -s or cluster including -s (-ks, -ðs, etc., including the nasals -ms, -ns, -ŋs) occurred in the ch'ü-sheng category. Later developments gave rise to the tones from these finals, assuming that -s and the clusters in -s became -h (voiceless vocalic ending?) and the -ms, -ns, -ŋs series became -mh, -nh, -ŋh (presumably -mm̥, -nn̥, -ŋŋ̊ or -m̥, -n̥, -ŋ̊).

Pulleyblank's evidence for this reconstruction is largely drawn from early transcriptions of foreign names using Chinese characters. Karlgren (1963) has objected that it is unclear what pronunciation most of these names may have been given in the source on which the Chinese transcription is based. Hence the foreign names give rather uncertain testimony on the pronunciation of contemporary Chinese graphs. Pulleyblank also cites the derivational suffix -s, which has been identified in Tibetan, as a possible parallel to the derivative function of the ch'ü-sheng tone. Many pairs of words demonstrate an alternation of tones with a 'derived' (e.g., nominalized) form in ch'ü-sheng. (cf. Downer, 1959; Benedict, 1972).

Benedict's observations on this are of particular interest. His conclusion is that two tones should be reconstructed back to Proto-Sino-Tibetan. These two correspond approximately to p'ing-sheng and shang-sheng categories. The ch'ü-sheng he regards as a secondary development arising from a derivational process in Chinese. There is good reason to accept the 'secondary' nature of the ch'ü-sheng tone.[11] But, whatever the ultimate origin of the three non-stopped tones, the point at issue is whether the categories were distinguished by characteristic endings when the split in reflexes of the Middle Chinese (breathy) voiced series of stops occurred.

If there were such characteristic final consonants then it is necessary to make the typologically difficult assumption that Middle Chinese had three series of contrasting final nasals (plain, glottalized, and voiceless or 'aspirated'). There are certainly more languages with pitch contrasts than three series of nasals. It also requires assuming that the phonologists of the period, who reported so carefully on the manners distinguishing initial consonants and had developed a technical vocabulary for doing so, failed to observe and report on manner differences in such final nasals.[12] Equally no trace of such differentiation can be found in the loans from Middle Chinese or earlier into Korean and Japanese (see, for example, Karlgren 1954; Tcheu, 1972).

Kennedy (1964) and Karlgren (1960) find evidence that rhyming practise and phonetic spellings (hsieh-sheng series) argue for purely tonal contrasts even in Old Chinese (although it is not clear that the analysis of rhymes can provide much information on the identity of the rhyming sounds), and Benedict argues on a strictly comparative basis that pitch differences can be projected all the way back to Proto-Sino-Tibetan.

On the other hand, the analogy with Vietnamese is a weak support for a theory of characteristic final consonants. As Denlinger (1972-3) and Mohring (1972) have pointed out, the comparative evidence cited by Haudricourt (1954b, 1961) for the origin of the tonal 'inflections' in Vietnamese is less than overwhelming. Also, comparison with closely related Muong suggests that the Vietnamese ngã tone (a high rising pitch with the vowel 'broken' by a glottal stop) has recently innovated the glottal property which Haudricourt wishes to propose as the cause of the rising pitch contour (cf. Thomas, 1966).

Furthermore Hyman's suggestion that "the pitch-raising effect of a syllable-final glottal stop and the pitch lowering effect of a syllable-final /h/" are "widely attested in the world's languages" is simply not true. The "attestation" of this process consists principally of Haudricourt's well-known, but disputed, theory of the origin of Vietnamese tones. There is one experimental study which supports this. Hombert (1975) showed that pitch rose before /ʔ/ and fell before /h/ in nonsense monosyllables spoken by four Arabic subjects. But there are contrary indications from various languages and from laboratory data. For example, recall Gradin's report on Jeh. In the Northern dialect final /h/ occurs, corresponding with /s/ and /h/ in related Mon-Khmer languages. A Southern dialect has an innovation replacing /h/ with final *rising* not falling pitch. In Palaung (Luce, 1965), where tone is largely predictable from the nature of the final segment, syllables in -h are mid-*level* whereas it is open or nasal- or liquid-final syllables which have a (mid-to-low) falling pattern. As for the glottal stop, in colloquial Amoy Chinese (Sung, 1973) the "upper" entering tone with a final glottal stop is a 32 *falling* contour (its sandhi variant is a 53 fall). In Shan (Egerod, 1957) final glottal stop occurs with low level, high-to-low falling, and high level tones. It does not occur with the low-to-high *rising* tone. Such examples could be multiplied. They provide no basis for asserting that 'h' lowers and 'ʔ' raises tone.

In sum, contempory phonological descriptions and ancient loanwords from Middle Chinese provide no evidence for characteristic endings for the non-stopped tones. No modern Chinese dialect provides evidence for such endings[13], including the conservative Wu dialects which preserve the voicing of initial obstruents. Comparative evidence suggests that the ch'ü-sheng tone has its origin in suffixation, but evidence for the origin of the other tones in characteristic endings at Old Chinese or earlier times is lacking. Internal evidence of various kinds in Old Chinese metrical practise and the writing system suggests that the non-stopped tones were truly tonal before Middle Chinese times. In Middle Chinese the (breathy) voiced stops had not yet diverged in the different tones. Therefore when Hyman claims to "accurately account for the different reflexes of original voiced stops" he is actually discussing an unsupported speculation made without considering the extensive data available.

Even if the 'reconstructed' finals proposed by Hyman could be supported there is no firm foundation for the theory that these finals will yield the desired tone reflexes or will cause the desired 'glottal dissimilation' with the initial consonants.[14] In particular, a serious difficulty with Hyman's casual speculations arises from the existence of a voiceless aspirated stop series which has come down unchanged to Mandarin from Middle Chinese. If Hyman was right, these original voiceless aspirated stops whould also have undergone the process of deaspiration (ph → p) in oblique tones. However, these stops do not have divergent reflexes in the different tones. Hyman's proposal therefore cannot be accurate. Rather, the most probable process involved here is indeed a tonal effect on the original voiced initial consonants.

The nature of this tonal effect is not clear. Gandour (1974) showed that stops in the aspirated series in Siamese tend to be followed by 'voiced aspiration' when high pitch follows. There may be a clue to the Chinese developments in Gandour's findings, but much more research of a similar detailed nature is required before this clue reveals very much.

Hyman does not appear to observe a similarity of the contemporary Siamese situation to the historical Chinese one. He could easily have devised a "glottal dissimilation"-style explanation for Siamese, but instead proposes a more interesting one based on a hypothetical 'perceptual strategy'. This explanation is intended to account for the *retention* of breathy voicing before high tone. Unfortunately, Hyman has overlooked some relevant historical evidence and his explanation cannot succeed.

The aspirated stops in Siamese do not only derive from earlier (breathy) voiced stops, where one could properly talk of retention, but also from an original voiceless aspirated series where Hyman's explanation is inapplicable. Equally, voiceless aspirated stops in words borrowed into Siamese from Pali are treated the same as those from earlier voiced stops. If Hyman's explanation was correct the voiced series would not have merged with the aspirated stops from these other sources. The tone-dependent allophones of aspirated initials must have developed after the merger had occurred since they appear in aspirated stops from all sources. (For details see Brown, 1962; Sarawit, 1973; etc.).

3. Summary

Hyman's attempts to deny the validity of the counterexamples to the generalization that

(1) consonants affect tone, but tone does not affect consonants

are not successful. The modifications and exceptions he proposed do not really rescue a generalization that is basically unsound.

The impression of an overwhelming number of cases of consonant influence on tone is not supported by a review of the literature on tone languages in at least two of the three main linguistic areas concerned. In contrast to an over-readiness to assume consonant influences, a widespread lack of awareness of the possibility of tone invluences on consonants is evident in the conversation and writings of linguists (compare the remarks of Ellington, 1970:148-9). This predisposition to consider one possibility to the exclusion of alternatives must contribute to any remaining apparent statistical bias in favour of consonant effects on tone. Hopefully the present controversy will alert more linguists to the fact and nature of tone effects on consonants.

Phonetic factors in tone/consonant interaction

In my earlier note I objected to the generalization (1) put forward by Hyman and Schuh because a) it seemed to overlook counterexamples, and b) it lacked any accompanying principle to explain the directionality of influence claimed. If a principled explanation had been available to suggest *why* tones should not be expected to affect consonants, the attempt to reinterpret the counterexamples would appear more justified. Hyman now makes a partial attempt to provide the missing explanatory principle.

Hyman's suggestion is that the effects of different consonants on tone are due only to differences in the rate of airflow through the glottis. I do not find it possible to agree with this suggestion. The rate of vocal cord vibration is dependent on the longitudinal tension and the degree of adduction/abduction of the vocal cords as well as the rate of airflow. Since consonants vary in vocal cord tension and aperture it seems improbable that these factors will not also have some effects on the pitch characteristics of adjacent segments. Co-articulation of this kind would appear capable of inducing both anticipatory and perseverative effects.

As for pitch-bearing segments, the principal active means of pitch control is the adjustment of vocal cord tension. It seems difficult to argue that differences of vocal cord tension do not extend to adjacent segments, given the extensive co-articulation that typically occurs in speech. Differences in vocal cord tension could be expected to have effects on the speed and degree of adjustment of glottal aperture, and on the timing of the cessation or beginning of vocal cord vibration (i.e., on voice onset and offset time). It seems equally as likely that such co-articulation effects will assume phonological standing as those due to different consonants.

We do not yet know the limits on these processes. Studies of laryngeal action and of the aerodynamic conditions surrounding the larynx, and studies of perception must both be continued to determine more precisely the mechanisms which can lead to consonant effects on tone and tone effects on consonants. Linguistic studies which assemble data on such effects and construct a typology of them are a prerequisite to the formulation of plausible hypotheses of the constraints on these effects.

A typology of consonant effects on tone would include at least (17) and (18):

(17) Voiced consonants cause lower following tone reflexes than voiceless ones (Songkhla, Mandarin and Southern Min p'ing-sheng, Cham, Camuhi).

(18) Voiceless consonants cause lower following tone reflexes than voiced ones (Siamese, Chiengmai, Southern Min ju-sheng, Nasu).

And a typology of tone effects on consonants would include (19) and (20):

(19) Low pitch causes voicing of adjacent consonants (Jingpho, Tankhur Naga, Igbo, Hausa).

(20) Rising pitch causes glottal closure (Jeh, Vietnamese, Latvian, Danish).

Although (19) is the converse of (17), the two sets of effects should not be expected to be the converse of each other throughout. For example, I doubt if the converse of (20) is found.

But (17) and (18) appear to be in conflict. Obviously the processes involved cannot be identical, but even in the same language both effects can be found. Compare the tone reflexes in Southern Min dialects shown in (21):

(21)

	p'ing-shen		*ju-sheng*	
	Amoy	Gaoxiong	Amoy	Gaoxiong
Middle Chinese *voiceless* initials	55	55	32	22
Middle Chinese *voiced* initials	24	24	55	55

In the p'ing-sheng category, tone reflexes are higher after voiceless consonants, but in the ju-sheng category they are lower. (The reflexes of these two groups of initials are not distinguished according to tone in Min). Solutions which require positing a 'flip-flop' change of voicing in initial consonants or of tonal values probably raise more problems than they solve. It seems more likely that small nonphonological differences in members of the same phonological category of consonants, correlated with variables such as syllable length, amplitude contour, etc., could produce these divergent results in syllables of different categories. There could therefore be a physiological basis for both (17) and (18). But not one which can be represented in terms of the broad categories of segment types listed in Hyman and Schuh's hierarchy of tone-raising and tone-lowering consonants.

Alternatively, the difference between (17) and (18) may have a perceptual explanation (in some instances). One model of a change of perceptual significance from a voicing distinction of initial consonants to a tonal distinction assumes pitch transitions like those shown in (22):

(22) i. following *voiceless* obstruent 

ii. following *voiced* obstruent 

Despite the fact that laboratory data do not generally reveal pitch transitions with this pattern (cf. Gandour, 1974), this model posits cuing on the elevated initial pitch following a voiceless obstruent, and on the depressed initial pitch following a voiced obstruent, and reinterpreting this as a difference of elevated and depressed pitch in the following segment. The effect in (17) would thus appear.

But in many tone languages a centralising tendency of initial pitch can be observed. For example, in Siamese, high and low tone begin from a pitch level closer to mid and move to a more extreme level, as seen in (23):[15]

(23) i. high tone 

ii. low tone 

Now, considering contour shapes alone, (23i) is like the contour in (22ii), and (23ii) is like the contour of (22i). If contour shape is taken as the perceptual cue, then the contour following obstruents would be identified with the low, as both are characterised by a descending pitch ramp. And the contour following voiced obstruents would be identified with high tone, as both are characterised by a rising pitch ramp. The effect in (18) would thus appear.

In Southern Min the non-stopped syllables, already contrastive in tone, may have undergone a split based on the first kind of perceptual reinterpretation, while the stopped syllables (ju-sheng), which had no pitch contrasts, were not subject to this split. Subsequently the ju-sheng tone was split in the fashion suggested by the second kind of perceptual reinterpretation, as an awareness of a characteristic ju-sheng pitch evolved.

Envoi

More research on laryngeal action, pitch perception and the construction of typologies of tone/consonant interaction are urgently needed. There are still many uncertainties concerning the way the larynx operates; no phonological interaction between tone and consonants is possible unless a perceptual mechanism exists to incorporate physical properties of the speech signal into psychological programmes for linguistic behaviour; collation and evaluation of linguistic evidence will show what data require explanation.

A higher-valued hypothesis is one that accounts for a greater range of data, rather than excludes certain data from consideration. Linguists are not responsible for languages having 'crazy rules' but they are responsible for designating certain rules as 'sane' and others as 'crazy'. As in other fields of human existence, I would make a plea for greater tolerance of variability in linguistic rules, and the formulation of hypotheses that are less restrictive than those of Hyman (and Schuh), but aim to include all natural properties of human language.

Acknowledgement

The majority of the work for this article was done while I was working on the Language Universals Project at Stanford University funded by the NSF.

Footnotes

1. There is experimental 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.

2. 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.

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

4. Hyman (1973:166) "The important fact in the above examples is that the consonants should not be seen as *motivating* the tonal assimilations, but rather as *permitting* vs *blocking* them."

5. The reasoning here is similar to that used by Hyman to argue that, under his interpretation, there is no tone effect on consonants in Siamese because what is involved is that certain tones *permit* and others *inhibit* a historical change that could have happened anyway. I agree with the reasoning involved although I disagree with the interpretation (see below).

6. It may be useful to point out that Nguni 'languages' such as Zulu, Xhosa, Ndebele and Siswati (Swazi) are little more than dialects of a single language, despite the establishment of separate literary forms. The Shona group is rather more diverse. I have not had access to some recent South African comparative work on these languages which might clarify my assumption of a single origin for the consonant influence on tone is this area.

7. I am grateful to Laura Meyers for telling me of her discovery. I have described my understanding of what were necessarily preliminary impressions of the data. She may not agree with my way of putting it.

8. A speculative proposal as to the nature of this role might start

from the greater ease of maintaining low pitch in a voiced obstruent compared with a nasal continuant.

9. In my earlier note I commented on the Danish case. Hyman misrepresents the process as having occurred in "many Scandinavian dialects". It has not. He also hopes to dismiss it as an "accentual phenomenon". However the interesting fact is that *two* accents, which contrast in terms of pitch exponents, occur in many modern Scandinavian dialects and must be reconstructed for Common Scandianavian dialects and must be reconstructed for Common Scandinavian (≃ Late Old Norse). These can both be distinguished from absence of accent in unstressed syllables. In Danish only one of these accent types gave rise to the glottal stop. It is reasonable to suppose that the nature of the difference between the accents was crucial to this process; and the difference was a pitch difference.

10. The word 'sheng' really means "sound" not the more specific "tone".

11. The ch'ü-sheng might arise from a *syllable* of opposite tone being suffixed, followed by contraction, rather than be due to the influence of a particular suffixed *consonant*. This might more readily explain the heterogeneous nature of the relationships between the morphologically related tonal variants, as a variety of unstressed syllabic suffixes could be hypothesized.

12. One can get an idea of the care to make the maximum number of distinctions and give them descriptive labels that went into compiling the Ch'ieh-yüh by reading Malmqvist (1968). Mei (1970) has examined some of the early descriptions of the tonal contours themselves.

13. A possible exception might exist in unpublished instrumental records of modern Pekinese made by Karlgren (1960). He found that in kɑn´ 'dare' (rising tone) the -n "was intense and richly voiced to the very end" whereas in k'ɑn` 'look' (falling tone) the -n "had a weak intensity... and lost its voiced nature halfway to the end (being a voiceless n̥ in its final moments)". Karlgren is of the opinion that the sharp fall in amplitude accompanying falling tone is the cause of this devoicing. I have not seen any other mention of this devoicing phenomenon in Pekinese. If it is so fugitive that it has escaped every other observer's attention, I find it easier to believe this is a sporadic conditioned variant than a phonological survival.

14. 'Glottal dissimilation' was proposed as an explanation of tonal developments in Lahu by Matisoff (1970). Its phonetic basis is not convincing. Attempts to compare the proposed 'dissimilation' with the distribution summed up by Grassmann's Law in Indo-European are also unconvincing. 'Plain' and 'aspirated' stop distribution in Indo-European can be accounted for on the basis of position, and it is not necessary (or desirable) to propose a dissimilatory process to explain this (cf. Gamkrelidze, 1975).

15. The centralising pitch tendency in Soyaltepec Mixtec noted by Pike (1956) might be interpreted as an instance of the same type, rather than as a consonant influence on pitch as suggested above.

Bibliography

Agnew, Arlene and Evelyn Pike (1957), "Phonemes of Ocaina (Huitoto)" *IJAL* 23.1:24-27.

Ansre, Gilbert (1961) *The Tonal Structure of Ewe* (Hartford Linguistic Monographs 1) Hartford Seminary Foundation.

Bartholomew, Doris A. (1965) *The Reconstruction of Otopamean (Mexico)* Ph.D. Thesis, University of Chicago, Illinois.

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

Benedict, Paul K. (1972) "The Sino-Tibetan tonal system" in *Langues et Techniques, Nature et Société I: Approche Linguistique* (ed. J.M.C. Thomas and L. Bernot) Klincksieck, Paris: pp. 25-33.

Bhat, D.N.S. (1969) *Tankhur Naga Vocabulary* (Building Centenary and Silver Jubilee Series No. 69) 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) Linguistic Research Inc., Edmonton, Alberta and Champaign, Illinois: pp. 153-174.

Brown, J. Marvin (1962) *From Ancient Thai to Modern Dialects: A Theory* Ph.D. Thesis, Cornell University, Ithaca, N.Y. Published in 1965 by Social Science Association Press, Bangkok, Thailand.

Burling, Robbins (1967) *Proto-Lolo-Burmese* (Publication 43) Research Center in Anthropology, Folklore and Linguistics, Indiana University, Bloomington. Appeared as *IJAL* 33.1 Part II.

Chistovitch, L. A. (1969) "Variations of the fundamental voice pitch as a discriminatory cue for consonants" *Soviet Physics - Acoustics* 14.3:372-378.

Denlinger, Paul B. (1972-3) "On Haudricourt's 'The origin of tones in Vietnamese'" *Monumenta Serica* 30:632-633.

Doke, C. M. (1954) *The Southern Bantu Languages* Oxford University Press, London , for the International African Institute.

Downer, G. B. (1959) "Derivation by tone-change in Classical Chinese" *Bulletin of the School of Oriental and African Studies* 22.2:258-290.

Dunstan, Elizabeth and G. E. Igwe (1966) "Two views of the phonology of the Ọhũhũ dialect of Igbo" *Journal of West African Languages* 3.2:71-75.

Dwyer, David J. (1973) "The comparative tonology of Southwestern Mande nominals" Ph.D. Thesis, Michigan State University, East Lansing.

Dwyer, David J. (1974) "The historical development of Southwestern Mande consonants" *Studies in African Linguistics* 5.1:59-94.

Egerod, Søren (1957) "Essentials of Shan phonology and script" *Bulletin of the Institute of History and Philology, Academia Sinica* 29.1:121-129 (volume titled *Studies Presented to Yuen Ren Chao on his Sixty-Fifth Birthday).*

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

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

Fisher, William (1970) "A study in the reconstruction of Proto-Yucatec" *Papers from the 6th. Regional Meetings, Chicago Linguistic Society:* 14-19.

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

Gamkrelidze, T. V. (1975) "Language typology and Indo-European reconstruction" Talk given to Language Universals Project, Stanford University.

Gandour, Jack (1974) "Consonant types and tone in Siamese" *UCLA Working Papers in Phonetics* 27:92-117.

George, Isaac (1970) "Nupe tonology" *Studies in African Linguistics* 1.1:100-122.

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

Grimes, James L. (1969) "Phonological History of Quichean Languages" Ph.D. Thesis, University of Texas, Austin.

Grimes, James L. (1971) "A reclassification of the Quichean and Kekchian (Mayan) languages" *IJAL* 37.1:15-19.

Gudschinsky, Sarah C. (1959) *Proto-Popotecan: A Comparative Study of Popolocan and Mixtecan* Indiana University Publications in Anthropology and Linguistics, Memoir 15, Bloomington, Indiana.

Haggard, Mark, Stephen Ambler and Mo Callow (1970) "Pitch as a voicing cue" *Journal of the Acoustical Society of America* 47.2.2:613-617.

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

Haudricourt, André-Georges (1954a) "Comment reconstruire le chinois archaique" *Word* 10.3:351-364.

Haudricourt, André-Georges (1954b) "De l'origine des tons en vietnamien" *Journal Asiatique* 24.1:69-82.

Haudricourt, André-Georges (1961) "Bipartition et tripartition des systèmes de tons dans quelques langues d'Extrême-Orient" *Bulletin de la Société de Linguistique de Paris* 56:163-180.

Hobley, June (1964) "A preliminary tonal analysis of the Bassa language" *Journal of West African Languages* 1.2:51-56.

Hombert, Jean-Marie (1975) "Development of tone from segmentals: evidence from contour tone perception" Paper presented to the 8th International Congress of Phonetic Sciences, Leeds. Forthcoming in the *Proceedings*.

Hulstaert, G. (1961) "Sur le parler Dɔkɔ" *AEquatoria* 24.4:121-135.

Hyman, Larry M. (1973) "The role of consonant types in natural tonal assimilations" in *Consonant Types and Tone* (ed. L.M. Hyman) University of Southern California, Los Angeles. pp. 151-179.

Hyman, Larry M. and Russell G. Schuh (1974) "Universals of tone rules: evidence from West Africa" *Linguistic Inquiry* 5.1:81-115 (Revised version of "Universals of tone rules" *Working Papers on Language Universals* 10:1-50 (1972)).

Karlgren, Bernhard (1954) *Compendium of Phonetics in Ancient and Archaic Chinese* Reprinted from *Bulletin of the Museum of Far Eastern Antiquities, Stockholm* 22.

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

Karlgren, Bernhard (1963) Postscript (pp. 18-24) to "Loan characters in Pre-Han texts" *Bulletin of the Museum of Far Eastern Antiquities, Stockholm* 35:1-128.

Kennedy, George A. (1964) "Tone in Archaic Chinese" in *Selected Works of George A. Kennedy* (ed. Tien-yi Li): pp. 135-150. Far Eastern Publications, Yale University, New Haven, Connecticut. (Reprinted from *Wennti Papers* 1, 1954).

Krauss, Michael E. (1973) "Na-Dene" in *Current Trends in Linguistics* 10.2 (Linguistics in North America) (ed. T. A. Sebeok) Mouton, The Hague, pp. 903-978.

Lanham, L. W. (1958) "The tonemes of Xhosa" *African Studies* 17.2:65-81.

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 Yu-hung and Chu Kwo-ray (Data Paper No. 88, Southeast Asia program) Cornell University, Ithaca, N.Y.

Longacre, Robert E. (1957) *Proto-Mixtecan* (Publication 5) Indiana University Research Centre in Anthropology, Folklore, and Linguistics, Bloomington, Indiana. Appeared as *IJAL* 23.4.

Longacre, Robert E. (1962) "Amplificaiton of Gudschinsky's Proto-Popolocan-Mixtecan" *IJAL* 28.4:227-242.

Longacre, Robert E. (1967) "Systemic comparison and reconstruction" Chapter 5 (pp. 117-159) of Vol. 5 *Linguistics* (ed. N. A. McQuown) of *Handbook of Middle American Indians* (general ed. R. Wauchope) University of Texas Press, Austin.

Loos, Eugene,V. (1969) *The Phonology of Capanahua and its Grammatical Basis* (Summer Institute of Linguistics Publications in Linguistics and Related Fields No. 20) S.I.L., Norman, Oklahoma.

Luce, G. H. (1965) "Danaw: A dying Austroasiatic language" in *Indo-Pacific Linguistic Studies Part 1: Historical Linguistics* (ed. G. B. Milner and E. J. A. Henderson) North-Holland, Amsterdam: pp. 98-129.

Lukas, Johannes (1969) "Tonpermeable und Tonimpermeable Konsonanten im Bolanci (NordNigerien)" in *Ethnological and Linguistic Studies in Honour of N.J. van Warmelo* Department of Bantu Administration and Development, South Africa: pp. 133-138.

Maddieson, Ian (1974a) "A note on tone and consonants" *UCLA Working Papers in Phonetics* 27:18-27.

Maddieson, Ian (1974b) "A possible new cause of tone-splitting - evidence from Cama, Yoruba and other languages" *Studies in African Linguistics*, Supplement 5:205-221, also *UCLA Working Papers in Phonetics* 27:28-46.

Maddieson, Ian, Tim Shopen and Jenny Okello (1973) "Lango tonology, suprasegmentally, and paradigms" Paper presented at the 4th Annual Conference on African Linguistics, Queens' College, New York, N.Y.

Malmqvist, Göran (1968) "Chou Tsu-mo on the Ch'ieh-Yün" *Bulletin of the Museum of Far Eastern Antiquities, Stockholm* 40:33-78.

Maran, La Raw (1971) *Burmese and Jingpho: A Study of Tonal Linguistic Processes* (Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics 4) University of Illinois, Urbana.

Matisoff, James A. (1970) "Glottal dissimilation and the Lahu high-rising tone: a tonogenetic case-study" *Journal of the American Oriental Society* 90.1:13-44.

Mclendon, Sarah (1973) *Proto-Pomo* (U.C. Publications in Linguistics 71) University of California Press, Berkeley.

Mei Tsu-Lin (1970) "Tones and prosody in Middle Chinese and the origin of the rising tone" *Harvard Journal of Asiatic Studies* 30:86-110.

Merrifield, William R. (1963) "Palantla Chinantec syllable types" *Anthropological Linguistics* 5.5:1-16.

Meyers, Laura (1976) *Tone Patterns in Hausa* Ph.D. Thesis *UCLA Working Papers in Phonetics* 32 (University of California, Los Angeles)

Miller, Wick R. (1965) *Acoma Grammar and Texts* (U.C. Publications in Linguistics 40) University of California Press, Berkeley.

Mohring, Hans (1972) "Einige Gedanken zum Ursprung der Tone im Vietnamesischen" *Zeitschrift für Phonetik, Sprachwissenschaft, und Kommunikationsforschung* 25.3: 232-242.

Newman, Stanley and R. J. Weitlaner (1950) "Central Otomian II: Primitive Central Otomian reconstructions" *IJAL* 16.2:73-81.

Olmsted, D. L. (1964) *A History of Palaihnihan Phonology* (U.C. Publications in Linguistics 35) University of California Press, Berkeley.

Olmsted, D. L. (1966) *Achumawi Dictionary* (U.C. Publications in Linguistics 45) University of California Press, Berkeley.

Pike, Eunice V. (1956) "Tonally differentiated allomorphs in Soyaltepec Mazatec" *IJAL* 22.1:57-71.

Pike, Eunice V. (1974) "A multiple stress system versus a tone system" *IJAL* 40.3:169-175.

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

Pulleyblank, Edwin G. (1962) "The consonantal system of Old Chinese" *Asia Major* 9.1:58-144 (Part 1) 9.2:206-239 (Part 2)

Rensch, Calvin Ross (1958) *Proto-Chinantec Phonology* (Serie Cientifica No. 10, Papeles de la Chinantla No. 4) Museu Nacional de Anthropologia, Mexico City, and Summer Institute of Linguistics, Huntington Beach, California.

Robbins, Frank (1961) "Quiotepec Chinantec syllable patterning" *IJAL* 27.3:237-250.

Sarawit, Mary (1973) "The Proto-Tai Vowel System" Ph.D. Thesis, University of Michigan, Ann Arbor.

Schneeberg (Tolessa), Nan (1973) "Sayanci Phonology" Ph.D. Thesis, Indiana University, Bloomington.

Schuh, Russell G. (1971) "Verb forms and verb aspects in Ngizim" *Journal of African Languages* 10.1:47-60.

Shorto, H. L. (1967) "The register distinctions in Mon-Khmer languages" *Wissenschaftliche Zeitschrift der Karl-Marx-Universität, Leipzig; Gesellschafts- und Sprachwissenschaftliche Reihe* 16.1-2:245-248.

Sivertsen, Eva (1956) "Pitch problems in Kiowa" *IJAL* 22.2:117-130.

Skinner, Leo E. (1962) "Usila Chinantec syllable structure" *IJAL* 28.4:251-255.

Sprigge, R.G.S. (1972) "An enquiry into the tonemic approach to tone" Paper presented to the 10th. West African Languages Congress, Legon, Ghana.

Sprigge, R.G.S. and Kevin Ford (1972) "General tone rules for nouns and verbs in Ewe" Paper presented to the 10th West African Languages Congress, Legon, Ghana.

Stahlke, Herbert (1971a) "The noun prefix in Ewe" *Studies in African Linguistics* Supplement 2:141-159.

Stahlke, Herbert (1971b) "Topics in Ewe Phonology" Ph.D. Thesis, University of California, Los Angeles.

Stanley, Richard J. (1969) "The Phonology of the Navaho Verb" Ph.D. Thesis, M. I. T., Cambridge, Mass.

Stewart, John (1970) "The lenis stops of the Potou lagoon languages" Paper presented at the 9th West African Languages Congress, Freetown, Sierra Leone.

Sung, Margaret M-Y. (1973) "A study of literary and colloquial Amoy Chinese" Ph.D. Thesis, Stanford University, California.

Tcheu, Soc-Kiou (1972) "Correspondences phonologiques du Sino-Coréen à l'Ancien Chinois" in *Langues et Techniques, Nature et Société I: Approche Linguistique* (ed. J.M.C. Thomas. and L. Bernot) (Klincksieck, Paris: pp. 237-256.

Thomas, David D. (1966) "A note on Proto-Viet-Muong tones" in *Studies in Comparative Austroasiatic Linguistics* (Indo-Iranian Monographs 5) (ed. N.H. Zide) Mouton, The Hague: pp. 26-27.

Thomas, David D. and Robert K. Headley Jr. (1970) "More on Mon-Khmer sub-groupings" *Lingua* 25.4·398-418.

Trager, Felicia (1971) "The phonology of Picuris" *IJAL* 37.1:29-33.

Trager, George L. (1942) "Historical phonology of the Tiwa languages" *Studies in Linguistics* 1.5:1-10.

Voorhis, Paul (1971) "Notes on Kickapoo whistle speech" *IJAL* 37.4: 238-243.

Welmers, William E. (1962) "The phonology of Kpelle" *Journal of African Languages* 1.1:69-93.

Westley, David O. (1971) "The Tepetotutla Chinantec stressed syllable" *IJAL* 37.3:160-163.

*Measuring Larynx Movement in Standard Thai Using the Cricothyrometer**

Jack Gandour and Ian Maddieson

[Linguistics Department, UCLA]

1.0 Introduction

This paper presents the results of an investigation of vertical adjustments of the larynx in a tone language. The language investigated is Standard Thai (or Siamese), which has five lexically contrastive tonal categories. The relative contributions of the following factors in determining the vertical position of the larynx are assessed: (1) tonal categories, (2) phonation type of consonants, (3) vowel quality, (4) pitch, and (5) position in the utterance.

It has often been reported that a positive correlation exists between pitch height and larynx height (Eijkman 1903, Kenyon 1927, Parmenter and others 1933, Curry 1937, Zemlin 1968). Many of these studies are of pitch adjustments during singing or sustained phonation of a vowel, and Zemlin's (1968:191) remark that "it is a rather common occurrence for the larynx to rise during the production of high-pitched tones and to lower during the production of low tones" is clearly intended to refer principally to the evidence from studies of this kind rather than from studies of linguistic tones.

More recently larynx height has been investigated in more speechlike material (Vanderslice 1967, Perkell 1969, Ohala 1972, Ewan and Krones 1972, Ewan 1973, Ewan and Krones 1974). While admitting that other factors are also involved, these studies do find a positive correlation between pitch and larynx height. Maran (1971) and Bird (1971) have gone so far as to propose phonological features of tone that are based on such a correlation. However, in another study (Lindqvist and others 1973) which used the most natural running speech no such correlation between pitch and larynx height was found.

---

*This research was supported in part by NSF Grant GS 37235X and by the Social Science Research Council and the American Council of Learned Societies. We wish to thank our Thai subject Sirichai Hamichart for his patience and cooperation. We also wish to thank Willie Martin for his invaluable technical assistance in the laboratory and Alice Agin of the UCLA Psychology Department for her assistance with the statistical analysis. Thanks also to Peter Ladefoged and Mona Lindau for helpful comments and suggestions.

Where no correlation exists it may be because other factors have combined to override the link between pitch and larynx height. For example, it has been shown that differences of larynx height can be associated with vowel quality (Parmenter and others 1933, Perkell 1969, Menon and Shearer 1971, Ewan and Krones 1974, Lindau 1975); with phonation type of consonant (Stetson 1951, Lindqvist and others 1973, Ewan 1973); with place of articulation of consonant (Ewan and Krones 1972, 1974); and the general pattern of the utterance (Lindqvist and others 1973). The effects of vowel quality, phonation type of consonant, and the general pattern of the utterance were included in the present investigation.

The conclusions presented are the first to result from an extensive use of the 'cricothyrometer'.

2.0 Methods and Procedures

Larynx displacements were measured by means of the cricothyrometer (see Figure 1) invented by Vanderslice (1967:74-76).[1] This device was constructed by appropriate modifications of a BBN Grafacon Rho-Theta transducer to enable it to track vertical larynx movements. The transducer was turned on its side, the rotating arm was counter-balanced and the telescoping rod was spring-loaded and provided with a rubber tip to engage the subject's larynx through the skin over the thyroid notch. Vertical excursions of the larynx move the arm up and down. As the angle of the arm varies, a variable DC voltage proportionate to the height of the larynx is produced by the rotary Theta potentiometer.

The subject is seated upright in a comfortable position. A specially contoured headrest made of gypsum plaster is used to support and hold the head and neck and prevent head movement during the recording.

This device has a number of advantages. It is relatively simple and easy to use; it makes it possible to obtain a large amount of data in a relatively short period of time; and finally, it permits the subject to assume a fairly normal sitting posture during the recording. There are, however, a couple of disadvantages. The subject must have a thyroid depression large enough to engage the rubber tip securely which effectively limits the number of eligible subjects. And, while recording, the rubber tip occasionally slips out of the thyroid notch and has to be repositioned. If, at the same time, the subject changes his position with respect to the cricothyrometer, the larynx height signal must also be recalibrated for subsequent measurement.

---

[1]Perhaps the device would be better named the 'thyroidometer' since it is capable of measuring displacements of the thyroid prominance only, not displacements of the cricoid cartilage.

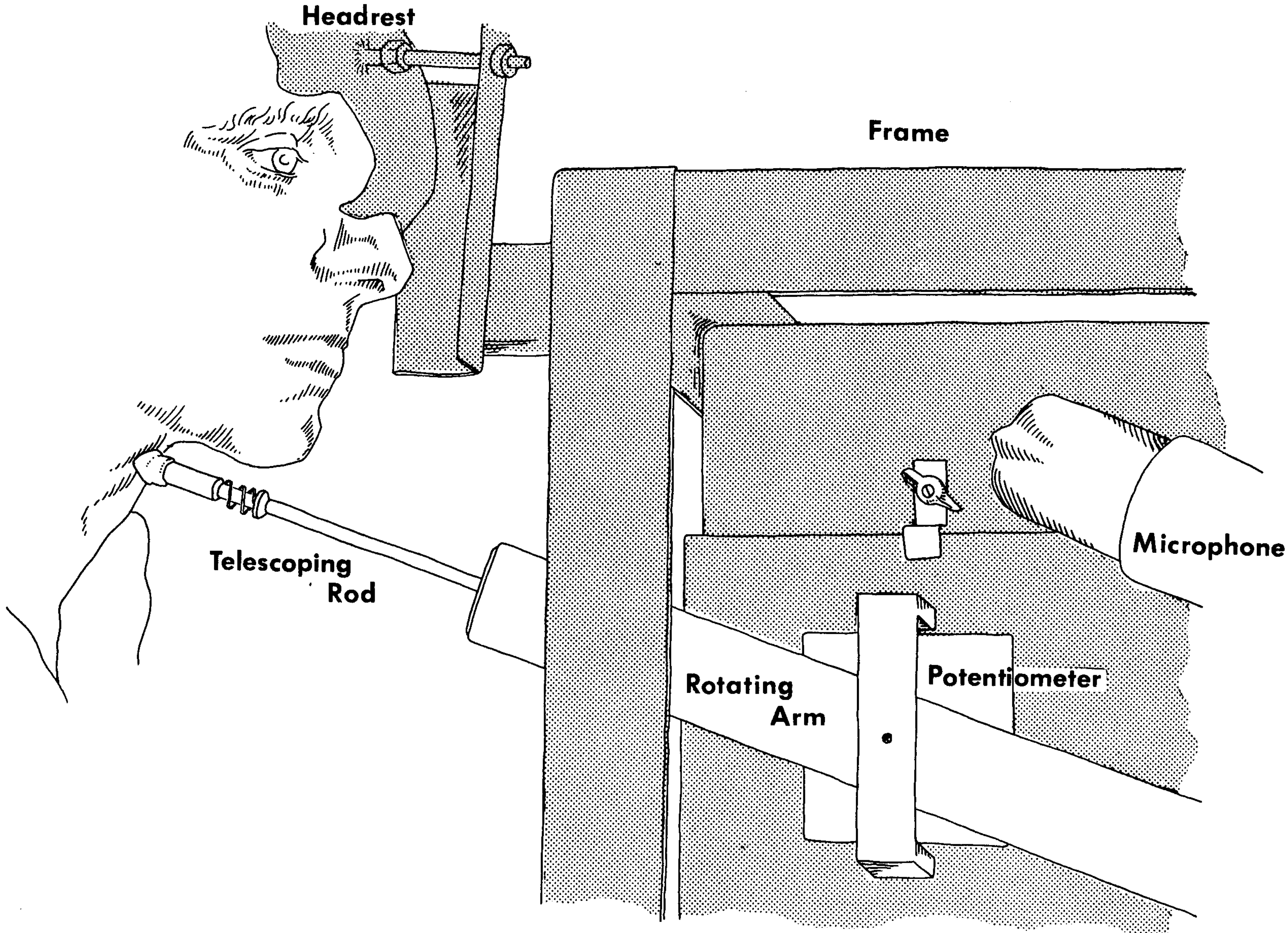


Figure 1. Sketch of the 'cricothyrometer' illustrating the major components and the position adopted by a subject when the device is in use. Drawn by V.S. Fox from a photograph by Willie Martin.

The data consisted of test syllables pronounced in the sentence frame

[nîi ___ khráp]

'This(is a) (polite particle)'

at normal conversational speed by a 24 year old male native speaker of Standard Thai. The test syllables were of the shape CVV^T where C represents either the voiceless unaspirated bilabial stop /p/ (often described as accompanied by glottal constriction), the voiceless aspirated bilabial stop /ph/, or the voiced bilabial stop /b/: VV represents the long low back unrounded vowel /aa/, the long high front vowel /ii/, or the long high back rounded vowel /uu/: T represents a tonal contour which is (1) mid-level, (2) low-falling, (3) high-rising-falling, (4) high-rising, or (5) low-rising - hereafter MID ( ), LOW (`), FALLING (^), HIGH (´), RISING (ˇ), respectively.

A set of flash cards, on which the following 45 types (3 consonants X 3 vowels X 5 tones) of test syllables were written in the Standard Thai orthography,[2] was prepared and arranged in random order. Although most of these test syllables are nonsense words, some of them actually do occur in the language.

paa	pàa	pâa	páa	pǎa
pii	pìi	pîi	píi	pǐi
puu	pùu	pûu	púu	pǔu
phaa	phàa	phâa	pháa	phǎa
phii	phìi	phîi	phíi	phǐi
phuu	phùu	phûu	phúu	phǔu
baa	bàa	bâa	báa	bǎa
bii	bìi	bîi	bíi	bǐi
buu	bùu	bûu	búu	bǔu

The subject was cued by the flash cards which were presented at approximately 5-second intervals.

A total of 450 test syllables (45 types X 10 tokens) were recorded at a single session in the soundproof booth in the UCLA Phonetics Laboratory. The recording was regularly interrupted after every 50 or so utterances in order to give the subject a brief rest. In addition, unscheduled interruptions occurred as a result of the rubber tip occasionally slipping out of the thyroid notch. When this happened, the recording was stopped temporarily while the rubber tip was repositioned. Because of these occasional

[2]The Standard Thai orthography does permit unambiguous representation of the tones of these syllables.

slippages as well as periodic coughs, deep swallowing, and misreadings of the flash cards about 100 of the recorded utterances were later found unsuitable for measurement.

The microphone signal and larynx height signal were recorded simultaneously on separate channels on an Ampex SP-300 recorder. The tape was then analyzed using the hardware Pitch Extraction System of the UCLA Phonetics Laboratory. A continuous oscillogram of the fundamental frequency contours, wave form, larynx height, and amplitude (simultaneously on separate channels) was produced on a Siemens oscillomink. Here, another 60 tokens of the utterances were discarded because of artifacts introduced by the Pitch Extraction System. As a result 290 tokens of the test syllables were finally obtained--an average of 6.4 tokens per test-syllable type.

Fundamental frequency values were read off the oscillogram with a pitch scale (0.5 mm = 2.5 Hz) prepared from the lines produced by a pitch calibration tape that was run immediately after the data had been processed through the oscillomink. The larynx height values (to the nearest 0.5 mm) were measured with a 1 mm ruler from a baseline established as the lowest level attained in a downward movement of the larynx. These latter values do not represent actual vertical movements of the larynx but magnify the actual movements by some small but unknown factor.

Simultaneous measurements of fundamental frequency and larynx height in the vowel were made for each token. These measurements were first taken at the onset and offset of the vowel in the test syllable. Vowel onset was determined primarily by the rapid rise in amplitude and vowel offset by the rapid fall in amplitude. The wave form and fundamental frequency curve were also used. The vowel in each test syllable was then divided into four equally-spaced intervals between vowel onset and vowel offset. This partitioning of the vowel yielded a total of 5 measurement points--one at the beginning of the vowel, another at the end, and 3 points in the middle. The points of measurement for fundamental frequency and larynx height and their respective scales are illustrated on the sample oscillogram in Figure 2. This measurement procedure was adopted in order to normalize the slight differences in duration across tokens. Visual inspection of both the fundamental frequency and larynx height tracks on the oscillograms led us to establish 5 as the number of points sufficient to adequately plot typical pitch curves for each of the tonal contours of Standard Thai (see Abramson 1962, Hiranburana 1971, Gandour 1974) and represent the vertical movements of the larynx.

Apart from these measurement points in the vowel, measurements of larynx height alone were taken at 3 additional points. These points are Lx (high)-- the highest position of the larynx during the closure phrase of the bilabial stops; Lx(low)--the lowest position of the larynx at the end of its downward movement in transition to the vowel of the test syllable; Lx(kh peak)--the highest position of the larynx at the end of its upward movement in transition to the initial voiceless aspirated velar stop in /khráp/.

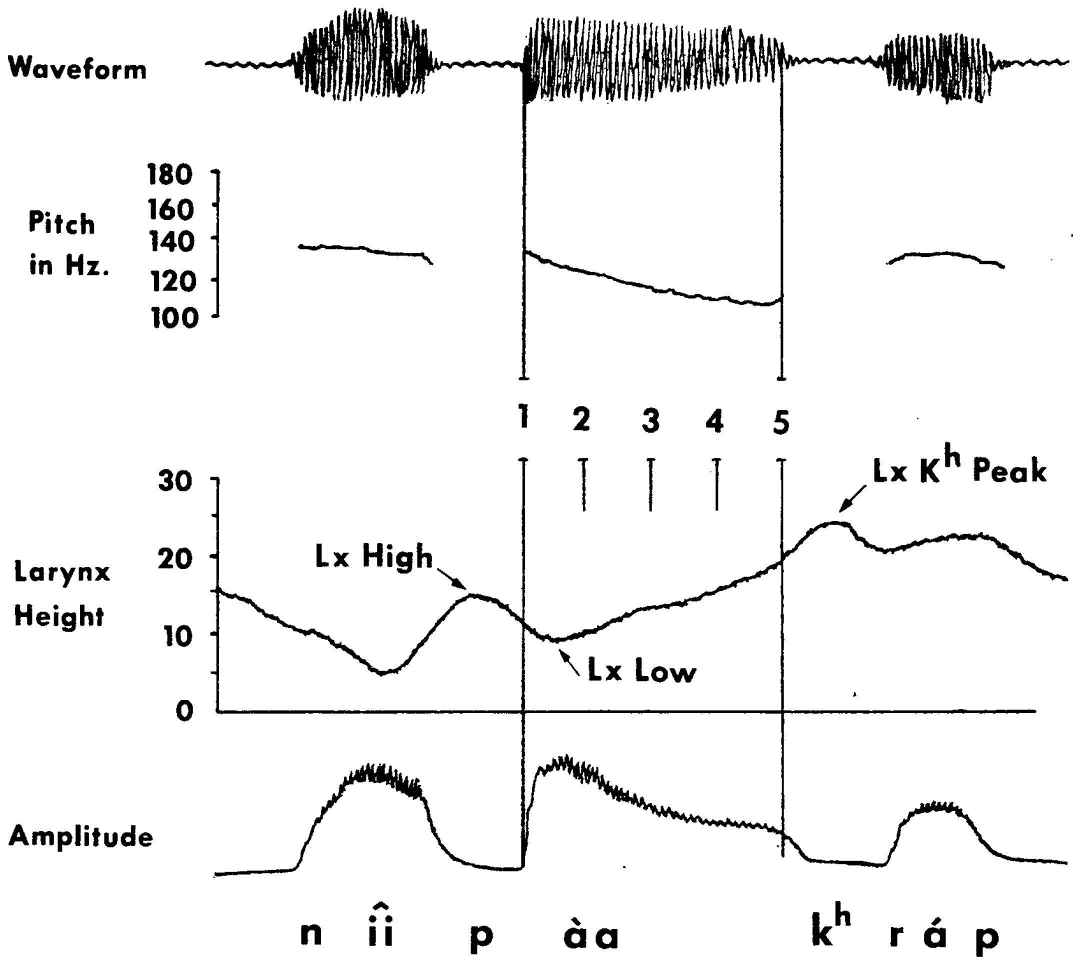


Figure 2. A sample oscillogram, retouched by hand, of the syllable pàa in the frame used for the experiment. The figure shows waveform, pitch, larynx height and amplitude tracings, and the scales for pitch and larynx height measurement. The baseline for larynx height measurements is drawn in. The five points in the syllable at which larynx height and pitch were measured are numbered. Number 1 coincides with vowel onset, number 5 with vowel offset. The additional three points at which larynx height was measured (Lx High, Lx Low and Lx Kh Peak) are also indicated.

3.0 Results

3.1 Consonants and Larynx Movement

From a rest position (about 15 on our scale) the larynx began to move down before the nasal in /nii/ of the carrier frame /nii...khráp/ was articulated. From a low point in the middle of the vowel /ii/ (/nii/) it rose to a high point during the closure of the bilabial stops before falling again. In the test syllable it rose steadily, reaching a peak during the closure phase of /kh/ of the word /khráp/. This typical sequence of movements can be seen in the sample oscillogram in Figure 2.

In the sentence-frame /nii...khráp/ the peaks in the rising movement of the larynx coincide with the consonants. The vertical position of the larynx during the closure phrase of the bilabial stops at the beginning of the test syllable is lower than that for the voiceless aspirated velar stop at the beginning of the word /khráp/. In general, the highest peak in the vertical larynx movement coincides with /kh/. In approximately 20% of the tokens, however, the highest peak occurs with the final /p/.

From Lx(low) to Lx(kh peak) the major part of the rise in larynx position occurs during the vowel of the test syllable. The mean larynx position at Lx(kh peak) is 12.8 units higher than Lx(low) while the mean larynx position at vowel offset is 9.5 units higher than Lx(low). Thus, of the rise from Lx(low) to Lx(kh peak) some 75% occurs within the vowel.

The mean larynx position (taken from the means of each test-syllable type) for Lx(high) and Lx(low) and the mean distance between Lx(high) and Lx(low) for the test-syllable initial bilabial stops are given below in Table 1.

Table 1. Larynx measurments for the bilabial stops.

	L(high)	L(low)	Distance L(high)-L(low)
p	18.3	12.3	6.0
p^h	17.6	12.5	5.1
b	14.4	8.9	5.5

The larynx movement from Lx(high) to Lx(low) for /p/ and /ph/ starts and finishes at significantly higher positions than that for /b/, but the distance traversed is of the same order of magnitude for all 3 bilabial stops. The larynx reaches a slightly higher Lx(high) and traverses a slightly greater distance for /p/ than for /ph/.

The results of paired-data T tests on the mean Lx(high) and Lx(low) values taken from the 15 test-syllable types associated with each phonation type of consonant) for /p/, /ph/, and /b/ are given in Table 2. These differences in larynx position for /p/, /ph/, and /b/ are broadly in line with the results obtained by Ewan and Krones (1974) for their Thai subject.

Table 2. Results of paired-data T tests on mean L(high) and L(low) values for the bilabial stops.

L(high)		L(low)	
b : p^h	.0005	p : b	.0005
p : p^h	.05	p : p^h	not significant

Some tendencies to distinct timing patterns for these consonants can also be observed. For /p/, Lx(low) occurs after vowel onset (see Figure 2, 2.0); for /b/, Lx(low) more or less coincides with vowel onset; for /ph/, the timing is less predictable but Lx(low) seems to occur before vowel onset most frequently.

Differences in larynx height during the vowel of the test syllable can be associated to some extent with the phonation type of the initial consonant (see Figure 3 below). At vowel onset, the position of the larynx is lower after /b/ than after /p/ or /ph/; at measurement points 3, 4, and 5 (vowel offset), however, the position of the larynx is higher after /b/ than after /p/ or /ph/. No marked differences in larynx position in the vowel after /p/ and /ph/ are observed.

3.2 Vowel Quality and Larynx Movement

The inverse correlation usually found between pitch and larynx height on the vowels /aa/, /ii/, and /uu/ may be seen by comparing the larynx height and pitch curves for each vowel in each tonal category presented in Figure 4.

The vowel /aa/ generally has a lower pitch than the high vowels /ii/ and /uu/. This difference is small or negligible in the intermediate range of pitch and is larger in both the higher and lower parts of the pitch range. For example, in the MID tone, the vowels are very close in pitch throughout the test syllable and, in the FALLING tone, the intrinsic pitch difference between the vowels is less when the pitch passes through the median pitch range and greater at both the higher beginning and lower ending of this tonal contour. Hombert (1975) found a different pattern in Yoruba in which the intrinsic pitch differences were least on the lowest pitch.

Although, in general, the vowel /aa/ has a higher larynx position than the high vowels /ii/ and /uu/, this difference tends to diminish towards the end of the test syllable and is even reversed at the end of test syllables with the RISING tone.

The rank order of the vowels in terms of larynx height /aa ii uu/ that we generally find in this study is in agreement with the findings of Chiba and Kajiyama (1958), Ohman (1967), Perkell (1969), Ladefoged (1972), and Lindau (1975). Our findings, however, are not in agreement with Scripture (1902), Amenomori (1961), Shimizu (1961), and Sundberg (1969), who found the rank order of the vowels by larynx height to be /i a u/. Ewan and Krones (1974) found the pattern /a i u/ for their Thai subject, but the /i a u/ pattern for their English and French subjects.

3.3 Pitch and Larynx Movement

The result of a Pearson product-moment correlation analysis of the linear relationship between pitch and larynx height shows that an increase in pitch is only weakly associated with an increase in larynx height (Pearson $r = 0.17546$, $p < .005$). Only 3% ($r^2 = .03079$) of the variation in larynx height can be associated with pitch. This weak positive linear relationship is illustrated in the scatter diagram in Figure 5.

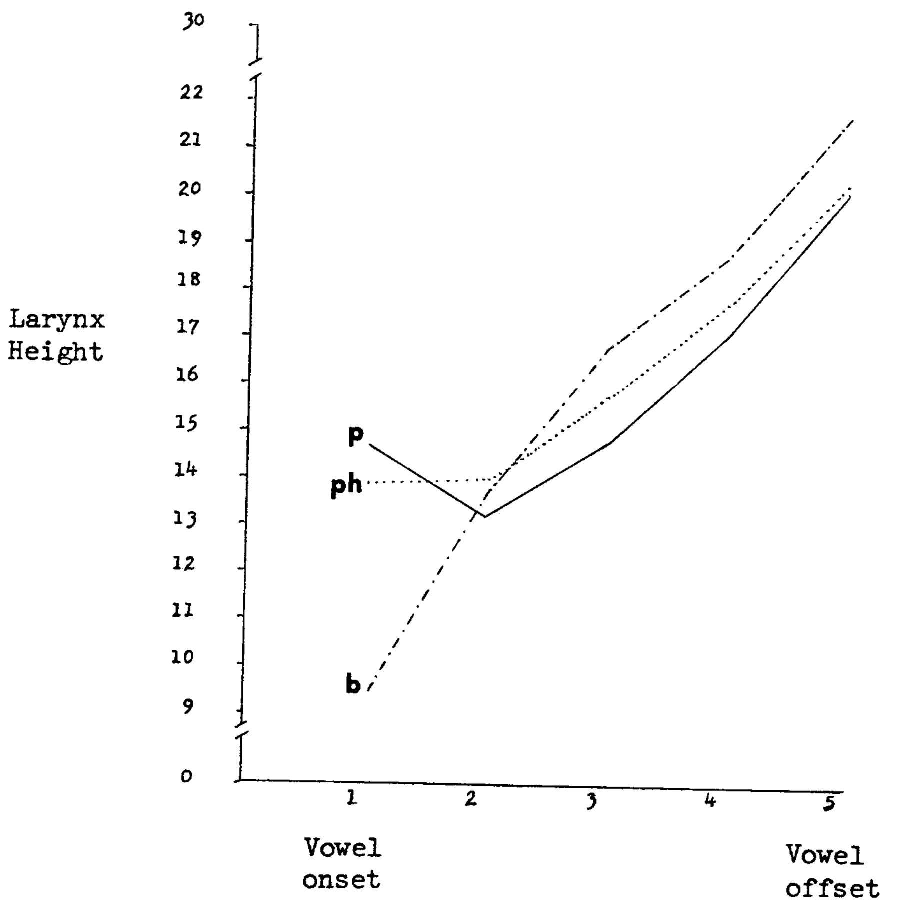


Figure 3. Mean larynx height, following /p/, /ph/ and /b/, during the vowel of the test syllable.

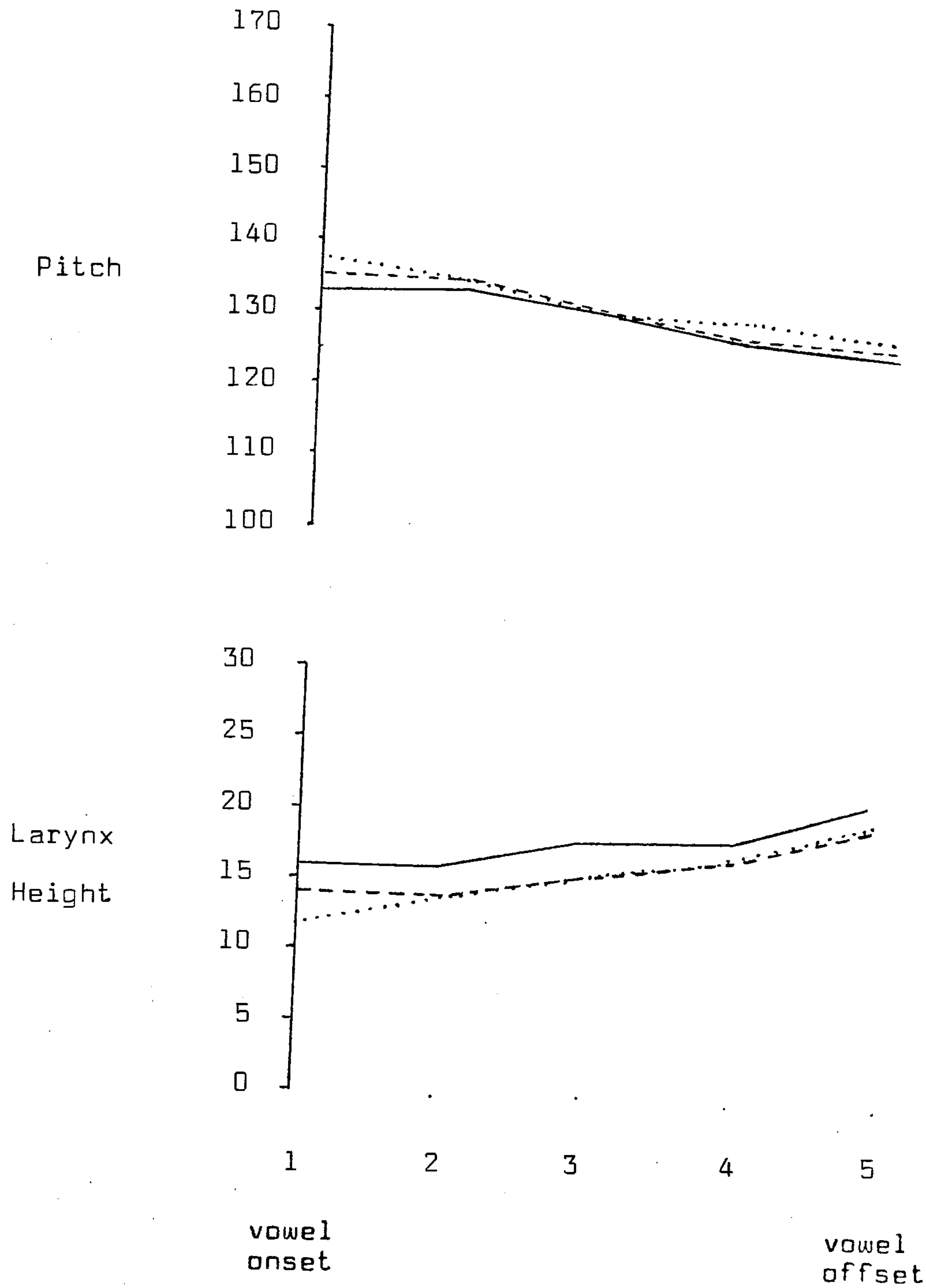
Figure 4 . Mean pitch and larynx height for the vowels
i *a* and *u* in each of the tonal categories.

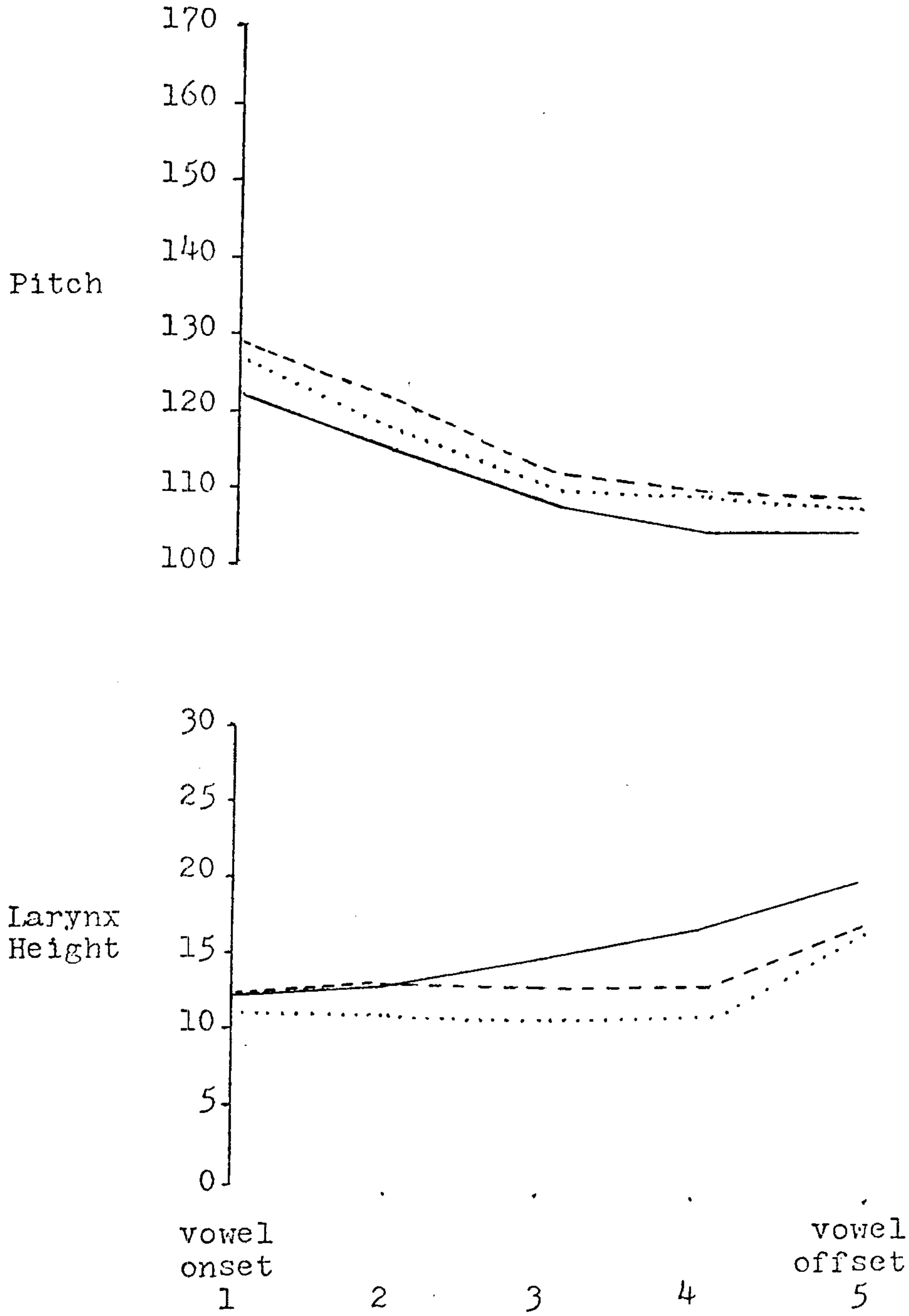
i = - - - - - - -

a = ———————

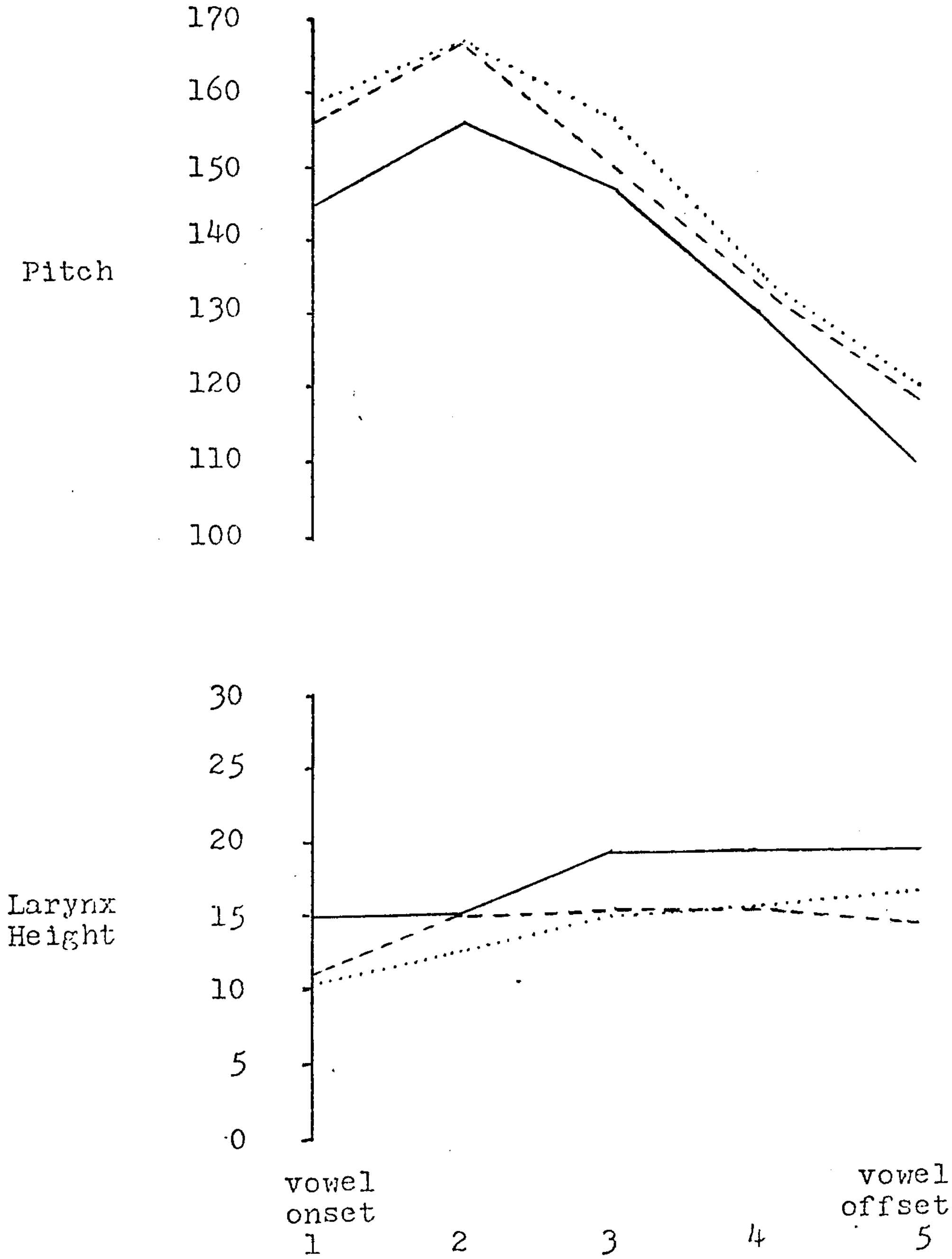
u = · · · · · · ·

(a) MID tone

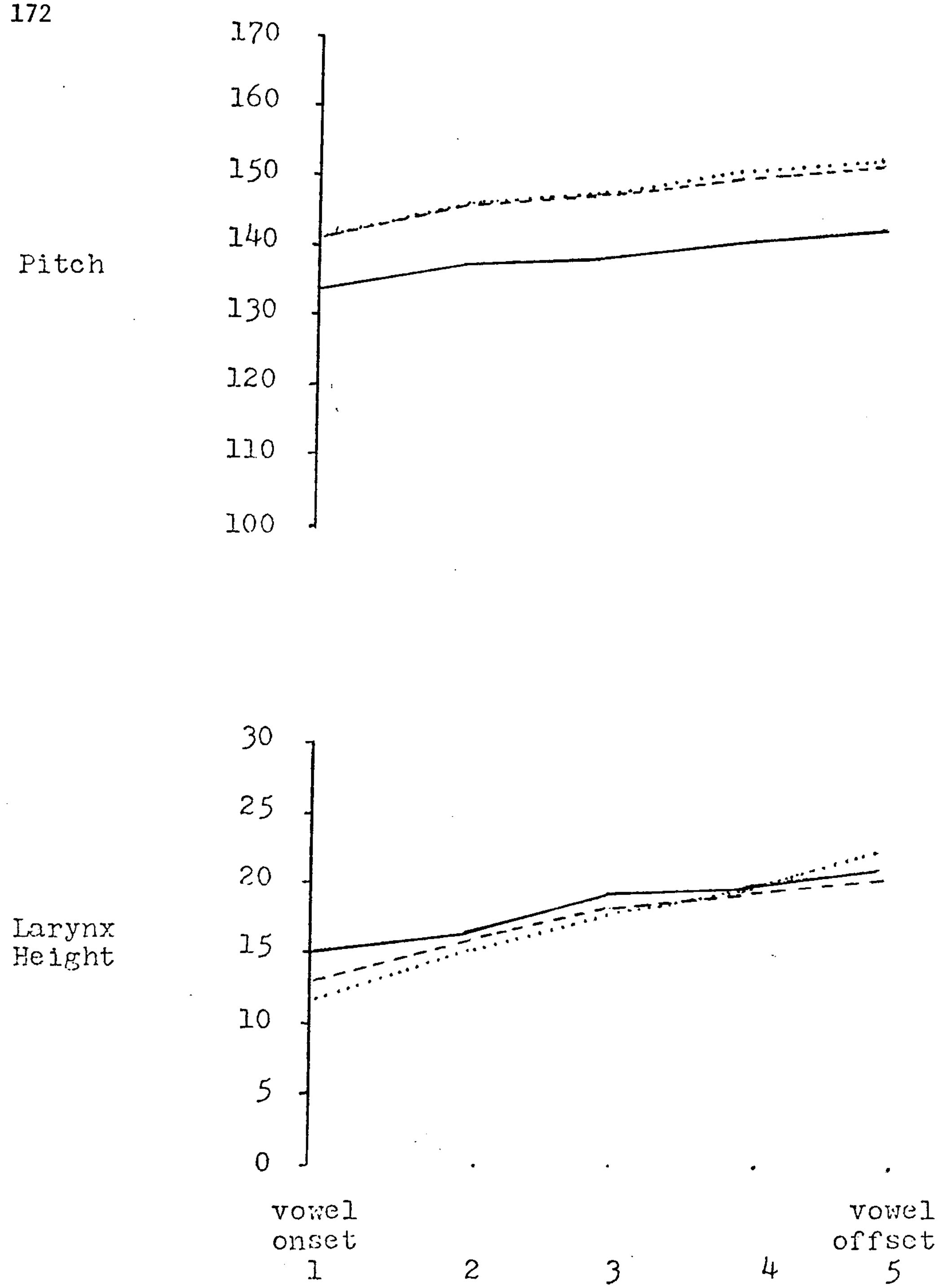



170
160
150
140
Pitch
130
120
110
100
30
25
20
Larynx
Height
15
10
5
0
vowel
onset
1
2
3
4
vowel
offset
5

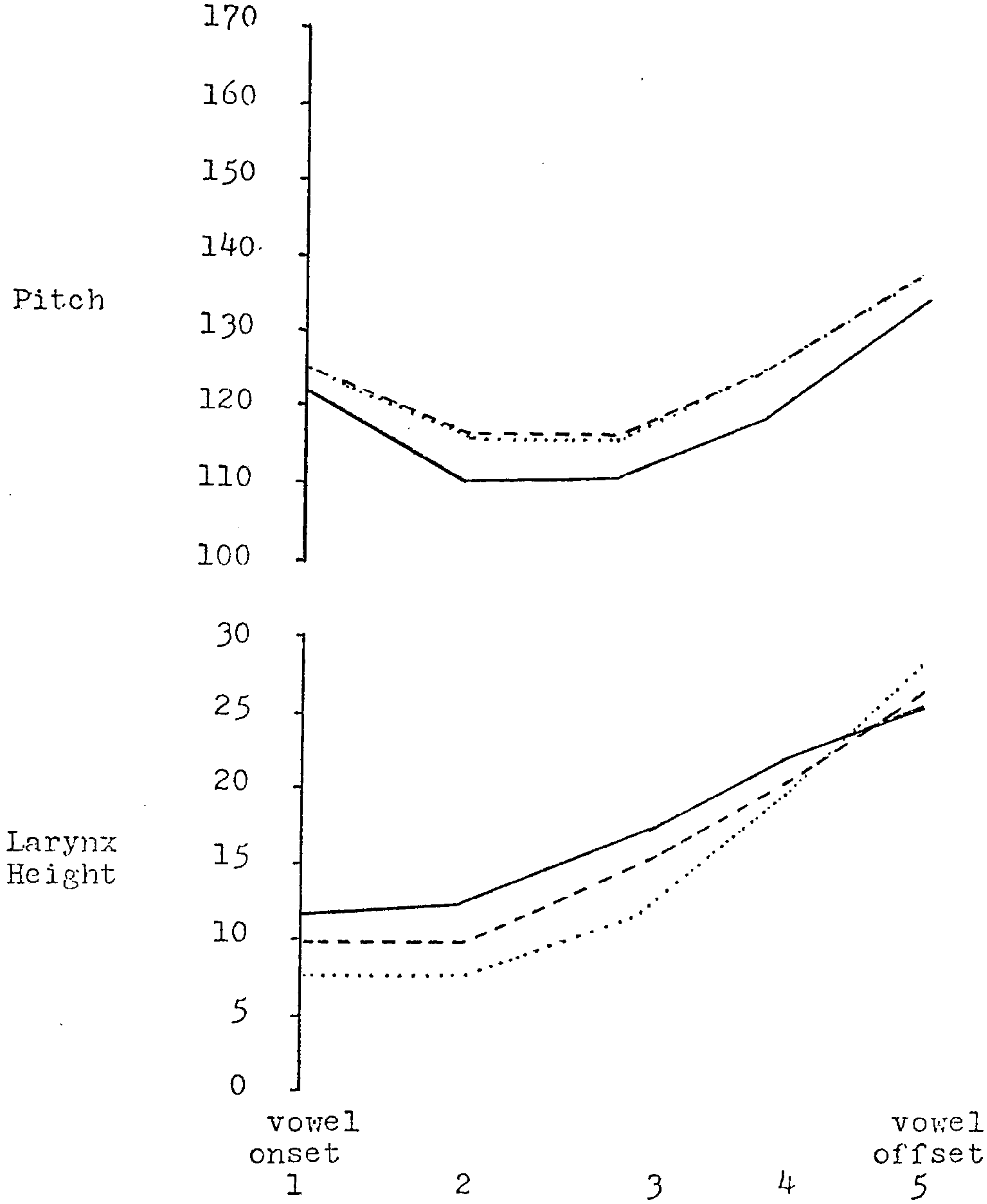
(b) LOW tone



(c) FALLING tone


170
160
150
140
Pitch
130
120
110
100
30
25
20
Larynx
Height
15
10
5
0
vowel
onset
1
2
3
4
vowel
offset
5

(d) High tone



(e) RISING tone

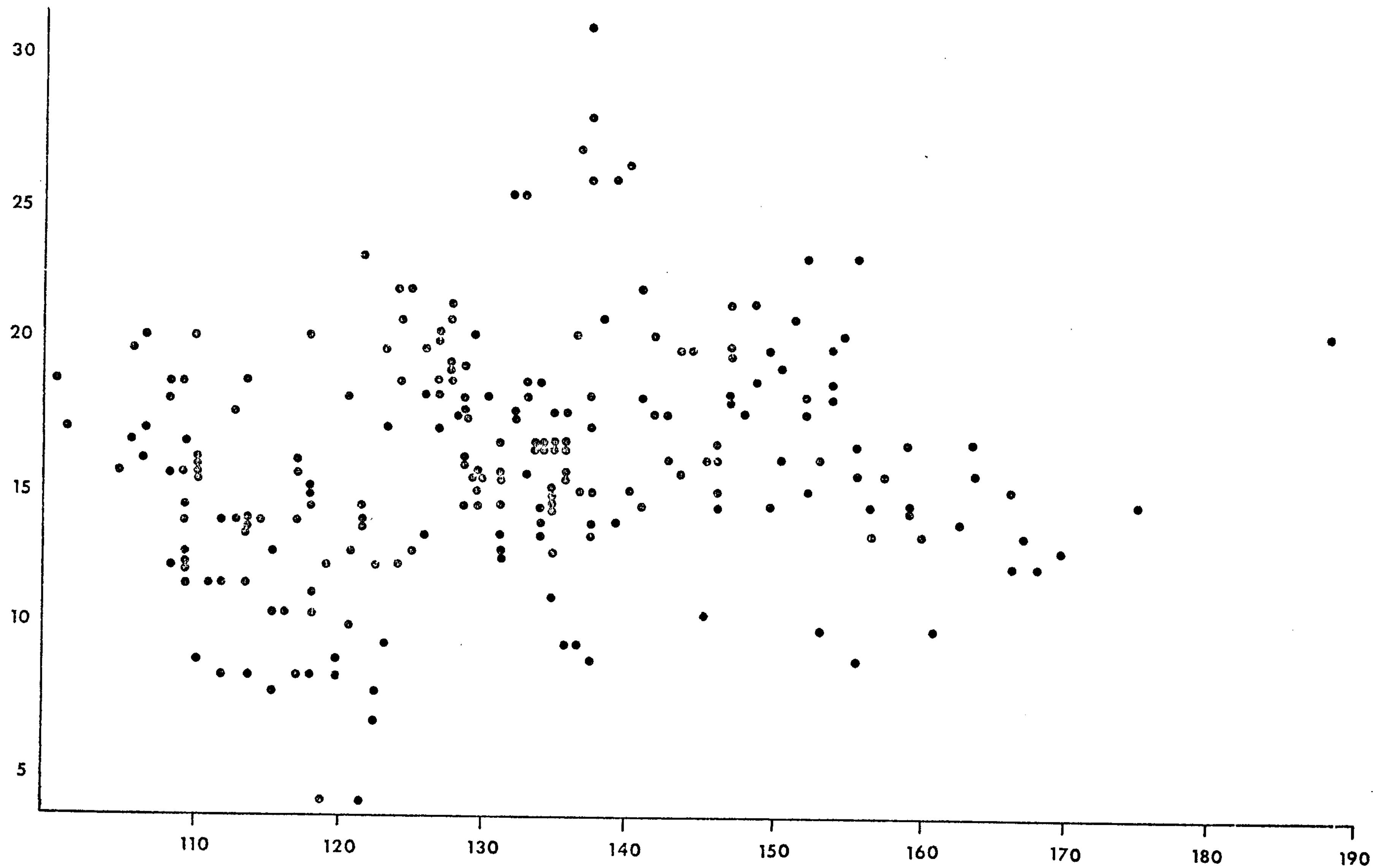


Figure 5. Scattergram plotting larynx **height** (vertical axis) against pitch (horizontal axis) for each test syllable type at five measurement points (225 points plotted). The pitch scale is in Herz.

The graphs plotting pitch and larynx height in each of the 5 Standard Thai tonal categories in Figure 6 further illustrate this weak association between pitch and larynx height. While pitch and larynx height generally move in the same direction on the HIGH and RISING tones, they move in opposite directions throughout the MID and LOW tones. On the FALLING tone the pitch rises and then falls but the larynx rises only.

At vowel onset, pitch (see Figure 7) and larynx height (see Figure 5, par. 3.1) are both lowest following an initial /b/; at measurement point 2, pitch remains lowest after /b/ yet the larynx position is higher after /b/ than after /p/; from measurement point 3 to vowel offset, the larynx position is always highest after /b/ but the pitch is never highest after /b/.

3.4 Standard Thai Tonal Categories and Larynx Movement

Although the present study does not find an overall correlation of pitch and larynx height, there is a general effect of the various tonal categories in the direction of the pitch/larynx height correlation found in other investigations. When the pitch contour associated with the tonal category rises, the larynx position shows a greater rise than when the pitch contour falls. This relationship may be illustrated by plotting the amount of larynx rise and pitch change in the test syllable.[3] Larynx rise was calculated by subtracting the larynx height value at vowel onset from the larynx height value at vowel offset, pitch change by subtracting the most extreme pitch value in the earlier part of the vowel from the pitch value at vowel offset (yielding minus values for the FALLING, LOW, and MID tones and plus values for the HIGH and RISING tones). The results of these calculations are given in Table 3 and represented graphically in Figure 8.

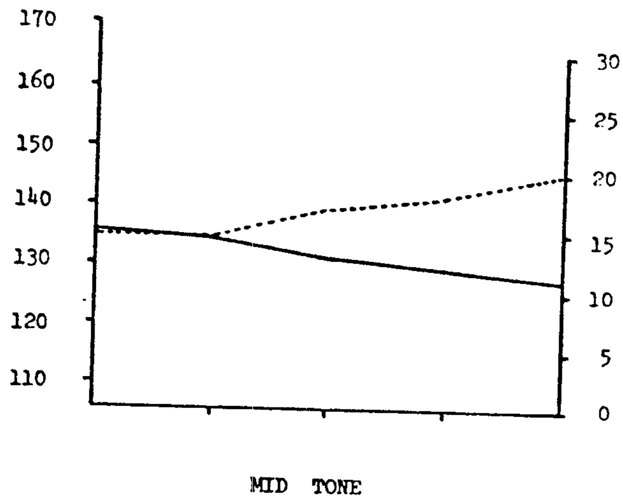
Except for the fact that the amount of larynx rise is the same for MID and LOW tones, the tonal categories are otherwise ranked in the same order in terms of both larynx rise and pitch change. However, when comparing one tonal category with another, the difference in larynx rise is in general not proportionate to the difference in pitch change. For example , the MID and LOW tones have the same amount of larynx rise but the LOW tone has a markedly larger pitch change. Only when we compare the RISING tone with the HIGH tone do we find that the larynx rise and pitch change values are in proportion to each other.

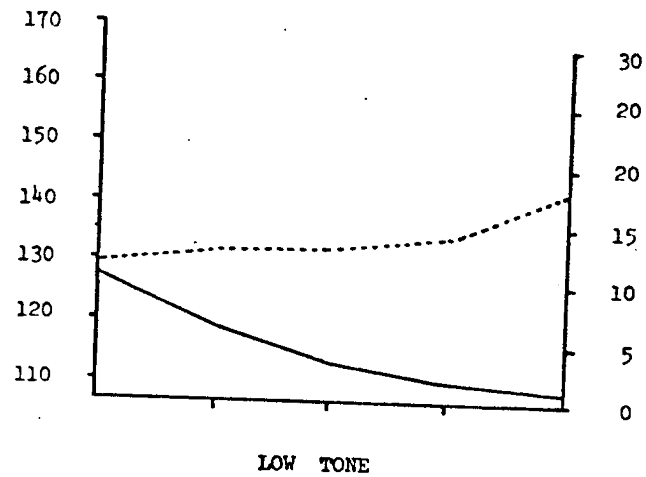
3.5 Relationship of Utterance Position, Phonation Type of Consonant, Vowel Quality, and Tonal Category to Larynx Movement

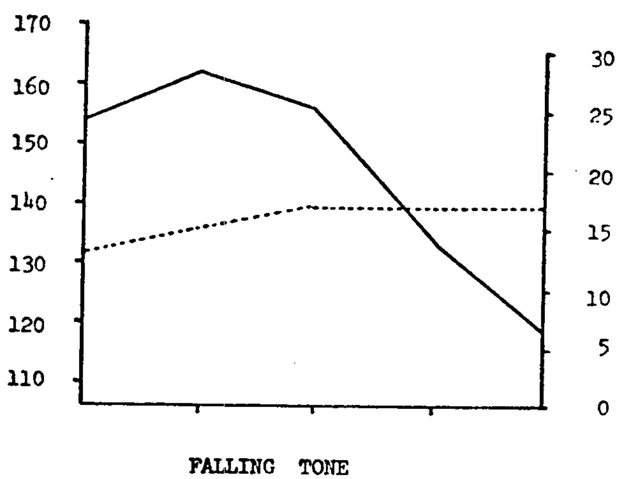
The results of a 4-way analysis of variance (harmonic mean = 5.68) used to investigate the relationship of the following 4 factors--(1) utterance position, (2) consonant phonation type, (3) vowel quality, (4) tonal category--to larynx height are summarized in Table 4.

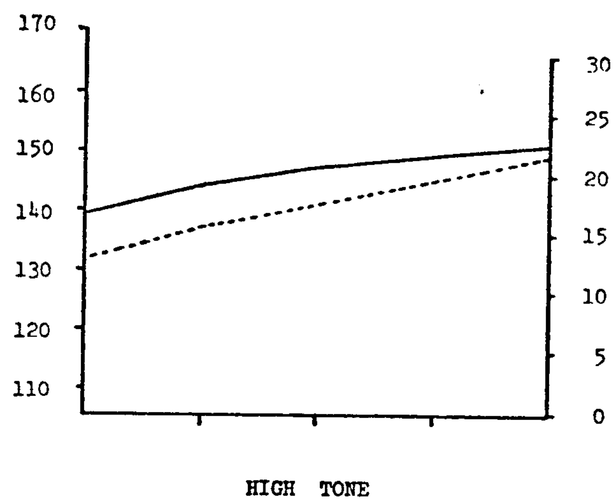
---

[3]We are grateful to Steve Anderson for pointing out how one might plot this relationship between pitch and larynx height in each tonal category.









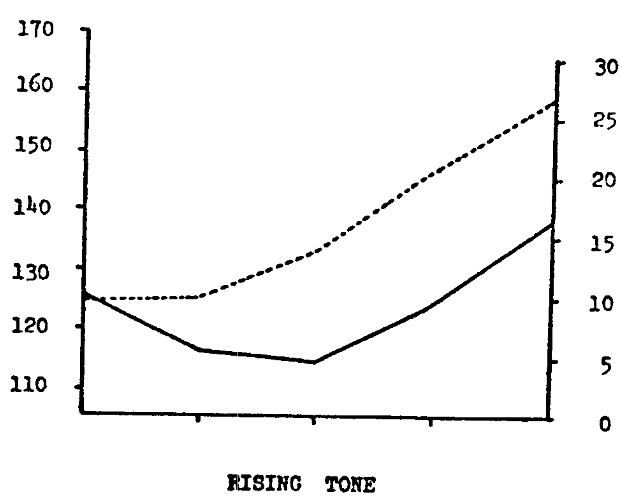


Figure 6: Pitch (continuous line, left-hand scale) and larynx height (broken line, right-hand scale) for each tonal category at five points of time in the vowel. Means of each type, pooled for consonants and vowels.

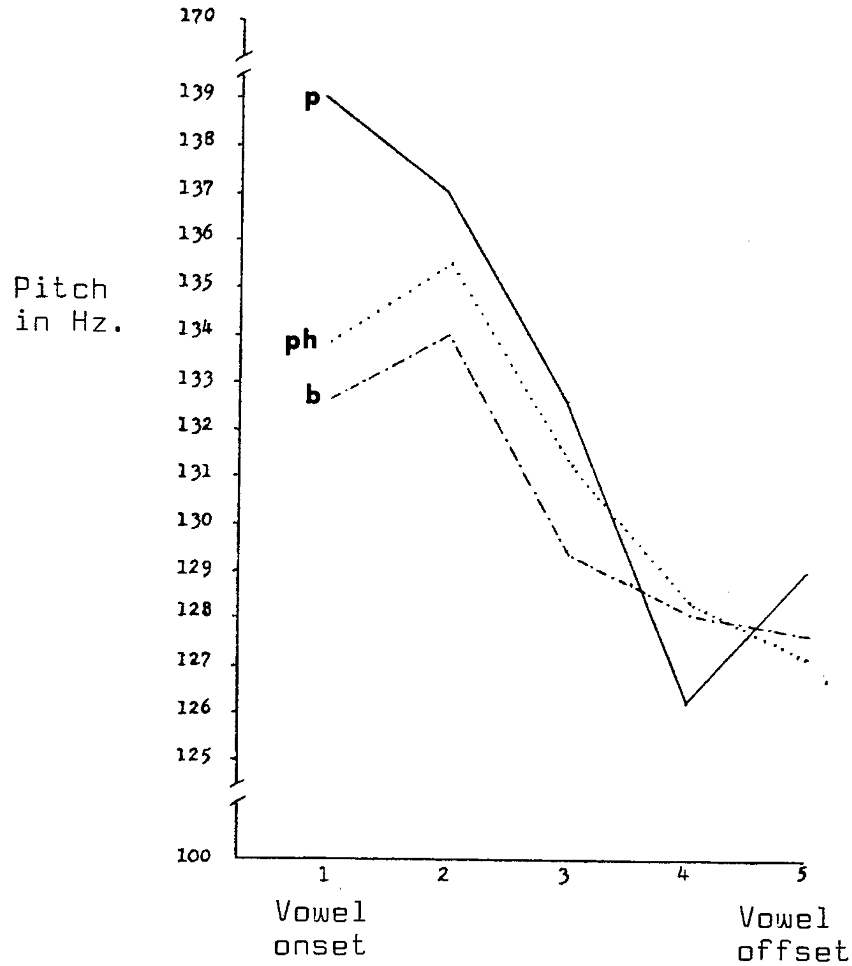


Figure 7. Mean pitch, following /p/, /ph/ and /b/, during the vowel of the test syllable.

Table 3. Larynx rise and pitch change values for each tonal category

	Falling	Low	Mid	High	Rising
Larynx rise	4.2	5.7	5.7	8.1	16.8
Pitch change	-44.0	-18.7	-8.1	10.5	22.4

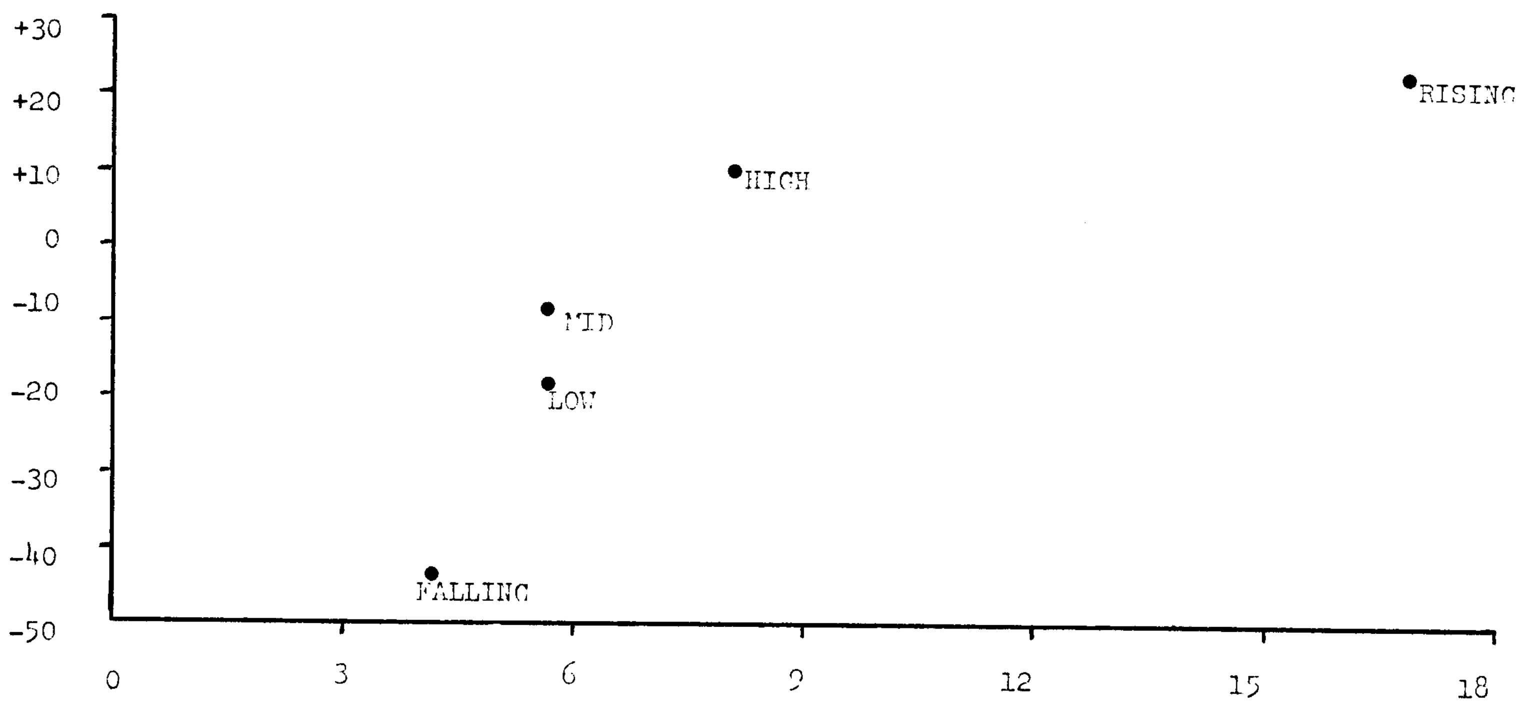


Figure 8. Pitch change(vertical axis) and larynx rise(horizontal axis) across tonal categories.

Variation in the vertical movement of the larynx is significantly affected by each individual factor and each interaction of 2 and 3 factors, but not by the interaction of 4 factors. The amount of variation due to each individual and combination of factors, however, is not the same. Based on the total sum of squares the partitioning of this variation among the individual and joint combinations of factors may be represented diagrammatically as shown in Figure 9.

Utterance position alone accounts for approximately 40% of the variation in larynx height. Another 28.5% of the variation is accounted for by utterance position in combination with other factors. About 17.5% of the variation can be associated with the interaction of utterance position and tonal category, 7% with the interaction of utterance position and phonation type of consonant. Tonal category and vowel quality individually can be associated with 7% and 5% of the variation, respectively. In sum, the position in the utterance stands out as the factor most clearly associated with the variations in larynx movement.

4.0 Discussion

4.1 Overall Pattern of Larynx Movement

The major finding in our study is that the larynx regularly moves upward throughout the utterance. The peaks in this upward movement coincide with the closures of the consonants, the highest peak occurring with the /kh/ or /p/ of the final word /khráp/. The larynx position for the bilabial stops at the beginning of the test syllable is thus lower than for /kh/. This might suggest that the upward movement in larynx position in our sentence-frame is due mainly to the presence of the velar consonant immediately following the test syllable. A few studies (Perkell 1969, Ewan and Krones 1974, Kim 1967) have found that the larynx reaches a higher position for velar consonants than for bilabial consonants. However, the differences in larynx height between bilabial and velar consonants found in these studies seems too small and too inconsistent to explain the large rise in larynx height found in our data.

The data published by Perkell permit the difference between the peaks for the English stops /p/ and /k/ to be calculated. From measurements taken at the front of the larynx, the peak following /k/ is about 1 mm higher than the peak following /p/. The distance that the larynx rises to reach these peaks is about 11 mm. The total distance between the lowest and highest position of the larynx in this study is about 15 mm.

Ewan and Krones also report on larynx height in English stops. In one of their English subjects, they found the mean larynx position for /p/ to be 1.6 mm higher than /k/, and /g/ to be 1.8 mm higher than /b/. Their other four English subjects, however show /k/ to have a higher larynx position than /p/ by 0.7 mm, 0.8 mm, 1.9 mm, and 2.2 mm, respectively. In this study the larynx moves a total of 18 mm vertically in the utterance /utu/ spoken by one of the English subject.

Table 4. Summary of 4-way analysis of variance

Factor	Sum of Squares	Mean Square	F-Test	Significance	Percent of Total Sum of Squares
Position	10821.590	2705.397	736.378	0.001	39.91
Consonant	38.668	19.334	5.263	0.006	0.14
Tone	1979.409	494.852	134.693	0.001	7.30
Vowel	1343.369	671.684	182.825	0.001	4.95
Position x Consonant	1847.010	230.875	62.842	0.001	6.81
Position x Tone	4769.078	298.067	81.131	0.001	17.59
Consonant x Tone	88.484	11.061	3.011	0.003	0.33
Position x Vowel	323.524	40.440	11.007	0.001	1.19
Consonant x Vowel	48.468	12.117	3.298	0.011	0.18
Tone x Vowel	196.406	24.551	6.682	0.001	0.72
Position x Consonant x Tone	284.235	8.882	2.418	0.001	1.05
Position x Consonant x Vowel	189.321	11.833	3.221	0.001	0.70
Position x Tone x Vowel	353.380	11.043	3.006	0.001	1.30
Consonant x Tone x Vowel	113.988	7.124	1.939	0.015	0.42
Position x Consonant x Tone x Vowel	216.807	3.388	0.922	0.500	0.80
Unit	4500.555	3.674	not tested		16.60
TOTAL	27114.270	18.712			100.00

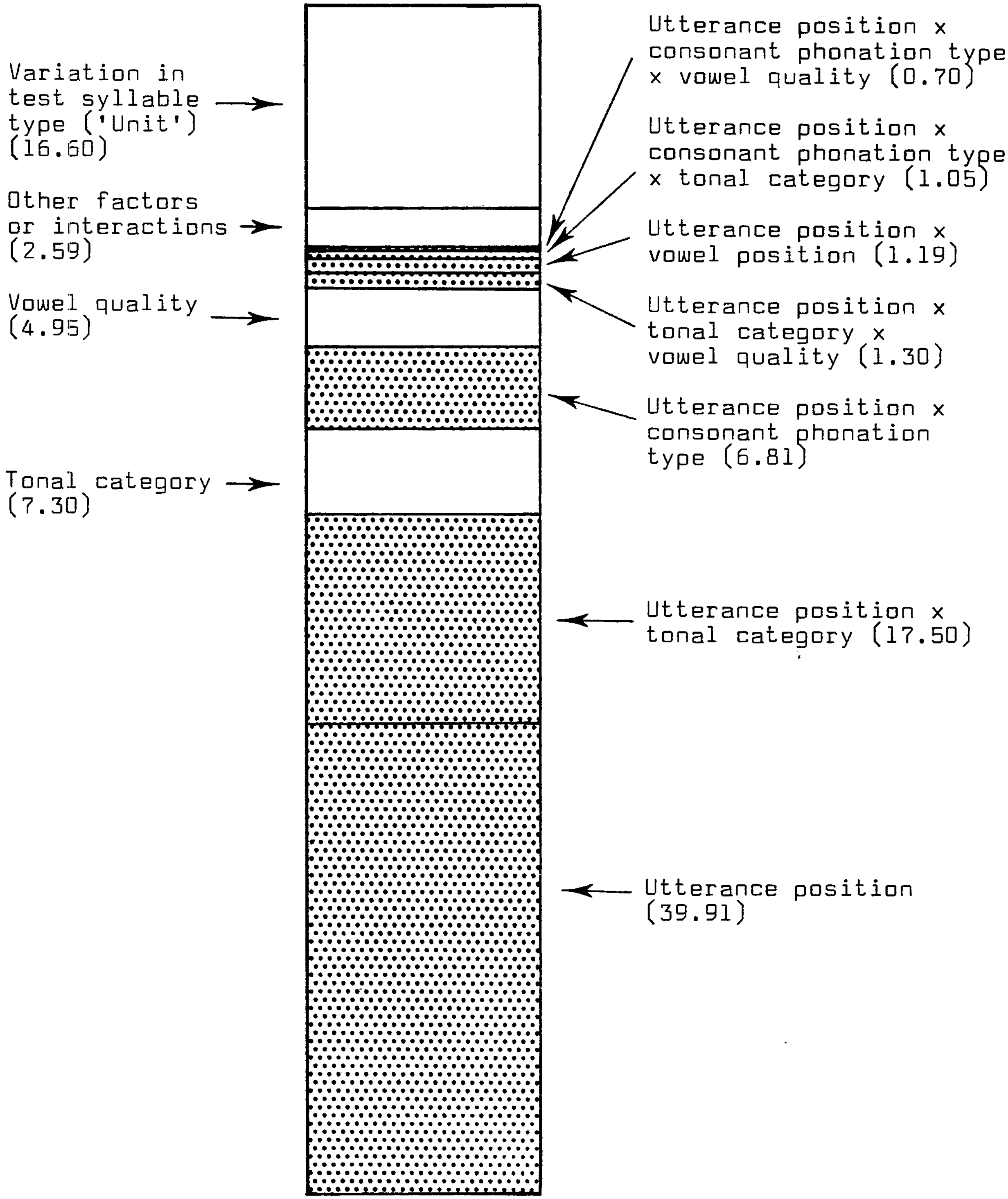


Figure 9. Relative contribution to the larynx position of different factors (separately and in combination) as shown by the percent of sum of squares in the analysis of variance. The shaded areas emphasize the proportion contributed by position in the utterance and combinations of factors including utterance position.

Kim measured the larynx movements over time in the three different stop series of Korean. His data permit the following observations: in the tense unaspirated voiceless series, the peak larynx height for the velar is about 1.5 mm higher than for the bilabial; in the lax unaspirated voiceless series, the velar and bilabial have the same peak larynx height; in the voiceless aspirated series, the peak larynx height for the velar is about 1 mm higher than for the bilabial. Total vertical movement of the larynx in this study ranges over a distance of 12 mm within 100 milliseconds before and after the release of these consonants.

From the results of these studies we can only conclude that the difference in vertical position of the larynx that may be attributed to the difference between the velar and bilabial place of articulation is small. This place of articulation difference would account for only 7% or less of the total larynx movement. Furthermore, we also observe that the distance from the highest position of the larynx to the following lowest position is 5 or 6 times greater than the size of the differences due to the velar and bilabial places of articulation.

In our study, the rising movement of the larynx from the *bilabial* stops at the beginning of the test syllable to the *velar* stop /kh/ is much larger than one could attribute to the difference in place of articulation. For example, the difference between the larynx position at Lx(low) for /p/ and at Lx(kh peak) is 5.1 units, which amounts to more than 17% of the total larynx movement (30 units). In the studies cited above, we found that only 7% of the larynx movement could be attributed to the difference between the velar and bilabial place of articulation.

It is further observed that the difference in larynx position between Lx(high) and Lx(low) is about the same as that between Lx(low) for /p/ and Lx(kh peak). The movement from Lx(high) to Lx(low) for the consonants may be compared to the above-mentioned downward movement that was found to be 5 or 6 times greater than the size of the differences due to place of articulation.

For these reasons, the overall rise cannot be attributed primarily to the difference in the place of articulation of the stops. It may also be noted that a different place of articulation cannot explain why, in a few cases, the final /p/ in /khráp/ has the highest larynx position in the whole utterance.

We may also argue that the overall rise from the bilabial stops to the velar stop /kh/ is not due to differences in phonation type between these consonants, although the size of the rise is affected by phonation type. The interval between the Lx(high) values for /p/, /ph/, and /b/ and the Lx(kh peak) value is given in Table 5.

Table 5. Larynx height interval between bilabial stops and velar stop [k^h].

$p - k^h$	5.1
$p^h - k^h$	5.9
$b - k^h$	10.6

A part of the /ph-kh/ interval may be due to difference in place of articulation, but not to difference in phonation type as both /ph/ and /kh/ are voiceless aspirated stops. A part of the /b-kh/ interval must be due to the phonation type difference as the Lx(high) value for /b/ is 3.2 units lower than that for /ph/. But the /b-kh/ interval is over three times as large as the difference due to phonation type in the bilabial stops (cf. Table 1, par. 3.1). The Lx(high) value for /p/, on the other hand, is slightly higher (0.7) than that for /ph/. We might expect this difference of phonation types to reduce slightly the /p-kh/ interval. Such a reduction is probably of the same order of magnitude as the difference due to a change in place of articulation. In this particular case, the effects of place of articulation and phonation type probably cancel each other out so that the *whole* of the rise in the /p-kh/ interval must be due to some other factor(s). Thus, for the /b-kh/ and /ph-kh/ intervals, at least 5 units must be due to that same factor(s).

The overall pattern of larynx movement in the sentence /nii...khráp/ can be seen as a rising trend to a peak reached in the word /khráp/ followed by a sharp fall. The effects of consonants, vowels, and tonal categories are superimposed on this overall pattern. Both the overall pattern and the superimposed effects can be seen in Figure 2 (see par. 2) on the sample oscillogram. The position in the utterance is by far the most significant contributing factor in determining the height of the larynx (see Table 4 and Figure 9, par. 3.5).

As the domain of the overall pattern of larynx movement is the whole utterance, it is best accounted for by some mechanism whose domain is the same, and moreover, which predicts the overall direction of larynx movement and the position of the fall. It is possible that this mechanism is a modification of a laryngeal movement that occurs during respiration. During normal respiration the larynx lowers when breathing in and rises when breathing out. The rising trend may be the rise that is normally observed during expiration. However, in our data, at the beginning of an utterance the larynx does not start to rise until some time after phonation begins (in fact, it falls initially), and near the end of an utterance the larynx falls toward the 'rest position' before phonation ceases. Thus during these utterances the larynx rise begins later and ends earlier than the outward movement of air. For this reason, the normal upward movement must be regarded as modified during speech. It may even be the case that in speech some other mechanism combines with or replaces the upward movement that occurs in normal respiration. Perhaps some alterations of the position of the larynx are related to maintaining phonation as the volume of air in the lungs decreases. The nature of the mechanism(s) that causes the overall pattern of larynx movement clearly deserves further research.

This overall pattern does not seem to have been commented on in earlier studies of larynx height in speech. However the same rise--fall pattern may be observed in several of the studies which have been interpreted as providing evidence of a correlation between larynx height and pitch. For example, in the utterance tested by Vanderslice (1967), and further analyzed by Ohala (1972), the larynx executes a general rise from the beginning of the utterance to the last stressed syllable and then falls rapidly. As the utterance was a list of three items spoken on the English 'list intonation' with a rise-rise-fall pattern, the pitch curve and the larynx height curve do follow each other. Perkell (1969) also investigated brief utterances with a rise-fall pitch pattern (statement intonation) and these utterances also display a rising-falling larynx movement. But the larynx movement in both these studies is similar to the pattern that we have found to occur on utterances with varying pitch patterns. As Vanderslice and Perkell investigated only one pitch contour, their studies do *not* establish that the relationship between larynx height and pitch is causal rather than coincidental.

The only other published data we know of on complete natural utterances is from Lindqvist (1973). In the Swedish test sentences examined in this study there seems to be neither the overall pattern of larynx movement found in our study, nor a correlation with the pitch pattern. Rather the movements of the larynx are largely predictable from the segments: the larynx is high in vocalic or approximant segments (assuming *j* is a palatal approximant) and low in stop consonants, with minor variations due to place of articulation or phonation type.

The studies of larynx height that used sung material are unsatisfactory as evidence for natural speech, particularly those that used trained singers, such as Curry (1937), because a singer is trained to overcome much of the natural movement of the larynx (Vennard 1967, Sundberg 1969).

4.2 Raised Larynx for Consonants

The raised position of the larynx which occurs during the closure of the stops in our data can be seen in the oscillogram in Figure 2 (see par. 2). There does not seem to be any reason for the articulation of a stop consonant to require an active raising of the larynx. Some factor causing a passive adjustment of larynx height would appear to be a more satisfactory explanation of this phenomenon. We have no persuasive suggestion to offer. But it does seem likely that whatever causes the larynx raising interacts with variations in the size of the glottal aperture, the size of the supraglottal cavity behind the oral obstruction, and the timing and duration of articulatory movements in the oral cavity, which, in turn, cause the differences in the absolute elevation and the timing of the larynx movements that are associated with difference in phonation type and place of articulation.

4.3 Compensatory Larynx Movements

Our study provides evidence that the larynx responds in an elastic way to displacement: the larynx reaches a higher elevation at the end of our test syllables when the initial starting position is lower. Three factors predict lower initial larynx positions: (1) initial /b/,

(2) high back vowel /uu/, (3) LOW or RISING tone compared with the test syllable with other consonants, vowels, and tones. For example, in syllables of the type /bùu/, which meets all three of the criteria above, the value at Lx(low) is 3.7 but at vowel offset the larynx height is 31.2; /páa/, on the other hand, is at the opposite extreme, having an Lx(low) value of 15.0 and a vowel offset value of only 17.6. It thus seems as if the extremely low position of the larynx at the beginning of the /bùu/ syllable is compensated by an extremely high position at the end of the syllable. Such a compensatory adjustment is more likely to be due to a passive elastic response mechanism than an active initiatory mechanism.

4.4 The Production of Pitch

The physical correlate of pitch is the fundamental frequency of the voice--the rate of vibration of the vocal cords. Changes in the fundamental frequency depend on two factors: one (physiological)--the tension of the vocal cords, the other (aerodynamic)--the pressure drop across them. The principal physiological forces controlling tension of the vocal cords are the intrinsic and extrinsic muscles of the larynx.

Tension on the vocal cords can be either in the longitudinal (anterior-posterior) or vertical dimension. Much evidence (Hollien 1960, 1962, Keenan and Barrett 1962, among others) suggests that pitch is controlled principally by adjusting the longitudinal tension of the vocal cords. Our data provides evidence that the longitudinal tension of the vocal cords is not regulated by a mechanism that operates principally by altering the vertical position of the larynx.

Furthermore, vertical tension does not appear to be a major determinant in regulating pitch. Ohala's (1972, 1973) suggestion that active variation in pitch (or at least lowering of pitch) is due primarily to vertical, not longitudinal, tension of the vocal cords seems implausible in light of the results of our study and the alternative interpretaiton of the results of Perkell's (1969) and Vanderslice's (1967) studies.

4.5 Historical Development of Tone

Comparative studies suggest that some languages have developed or elaborated tonal contrasts on vowels depending on differences in phonation type of preceding stop consonants. When such a development occurs, lower tones often appear following the previously voiced series of consonants, and higher tones following the previously voiceless or voiceless aspirated series. Experimental studies suggest an explanation in terms of intrinsic pitch perturbations induced by the prevocalic consonants. Voiced consonants display a *rising* slope in fundamental frequency in transition to the following vowel; voiceless consonants, on the other hand, display a *falling* slope (Lea 1973, Gandour 1974, Hombert 1975). With the loss of the prevocalic voicing distinction these intrinsic pitch perturbations may be reinterpreted as relatively lower tones following the earlier voiced series of consonants and relatively higher tones following the earlier voiceless series.

The phonetic explanations that have been proposed attribute the intrinsic pitch perturbations induced by prevocalic consonants to factors that are either aerodynamic (rate of air flow) or physiological (longitudinal or vertical tension of the vocal cords). Ewan (1973, 1975b) suggests that the vertical position of the larynx associated with the consonant phonation types could be the laryngeal mechanism responsible for the intrinsic pitch differences between voiced and voiceless stops.

Numerous cases of tone splits conditioned by the presence or absence of aspiration as well as voicing of the prevocalic consonants have been reported in the languages of Asia (Brown 1965, Haudricourt 1972, Sarawit 1973). In our data, at vowel onset, there is a comparatively large difference in pitch but only a small difference in larynx height between /p/ and /ph/ (see Par. 3.3). Thus, one cannot attribute the different pitch perturbations following /p/ and /ph/ to different vertical positions of the larynx. This experimental evidence on larynx movement in Standard Thai, coupled with the historical evidence on tone splits in Asian languages, casts strong doubt on vertical larynx position as the physiological mechanism underlying these sound changes. An aerodynamic mechanism, on the other hand, relying on different rates of air flow across the glottis associated with different consonant phonation types that affect the fundamental frequency of phonation, is consistent with the experimental data in this study.

4.6 Tone Features in a Generative Phonology

Maran (1971) and Bird (1971) propose distinctive features of tone defined articulatorily in terms of raising or lowering of the larynx. The weak positive relationship between pitch and larynx height found in the present investigation of larynx movement in Standard Thai fails to support those proposals for tone features based wholly on the vertical position of the larynx. We believe that articulatory correlates of tone features should be more directly related to longitudinal, not vertical, tension of the vocal cords.

Maran and Bird further propose to handle the voicing of consonants as well as tone with the same set of features defined in terms of larynx height. Since we have already found that pitch and larynx height are not strongly correlated, we may conclude that voicing of consonants and tone are controlled by different mechanisms.

4.7 Standard Thai Voiceless Unaspirated Stops and Affricates

The Standard Thai voiceless unaspirated stops and affricates have been described as being pronounced with simultaneous oral and glottal closure and release (Brown 1965) accompanied by a slight upward movement of the larynx (Henderson 1964, Harris 1972). The effect of this glottal articulation, according to Henderson, underscores the contrast between the two series of voiceless stops--the unaspirated series articulated with 'tense' phonation and clear onset of the following vowel. The present investigation of larynx movement in Standard Thai provides relevant data to supplement these earlier descriptions based mainly on auditory impressions.

In our data, the larynx moves upward for /ph/ as well as /p/. True, the larynx moves up higher for /p/ than /ph/ at Lx(high) and vowel onset, but the absolute difference in cricothyrometer units at these measurement points is quite small. These results suggest that upward movement of the larynx is not a distinguishing articulatory property of the unaspirated series. In addition, we also find a comparable upward movement of the larynx for the voiced series.

Both Henderson and Harris point out that these voiceless unaspirated stops are not ejectives because the upward movement of the larynx is too slight to build up sufficient air-pressure behind the oral closure to produce the typical auditory effect of an ejective. This qualification is necessary in their descriptions of this stop series because they posit simultaneous oral and glottal closure accompanied by a slight upward movement of the larynx. But as our data shows, in the articulation of /p/, the larynx is sharply raised. The reason /p/ does not produce the typical auditory quality of an ejective must be because the glottis is *not* closed, contrary to these earlier descriptions, at the release of the voiceless unaspirated stops.

We would rather suggest that the 'tense' quality be attributed to pharyngeal rather than glottal constriction. By positing pharyngeal constriction, we can easily explain the commonly observed 'dark' quality of vowels, especially the high front vowel, following this stop series. A tightly closed glottis would not produce such an auditory effect.

5.0 Conclusion

Our investigation of larynx movement in Standard Thai reveals an overall rise-fall pattern throughout the utterance that strongly suggests the existence of a mechanism capable of spanning the utterance domain. The relative position in the utterance is clearly the most important factor in determining the vertical position of the larynx. This overall pattern of larynx movement occurs on utterances of varying pitch patterns which leads us to conclude that larynx height is *not* a principal factor in regulating pitch in connected speech.

Our study further shows that:

a) /aa/ generally has a higher larynx position than /ii/ and /uu/, this difference tends to diminish towards the end of a syllable.

b) the larynx position for /p/ and /ph/ is higher than that for /b/, the pitch following /ph/ is closer to the pitch following /b/ rather than /p/

c) the peaks in the rising movement of the larynx coincide with the consonants, the magnitude of the larynx rise is far greater than one could attribute to differences in place of articulation and/or phonation type.

References

Abramson, A.S. (1962) The vowels and tones of Standard Thai: acoustic measurements and experiments. *Int. Journ. of American Linguistics*, Publication 20.

Amenomori, Y. (1961) Some studies on the 'suspending mechanism of the larynx'. (in Japanese) *Studia Phonologica* 1.95-110.

Bird, C.S. (1971) Observations on initial consonant change in Southwestern Mande. *Papers in African Linguistics*, ed. by C-W. Kim and H. Stahlke, 163-174. Linguistic Research, Inc.: Champaign, Illinois.

Brown, J.M. (1965) From Ancient Thai to modern dialects. Social Science Press of Thailand: Bangkok.

Chiba, T. and M. Kajiyama. (1958) The vowel, its nature and structure. Phonetic Society of Japan: Tokyo.

Curry, R. (1937) The mechanism of pitch change in the voice. *Journ. of Physiology*, 91.251-258.

Eijkman, L.P.H. (1903) Radiographie des Kehlkopfes. *Fortschritte auf der Gebiete der Röntgenstrahlen* 7.196-206.

Ewan, W.G. (1973) A cross-language study of larynx height in stops and its implications for explaining sound change. Paper presented at 3rd California Linguistics Association Conference, Santa Cruz.

Ewan, W.G. (1975a) Explaining the intrinsic pitch of vowels. Paper presented at the 5th California Linguistics Association Conference, San Jose.

Ewan, W.G. (1975b) Laryngeal behavior in speech. To appear in the *Proceedings of the 8th International Congress of Phonetic Sciences*, Leeds.

Ewan, W.G. and R. Krones. (1972) Laryngeal height during VCV utterances: umbrometer measurements. Paper presented at the 2nd California Linguistics Association Conference, Santa Cruz.

Ewan, W.G. and R. Krones. (1974) Measuring larynx movement using the thyroumbrometer. *Journ. of Phonetics*, 2.327-335.

Gandour, J.T. (1974) Consonant types and tone in Siamese. *Journ. of Phonetics*, 2.337-350.

Harris, J.G. (1972) Phonetic notes on some Siamese consonants. *Tai Phonetics and Phonology*, ed. by J.G. Harris and R.B. Noss, 8-22. Central Institute of English Language, Mahidol University: Bangkok.

Henderson, E.J.A. (1964) Marginalia to Siamese phonetic studies . *In Honour of Daniel Jones*, ed. by D. Abercrombie and others, 415-424. Longmans: London.

Hiranburana, S. (1971) The role of accent in Thai grammar. University of London Ph.D. dissertation.

Hollien, H. (1960) Vocal pitch variation related to changes in vocal fold length . *Journ. of Speech and Hearing Research*, 3.2:150-156.

Hollien, H. (1962) Vocal fold thickness and fundamental frequency of phonation . *Journ. of Speech and Hearing Research*, 5.3:237-243.

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

Keenan, J.S. and G.C. Barrett. (1962) Intralaryngeal relationships during pitch and intensity changes . *Journ. of Speech and Hearing Research*, 5.2:173-178.

Kenyon, E.L. (1927) Relation of oral articulative movements of speech and of extrinsic laryngeal musculature in general to function of vocal cords . *Archives of Otolaryngology*, 5.481-501.

Kim, C.W. (1967) Cineradiographic study of Korean stops and a note on aspiration . *Quarterly Progress Report* (MIT Research Laboratory of Electronics) 86.259-272.

Ladefoged, P., J. DeClerk, M. Lindau, and G. Papçun. (1972) An auditory-motor theory of speech production . *Working Papers in Phonetics* 22.48-75. University of California, Los Angeles.

Lindau, M. (1975) Larynx height in Kwa. Paper presented at the 90th meeting of the Acoustical Society of America, San Francisco.

Lindqvist, J., M. Sawashima, and H. Hirose. (1973) An investigation of the vertical movement of the larynx in a Swedish speaker . *Annual Bulletin of the Research Institute of Logopedics and Phoniatrics* 7.27-34. University of Tokyo.

Maran, L. (1971) Burmese and Jinghpo: a study of tonal linguistic processes . *Occasional Papers of the Wolfenden Society on Tibeto-Burman Linguistics 4*. University of Illinois, Urbana.

Menon, K.M.N. and W.M. Shearer. (1971) Hyoid positions during repeated syllables . *Journ. of Speech and Hearing Research* 14.4:858-864.

Ohala, J. (1972) How is pitch lowered? *Journ. of Acoustical Soc. of Amer.* 52.1:124 (Abstract). Paper presented at the 83rd meeting of the Acoustical Society of America, Buffalo.

Ohala, J. (1973) The physiology of tone . Consonant Types and Tone. *So. California Occasional Papers in Linguistics* 1, ed. by L.M. Hyman, 1-14. Linguistic Program, University of Southern California: Los Angeles.

Ohala, J. (1973) Explanations for the intrinsic pitch of vowels . *Monthly Internal Memorandum* (January) 9-26. Phonology Laboratory, University of California at Berkeley.

Ohala, J. and W.G. Ewan. (1973) Speed of pitch change . *Journ. of the Acoustical Soc. of Amer.* 53.345 (abstract).

Parmenter, C., S. Trevino, and C. Bevans. (1933) The influence of a change in pitch on the articulation of a vowel . *Language* 9.1:72-81.

Ohman, S. (1967) Numerical model of coarticulation . *Journ. of the Acoustical Soc. of Amer.* 41.310-320.

Perkell, J.S. (1969) *Physiology of speech production: results and implications of a quantitative cineradiographic study.* MIT Press: Cambridge, Mass.

Scripture, E.W. (1902) *The elements of experimental phonetics.* Scribner's: New York.

Shimizu, K. (1961) Experimental studies on movements of the vocal cords during phonation by high voltage radiograph motion pictures . *Studia Phonologica* 1.111-116.

Stetson, R.H. (1951) *Motor Phonetics.* 2nd edition. North-Holland Pub. Co.: Amsterdam.

Sundberg, J. (1951) Articulatory differences between spoken and sung vowels in singers . *Quarterly Progress and Status Report of the Speech Transmission Laboratory* 1.33-46. Royal Institute of Technology, Stockholm.

Vanderslice, R. (1967) Larynx vs. lungs: cricothyrometer data refuting some recent claims concerning intonation and "archetypality". *Working Papers in Phonetics* 7.69-79. University of Calif. Los Angeles.

Vennard, W. (1967) *Singing, the mechanism and the technic* (revised edition). Fischer: New York.

Zemlin, W.R. (1968) *Speech and Hearing Science.* Prentice-Hall: Englewood Cliffs, New Jersey.

The intrinsic pitch of vowels and tones in Foochow

Ian Maddieson

[Department of Linguistics,

University of California, Los Angeles]

Introduction

A number of studies have shown that there is a tendency for higher vowels to be uttered with a higher pitch than lower vowels in similar environments (Peterson and Barney 1952, House and Fairbanks 1953, Lehiste and Peterson 1961, Ladefoged 1964, Mohr 1969, Lea 1972, 1973, Hombert 1976). This correlation of higher vowels with higher pitch has generally been understood to result either from physiological or acoustic factors.[1] In both cases the correlation is understood to occur automatically rather than voluntarily. As Mohr (1969) says "Given the basic similarity of the speech mechanism of normal human speakers, intrinsic variations must be universal." Such an intrinsic connection of pitch height and vowel height might be expected to leave its imprint in phonologically significant co-variation of tone and vowel quality.

In recent years several authorities have made the interesting suggestion that a well-known phenomenon in the Northern Min Chinese dialect of Foochow is such a case of co-variation of tone and vowels and can be explained in the light of the experimental evidence referred to above. Curiously, it has been suggested both that an alternation of vowels of differing heights might induce an alternation of pitch, and that an alternation of pitch might induce an alternation of vowels. Wang (1968) extracted vowel alternations in Foochow from the *Outline of Chinese Dialects* (Yuan et al. 1960) and presented them as in (1):

---

[1] For further discussion of the cause of intrinsic pitch differences see Ewan (1975) and Hombert (1976) and references there. For the purposes of my paper it does not matter which of a number of possible explanations may be correct. The important thing is that the pitch differences are an automatic consequence of the production of the vowels. I will therefore avoid advocacy or criticism of the various theories of intrinsic pitch in vowels.

(1) œ → ø → y

o → u

a → e → i

In his view, these alternations operate together with alternations of tone in such a way that "each time a tone with a lower [fundamental frequency] changes to one with a higher [fundamental frequency] the vowel also changes to a higher vowel." He argues that although it is synchronically "more economical to let the [tone] sandhi environment condition the vowel raising rule," historically "it is more likely that the vowel raising brought about the tone sandhi."

On the other hand, Lehiste (1970:79) believes that an alternation of tones preceded the alternation of vowels. On the basis of Wang's observations, she notes that "in the Foochow dialect of Chinese high tones morphophonemically raise vowels from low to mid and from mid to high ... one is led to assume, however, that a stage with allophonic raising of the vowels under high tones may have preceded the present stage." Mohr (1969) comments: "whatever the order of events, explanations of the particular developments in the Foochow dialect have to be based on the intrinsic relationship between segment and tone, tongue height and [fundamental frequency]."

Foochow remains the best-known example of a claimed connection of vowel height and tone.[2] Two explanations have been given for this phenomenon, as in (2):

(2) a. higher vowels can raise tones

b. higher tones can raise vowels

This paper will consider if either of these explanations accounts for the data. It will propose an alternative synchronic view of the vocalic alternations, speculate on their historical origins and propose a rationale

---

[2] Other cases have been proposed. Pilszczikowa-Chodak (1972) suggests that tones of suffix vowels in Hausa are largely predictable from vowel height, but this is only possible if a simplified and abstract account is given, and cannot stand as a historical explanation (cf. Newman, 1975). In Siamese, there are a few instances of words with a high vowel in the first syllable that have a high tone variant, but these words are not representative of a general process and some words with low vowels show the same variance. In Cantonese (and other Yueh dialects) a tone-splitting process in checked syllables which originally had a voiceless initial seems to have resulted from a vocalic distinction, but this involves two groups of vowels of mixed heights. O.Y. Hashimoto (1972) explains it as a lower tone (level 4) emerging with tense vowels, while lax vowels have a higher tone (level 5).

(additional to the one offered by Hombert 1975) for believing that vowel alternations are unlikely to induce tone alternations.

Synchronic alternation

For ease of exposition the synchronic vowel alternations in Foochow will be discussed first. The data are taken from Chen and Norman's (1965) *Introduction to the Foochow Dialect*. This evidence agrees largely with alternations cited from Yuan (1960) by Wang (1968) and those given by Chao (1934) and Egerod (1967, 1971). Some sources imply that more vowels alternate - for the moment this problem will be set aside.

Foochow maintains a contrast between 7 of the 8 traditionally distinguished tone categories of comparative Chinese phonology, tone 4 having merged with tone 6. The pitch contours of the tones, from Norman (1973), are given in (3):

(3)

tone number:	1	2	3	5	6	7	8
pitch :	55	52	22	12	242	<u>24</u>	<u>55</u>

The two ju ("entering") tones, 7 and 8, occur only in checked syllables: the underlining represents this.[3] Egerod (1956) points out that tone 8 can be 'tonemicaly' equated with tone 1, and tone 7 with tone 5. Other views of the tonology of Foochow could be proposed but for the purposes of this paper it is not necessary to choose between them.

The underlying tone of syllables, which occurs in citation and phrase-final positions, alternates with another tone when the syllable is in other positions. Only four tones contrast in these 'sandhi' positions. One of these is a mid-high rise [35] which does not occur in isolation. No syllables in sandhi positions are checked. Thus it can be seen that there is a reduction in contrastiveness in the sandhi positions.

Sandhi alternations of tones 5, 6 and 7 [12, 242, <u>24</u>] (enclosed in the box in (3)) co-occur with vowel changes. These tones change to tones 1, 2, or 3 [55, 52, 22] or the special sandhi tone [35] depending on the following tone. When these tone changes occur, the vowel changes given in (4) also occur. (4) is retabulated from Chen and Norman (1965), and shows that the alternations involve diphthongs rather than simple vowels. Wang's presentation of the data does not draw attention to this fact, as he shows the vowels outside of their context of a second vocalic element. The alternations consist of the modification of the first or of both elements of the diphthong (I), or simplification of the diphthong (II):

---

[3]A five-point pitch scale is used to represent the tones; the highest level is 5, the lowest 1. Gliding pitches are represented by juxtaposition of different numbers, level pitches by repeating the same number.

(4) I. a) ai → ei

b) au → ou

c) oi → øy

II. a) ei → i

b) ou → u

c) øi → y

The notation employed is a 'broad phonetic' one which serves to indicate probable underlying forms.[4]

It can be seen that the alternations in I involve the assimilation of properties of the component parts of the diphthong. Ia) involves /a/ moving from low to mid when /i/ follows; Ib) involves /a/ moving from low to mid also, but in addition the rounding and backness of the second vowel spread to the first element. The alternation in Ic) is similar, since /i/ → /y/ under the influence of the preceding rounded vowel /o/, whereas /o/ → /ø/ under the influence of the following front vowel /i/. The alternations in I show that three kinds of modification occur. Since these involve the spreading of the property of a given segment to its neighbour, these are justifiably regarded as assimilatory processes. The three modifications can be classified as in (5):

(5) 1. assimilation of vowel height upwards

2. assimilation of rounding

3. assimilation to position on the front/back axis.

The alternations in II can also be regarded as cases where assimilation occurs, producing a sequence of identical vowels. IIa), b) and c) require an assimilatory raising of mid vowels to high vowels somewhat similar to the raising of low to mid in Ia) and b). Note that in Ic) /oi/ does *not* raise the first element, presumably in order to maintain

---

[4] I have substituted /øy/ for Chen and Norman's broad /oi/ (narrow IPA [œy̌]) in order to illustrate more clearly the assimilation of rounding and the parallelism of I and II. Note that Chao (1934) would reverse the direction of the changes shown in (4). He regards the process involved as a lowering or diphthongization of vowels in the environment of lower tones. This view is not further discussed as the etymologically and lexically basic tones are the other members of the pairs which alternate in sandhi than those which he assumes. M.J. Hashimoto (1972) also discusses the vowel alternation as a lowering, talking of "the change [e] to [a], [ø] to [ɔ] and [o] to [ɒ]" when a tone starts with a nonlow pitch. The motive seems to be that if the change is read in this direction it is easier to state an environment in which it occurs.

the lexical contrast /oi/ ~ /øi/ which would otherwise be lost (compare Ic) with IIc)). Similarly the outputs of Ia) and b) do not undergo the further raising processes in IIa) and b) to which they could serve as inputs. IIc) shares with Ia) and b) the spread of rounding. The alternations in II can therefore be presented as assimilatory processes, as in (6):

(6) a) ei → ii

b) ou → uu

c) øi → yi → yy

In this light the alternations of vowels in Foochow seem like familiar assimilation and coalescence processes known both from well-known languages like Sanskrit and Greek, and obscure ones like Central Kambari (Hoffman 1972) and Owon Afa (Awobuluyi 1972 & cf. Maddieson 1972). Some samples of vowels alternating in Kambari and Sanskrit are given in (7):

(7)

Central Kambari	Sanskrit (adapted from Benfey 1868)
a + i: → ee	ai (→ ei) → ee
a + o: → oo	au (→ ou) → oo
a + u: → oo	
ɔ + u: → uu	

The vowel alternations in Foochow are understandable, as assimilation often is, in terms of an economy-of-effort principle. They involve a reduction in the distinctness of the articulatory gesture required to produce the diphthongs in their full forms. This less distinct articulation occurs naturally in the less prominent non-final position of syllables. This is the position in which tone sandhi occurs - which also reduces the contrastiveness of the syllables (from 7 syllable types to 4). Tone sandhi and vowel assimilation can be seen as aspects of the same reduction of contrastiveness.

Synchronically, Foochow vowel alternations seem most naturally explained as assimilative. Vowel height is only one of three parameters of the assimilation (see (5) above), and all three of the usual parameters along which vowels are distinguished take part. We can therefore see that an intrinsic connection of vowel height and tone could at best only be partially successful at accounting for the observed alternations, even before we examine the tone levels involved and consider the possible historical connection of vowel and tone.

Historical origin

The presentation of synchronic alternations has been simplified by limiting consideration to diphthongal cases. Some writers (e.g. Mohr 1969, Egerod 1956, 1971) imply that Foochow simple vowels also have synchronic alternating forms; others imply that simple vowels do not alternate (Wang 1968, Egerod 1967, Chao 1934). Perhaps these disagreements result from the fact that some accounts only record alternations if "phonemic overlapping" is involved, while others record subphonemic alternations also. An explicit distinction of this kind seems to be made in Chen and Norman (1965). Whatever the synchronic picture may be, it is generally agreed that the alternations are the result of a 'splitting' of vowels in syllable types defined by the traditional labels for Chinese tonal categories, and that this historical process affected both diphthongs and simple vowels (Karlgren 1915-1926, Lan 1953, Forrest 1948, M. J. Hashimoto 1972, etc.). This section of this paper evaluates whether the historical process involved can be explained by an appeal to the intrinsic pitch of vowels.

From a historical perspective, Foochow shows that syllables that are reported to rhyme in Middle Chinese (= Karlgren's 'Ancient Chinese') have two different vocalic reflexes (and sometimes different final consonants) according to the tonal categories they occur in.[5] Although Middle Chinese is not directly ancestral to the Min group of dialects, it nevertheless seems correct to talk of vowel splitting as far as these dialects are concerned. In Foochow, one set of vocalic reflexes occurs with tones 5, 6 and 7 (those in the box in (3)), and another set of reflexes with the other tones. These sets of tones and vowels can be conveniently labelled Set I and Set II respectively. The alternation of vowels under sandhi conditions is a consequence of the historical vowel split, as Set I tones are replaced by Set II tones in sandhi environments.

Some other closely related Northern Min dialects have even more extensive splitting of vowels and they have preserved more distinctions among finals. It seems likely that one and the same historical process has produced the vowel splits in these related dialects.[6] The tones of three dialects are shown in the table in (8); Foochow, and the closely related

---

[5] Early Chinese poetry does not usually rhyme syllables belonging to different tone categories, but the rhyme tables such as the Ch'ieh Yun are compiled to show the similarities between syllables that are alike except for the difference of tonal category. Therefore each rhyme category is assumed to include only syllables with the same nuclear vowel and post-nuclear segments in non-checked syllables.

[6] At least one Southern Min dialect, Swatow (Chaochow), has some differentiation of vocalic reflexes in alternants of tone 5 (Egerod 1956). Among Mandarin dialects, Pekingese also has a trace of a somewhat similar vowel splitting (Forrest 1948, Egerod 1971). These splits do not seem to be due to the same process as in Northern Min but have a separate historical origin.

Ningteh and Fuan dialects from Norman (1972) - column 4 has been're-constructed' by inference.[7] The jagged line divides the sets of tones.

(8)

Set II "higher"

tone number:	1	2	3	4	5	6	7	8
Foochow	55	52	22	242	13	242	24	55
Ningteh	33	11	41	41	35	41	33	55
Fuan	43	11	41	?41	35	13	33	55

Set I "lower"

It can be seen that tones 4, 5, 6 and 7 form a block in Set I for all three dialects. Ningteh adds tone 2 to Set I; and Fuan also adds tones 1 and 8 to Set I.

If the theory, advanced by Lehiste (1970), that higher tones caused higher vowels is to be upheld then the tones of Set I must be shown to be lower, or derive from lower tones than set II. Examining the current tonal values reveals that Set I tones are not now lower than Set II tones. In Foochow tone 3 [22] cannot be considered higher than tone 6 (and 4) [242] or tone 7 [24]. In the related Ningteh and Fuan dialects the distinction between the tone sets is even less clearly one of higher vs. lower tones. In Ningteh tones 3 and 6 are the same, [41], but tone 3 is in Set II and tone 6 is in Set I. Tone 1 [33] is in Set II, tone 7 [33] is in Set I; and tone 5 [35] is in the set which should be lower while tone 1 [33] is in the "higher" Set II. In Fuan, Set II consists only of tone 3 [41] whereas tone 1 [43] is in the "lower" Set I. Simple inspection of the modern tones in the different sets does not support the view that higher pitch levels induced the splitting of vowels.

Of course this does not rule out the possibility that earlier values for tones divided the pitch range in a different way. A simplified reconstruction of Proto-Min tones would establish four syllable types from each of which two tones have developed; pairs 1/2, 3/4, 5/6 and 7/8 each having one antecedent syllable type. Norman (1973, 1974) has traced a fairly complex set of correlations of reconstructed syllable initial consonants with the modern tone reflexes in Min dialects, but has offered no reconstructions of the pitch values of the tones. The task of reconstructing intermediate tone levels which might provide a correlation of tone height with the vowel sets, stand as plausible antecedents for the present tonal values, and lead back to the four syllable-type contrast of Proto-Min is complex. The three dialects agree that tones 4, 5, 6 and 7 occur with Set I vowels. Thus the pairs 3/4, 5/6 and 7/8 are involved.

---

[7] Norman differs slightly in the transcription of tone 5 in various sources making it either 12 or 13.

The modern tone 7 values are lower in pitch than tone 8, and a low original pitch can be posited for the antecedent of the pair 5/6. But the modern tone 4 values are not lower than tone 3 - indeed the pair 3/4 does not seem to have split in Ningteh or Fuan (though Norman's comment on Fuan is ambiguous). It therefore seems unlikely that the distinction of vowel reflexes between tones 3 and 4 is due to a lower pitch for tone 4. There are also difficulties in the way of understanding the extension of vowel splitting to other categories in Fuan and Ningteh in relation to pitch height. The impression is that pitch height distinctions cannot be responsible for inducing the vowel splits.

Two further observations reinforce the evidence from comparison of tones that a correlation of pitch height with vowel height is not responsible. In the first place, not all the paired vowels differ in height; and, among those that do, the height difference is frequently not the only difference. As noted above, the Foochow alternation /oi/ - /øy/ does not involve change in vowel height. And in Ningteh /i:/ alternates with /E/, a lowering, and both /u:/ and /ɔ:/ alternate with /o/ - one is a lowering, the other a raising. Secondly, the experimental work cited above does not provide evidence that the existence of pitch differences can lead to vowel splits. This work shows that, given vowel height differences, pitch differences might be the consequence. We may therefore reject the suggestion that Foochow provides an example of pitch differences inducing vowel height differences.

However, Wang's proposal, supported by Mohr, is rather that the vowel alternations are historically prior to the tonal differences. Wang proposed this specifically to explain the alternation of tones in the sandhi process. But the difference in vowel quality in Foochow accompanies the different tones whether these are lexical or derived from sandhi. An explanation must cover all cases, including the lexical ones, not just the alternations under sandhi conditions. The comparative evidence shows that rhyme classes with the same vowel in Middle Chinese are split into two reflexes in Northern Min; whereas rhyme classes with different vowels retain their membership in the same tone class in Middle Chinese. If vowel differences had induced pitch differences, the tone classes would be divided according to the vowels; e.g. *pim having a higher tone than *pam. This is not the way the correspondences work. Unless comparisons with Middle Chinese are more misleading than is usually believed, the Foochow tones do not seem to have been influenced by vowel height.

Wang's suggestion that tone sandhi is a consequence of the vowel alternation in Foochow is not only countered by the occurrence of the vowel variants in lexically distinct sets of tones, but also because some of the tone alternations in Foochow do not co-occur with vowel alternation, and because tone sandhi occurs in all varieties of Chinese. Those forms with Set II vowels and tones in Foochow do not change their vowels in sandhi, but do change the tones (in most environments). The Foochow sandhi rule [<u>55</u>] → [22], which applies to tone 8 (the tone in checked syllables with an originally voiced consonant) and involves no vowel change, can be reconstructed for an early form of Min, as the same rule occurs in Southern

Min varieties such as Amoy, Gaoxiong and Hainan. Other sandhi rules could similarly be reconstructed. By contrast, the vowel splitting in Northern Min has a much narrower distribution and appears as a local innovation.

I would speculate that the vowel splitting originates from differences in the nontonal environment. The four 'tone categories' of Proto-Min and Middle Chinese are syllable types of which some were not distinct by pitch (or not by pitch alone) but by the nature of the syllable offset. These categories were split according to various distinctions in the syllable initial consonants, so that the modern dialects typically have more than four categories. We know that adjacent consonants can affect vowel quality (e.g. 'breaking' in Old English), and in Northern Min distinct initial and/or final consonants distinguish those tonal categories in which different vowel reflexes occur. Tone sandhi seems to have three phonological sources; de-stressing, assimilatory changes among tone levels, and tonal divergence resulting from 'consonant sandhi' changes. There may also be morphological reasons for tone change in sandhi (e.g. suffixation or other morphological marking of the relationship between tye components). Destressing and consonant sandhi probably explain the co-occurrence of tone sandhi with vowel alternation.

In short, the appeal to the intrinsic pitch of vowels does not explain developments in Foochow. Other factors are more plausibly responsible for these developments.

Conclusion

This paper demonstrates that the best example of a claimed influence of an intrinsic association of vowel height and pitch on the fules of language - the Foochow case - does not stand up as a synchronic or historical explanation. While other examples of this influence may await discovery, the present situation raises the question whether vowel height differences can ever induce linguistically significant pitch variation. (It has not been shown that there is any intrinsic reason for pitch differences to result in vowel height differences.) But the experimental evidence of intrinsic variations in the pitch of vowels does not necessarily suggest that these variations can become significant. It may be that they function only as perceptual cues to the identification of vowel quality, or are simply negligible. The scale of the pitch differences between vowels of adjacent degrees of vowel height is very small. The data presented by Hombert (1976) illustrate this well. Using two Yoruba subjects speaking in the same style, Hombert finds something around 4 Hz as the maximum difference between adjacent vowel heights or about 17% of the difference between the mean pitch of different tones; whereas pitch following voiced and voiceless consonants starts at points some 30 to 40 Hz apart or over 120% of the difference between tones.[8] Hombert points to the dynamic nature of post-

---

[8]Hombert's results are especially interesting in showing that with low tone vowels a very different correlation of pitch and vowels occurs than

consonantal pitch effects to explain why they do yield extrinsic results when the static vowel pitch differences do not. The absolute scale of these differences is probably also relevant to the lack of cases of 'phonologization' of the intrinsic pitch of vowels.

It has often been assumed that the step from non-significant (extrinsic) variation is a straightforward one. I do not believe that all universal properties of the speech mechanism are equally likely to make this step. It is important to build a typology of 'extrinsifiable' and 'non-extrinsifiable' intrinsic phonetic variations. As a first contribution, I offer (9):

(9) Intrinsic pitch of vowels is a universal that does not lend itself to extrinsic exploitation.

References

Awobuluyi, O. (1972) 'The morphophonemics of Owon Afa' *Research Notes* University of Ibadan, Nigeria 5.2-3:25-44.

Benfey, T. (1868) *A Practical Grammar of the Sanskrit Language* (2nd. edition), Trubner, London.

Chen, L. and J. Norman (1965) *An Introduction to the Foochow Dialect* San Francisco State College in co-operation with the U.S. Office of Education.

Chao, Y-R (1934) 'The non-uniqueness of phonemic solutions of phonetic systems' *Bulletin of the Institute of History and Philology, Academia Sinica* 4.4:363-397 (reprinted in *Readings in Linguistics I* ed. M. Joos pp. 38-54, American Council of Learned Societies, Washington D.C. 1957).

Egerod, S. (1956) *The Lungtu Dialect*, Munksgaard, Copenhagen.

Egerod, S. (1967) 'Dialectology' in *Current Trends in Linguistics 2: Linguistics in East Asia and South East Asia* (ed. T. Sebeok) Mouton, The Hague, pp. 91-129.

---

with mid and high tone vowels. It is too soon to say if this is a universal pattern. Ladefoged (1964) is silent about anything similar in Itsekiri, which is close enough to be considered a dialect of the same language as Yoruba. Incidentally, of the languages investigated, Yoruba and Itsekiri, which are tonal, show the lowest figures for intrinsic pitch variation, English shows the highest, while the quasi-tonal Serbo-Croatian shows intermediate figures (Ivic and Lehiste 1963, quoted in Lehiste 1970).

Egerod, S. (1971) 'Phonation types in Chinese and South East Asian languages' *Acta Linguistica Hafniensa* 13.2:159-171.

Ewan, W.G. (1975) 'Explaining the intrinsic pitch of vowels' Paper presented at the 5th California Linguistic Association Conference, San Jose.

Forrest, R.A.D. (1948) *The Chinese Language* Faber, London.

Hashimoto, M.J. (1972) 'The linguistic mechanism of flip-flop' *Unicorn* 10:1-19.

Hashimoto, O-k.Y. (1972) *Studies in Yue Dialects I: Phonology of Cantonese* (Princeton-Cambridge Studies in Chinese Linguistics III), Cambridge University Press, Cambridge.

Hoffmann, C. (1972) 'A note on vowel contraction in Central Kambari' *Research Notes* (University of Ibadan, Nigeria) 5.2-3:73-92.

Hombert, J-M. (1976) 'Consonant types, vowel height and tone in Yoruba' in this volume.

House, A.S. and G. Fairbanks (1953) 'The influence of consonant environment upon the secondary acoustical characteristics of vowels' *Journal of the Acoustical Society of America* 25:105-113.

Ivic, P. and I. Lehiste (1963) 'Prilozi ispitivanju fonetske i fonološke prirode akcenata u savremenom srpskohrvatskom književnom jeziku' *Zbornik za Filologiju i Lingvistiku* 6:33-73.

Karlgren, B. (1915-1926) *Etudes sur la Phonologie Chinoise (Archives d'Etudes Orientales*, various issues) Lundell, Uppsala.

Ladefoged, P. (1964) *A Phonetic Study of West African Languages: An Auditory-Instrumental Study*, West African Language Monographs 1. 2nd. edition 1968, Cambridge University Press, Cambridge.

Lan, Y-h. 'Fo-chou yin-hsi' [phonetic system of the Foochow dialect] *Bulletin of the College of Arts, National Taiwan University* 5:241-331.

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

Lea, W.A. (1973) 'Segmental and suprasegmental influences on fundamental frequency contours' *Consonant Types and Tone* (ed. Larry M. Hyman) University of Southern California, Los Angeles, pp. 15-70.

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

Lehiste, I. and G.E. Peterson (1961) 'Some basic considerations in the analysis of intonation' *Journal of the Acoustical Society of America* 33: 419-425.

Maddieson, I. (1972) 'Some tentative general claims on vowel contraction and assimilation' *Research Notes*, University of Ibadan, Nigeria, 5.2-3: 103-112.

Mohr, B. (1969) 'Intrinsic variations of Acoustical parameters of speech sounds' *Project on Linguistic Analysis Reports*, 2nd. series 9: M1-M44. Revised version published as 'Intrinsic variations in the speech signal' *Phonetica* 23: 65-93 (1971).

Newman, P. (1975) 'The non-correlation of tone and vowel height in Hausa' *Studies in African Linguistics* 6.2: 207-213.

Norman, J. (1972) 'A preliminary report on the dialects of Mintung' *Unicorn* 10: 20-35.

Norman, J. (1973) 'Tonal development in Min' *Journal of Chinese Linguistics* 1.2: 222-238.

Norman, J. (1974) 'The initials of Proto-Min' *Journal of Chinese Linguistics* 2.1: 27-36.

Peterson, G.E. and H.L. Barney (1952) 'Control methods used in a study of vowels' *Journal of the Acoustical Society of America* 24: 175-184.

Pilszczikowa-Chodak, N. (1972) 'Tone-vowel height correlation and tone assignment in the patterns of verb and noun plurals in Hausa' *Studies in African Linguistics* 3.3: 399-421.

Wang, W. S-Y. (1968) 'The many uses of F_0' *Project on Linguistic Analysis Reports*, 2nd. series 8: W1-W35. Revised version in *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre* (ed. A. Valdman), Mouton, The Hague, 1972: pp. 487-503.

Yuan Chia-hua et al. (1960) *Han-yu fang-yen kai-yao* [Outline of Chinese Dialects] Peking.

Acknowledgements

* This work partially supported by the Project on Linguistic Universals, Stanford University (funded by the National Science Foundation).

A Hybrid Multiple-Channel Pitch-Frequency Analysis System

Norris McKinney and Ronald Carlson

[University of California, Los Angeles]

In order to obtain more accurate measurements of the pitch-frequency of speech waves, we have undertaken development of an improved pitch-frequency analysis ("pitch analyzer") system. It combines the outputs of two or more differently configured simple analyzer channels in order to utilize the advantages of each channel configuration over the others. The system is a hybrid of analog and digital hardware and digital software.

The first phase of the project involved primarily the redesign of sections of pitch analyzer circuits already in use in the U.C.L.A. Phonetics Lab, utilizing circuit components which were not readily available when the circuits were originally designed. In the second phase of the project we are studying the conditions under which each of several selected single-channel configurations yields a more reliable output than the other single-channel configurations. We are also developing a program for our PDP-12 computer which will utilize some results of this study by smoothing and combining the outputs of two or more simple channels to yield a system output significantly more accurate than the raw output from any one of the channels alone.

We expect that, for a given degree of reliability, such a hybrid system, in which both continuous-function circuits (e.g., wave filters) and a computer are utilized, will be much faster than an all-digital system in which the speech wave is sampled by the PDP-12 and a program does all the signal analysis. We also expect that such a hybrid system will be much more effective than any non-digital analysis system simple enough for us to assemble conveniently.

Channel configuration

We selected as our initial channel configurations the two shown in the system block diagram in Figure 1. Both channels incorporate principles from the simplest and oldest pitch analyzers, such as those described by Hunt (1935) and Dudley (1937). Their basic design has been modified 1) by incorporating an interval-to-voltage (t-V) converter instead of either the commonly used frequency-to-voltage converter (pulse rate converter) or a Grützmacher-Lottermoser (1937) sawtooth generator, and 2) by non-linear smoothing with the computer instead of the typical linear smoothing by a filter. Detailed analysis and rationale for our choice of these two channel configurations for the system are given by

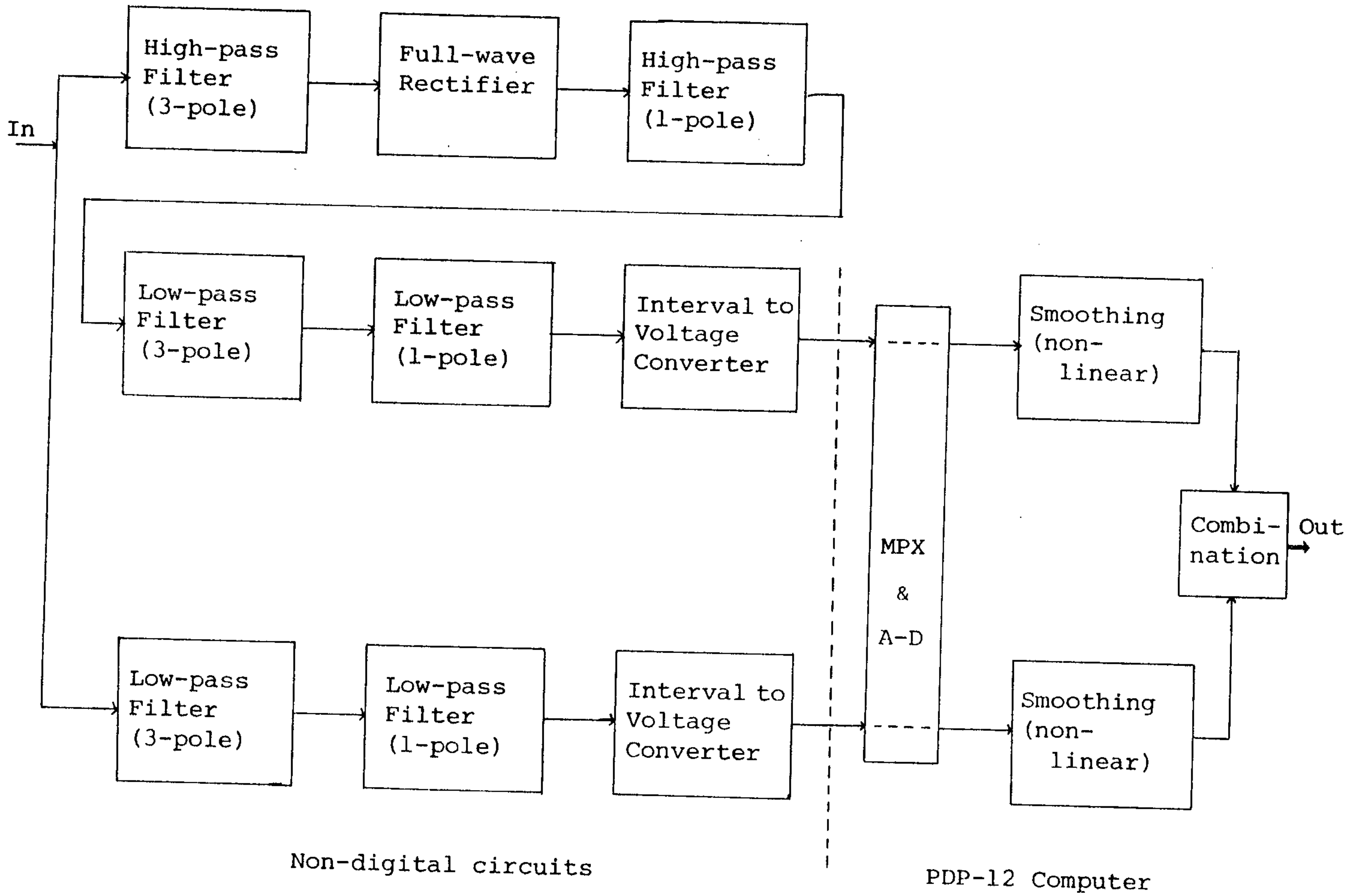


Figure 1. Pitch frequency analyzer. The Input is a voiced speech wave, the Output is a measurement of its pitch frequency.

McKinney (1965; 245). A brief resume follows.

The assumption or expectation in the design of an analyzer channel with a "frequency meter" circuit is that during much of the time when a voiced speech wave is being applied at the input to the channel there will be exactly one pair of zero crossings in the input to the t-V converter (in our design) or other frequency meter (in older designs) for each cycle of the speech wave. If there are not exactly two zero-crossings per cycle, the channel will give an incorrect indication of the fundamental frequency.

The preprocessor sections preceding the t-V converter are required to modify the signal before it is applied to the converter or frequency meter, because the unmodified speech wave is a very complicated function of time which very seldom meets the requirement of two zero-crossings per period. If the time-domain requirement that the signal have exactly two zero-crossings per period is to be met, then in the frequency domain we should make sure that the signal meets the criterion

$$A_1 > \sum_{n=2}^{M} nA_n ,$$

where A_n is the amplitude of the n^{th} harmonic component and M may be arbitrarily large (McKinney, 1965, Chapter 6). This is the reason low-pass filtering always precedes the t-V converter or other converter.

The earliest pitch-frequency analyzers consisted of a single channel much like the sequence of the first three circuits in the lower channel in Figure 1, with only a low-pass filter as the preprocessor. But often the first harmonic component of the incoming speech wave attains such a low relative level that it becomes impossible for such a simple preprocessor configuration to meet the criterion of producing a wave with only one pair of zero-crossings per pitch-frequency period.

An early attempt at refinement of the proprocessor was to add a nonlinear circuit to it, as in the upper channel of Figure 1, in order to increase the level of the first harmonic component relative to that of the second. Experimentally it has been found that on the average the full-wave rectifier significantly increases the level of the fundamental component (relative to the second harmonic component) when it is low (McKinney, 1965:265 ff.), and that augmentation of the fundamental component in this way results in a signal with the required two zero-crossings per period during much of the time when the undistorted signal fails to do so.

However, because of the range of variation of spectral structure and pitch frequency of the speech wave, it has not been found possible to design a simple single-channel system which works satisfactorily. It may work well for some of the voiced sections of the speech wave, but it fails for an annoying proportion of the time.

Often, when a channel of one simple configuration fails for a particular waveform because the input to the frequency meter contains more than one pair of zero-crossings per pitch-frequency period, a channel of another simple configuration avoids that failing, and the union of successful measurements on a set of periods of a voiced speech wave by the two channels much more nearly covers the whole of the voiced portion of the speech wave than does the measurement by either of the channels individually.

We must now consider the criteria for considering one channel rather than the other to be giving the correct pitch output. It seems most appropriate to choose whichever channel is most stable. Accordingly the criteria for best channel selects the lowest weighted history of pitch change (Δt) since the current channel was selected. This "rewards" stable channels and "punishes" errors such as second harmonic tracking, single period dropouts, and double zero crossings with increasing severity. A double zero-crossing will receive two periods of large excursion compared to a correct channel with one corresponding correct pitch period. Each channel is evaluated by:

$$\Delta t_b = \sum_{n=-a}^{-1} \frac{-(t_{b_{n+1}} - t_{b_n})}{n}$$

$a \leq 5$ determined by time of current channel selection
$b =$ channel 1 or channel 2
$t_n =$ pitch at period n

The maximum value of "a", the maximum memory length, is an insightfully chosen constant.

A smoothed output form can be selected from the computer which uses a three point median smoothing algorithm.

The potential increase of frequency of success in measuring the duration of pitch-periods of a speech wave by employing a multi-channel analyzer system with preprocessors which are configured as in Figure 1 has been shown experimentally (ibid, Chapters 7 and 8). It was found that a channel in which the signal is distorted with an even function, such as full-wave rectification or squaring, is a good complement to a channel without distortion, like the upper one in Figure 1. This is because the distortion often considerably increases the level of the first harmonic component of the signal relative to the level of the second harmonic component if the relative level of the first harmonic component is too low for the channel with linear preprocessing. Full-wave rectification is more practical than squaring because squaring doubles the dynamic range of the signal, as measured in decibels.

A high-pass filter is necessary following the rectifier to attenuate the low-frequency components in the rectifier output. Its frequency setting must compromise between being high enough to block these unwanted components and low enough to pass the pitch-frequency fundamental component. The high-pass filter preceding the rectifier simultaneously reduces the first harmonic component and any low-frequency noise components going into the rectifier.

Filters

Our 3-pole low-pass filters were designed to have a pole configuration given by Voss (1954), further treated by Fant (1959), and adopted

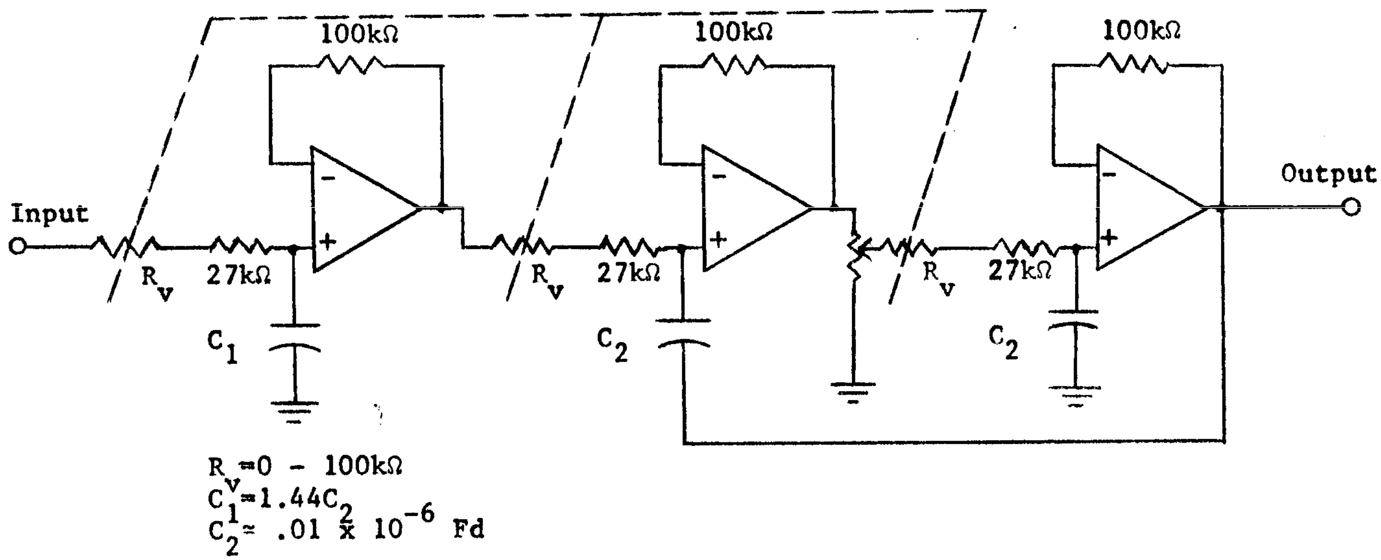


Figure 2. Circuit diagram of the 3-pole low-pass filters.

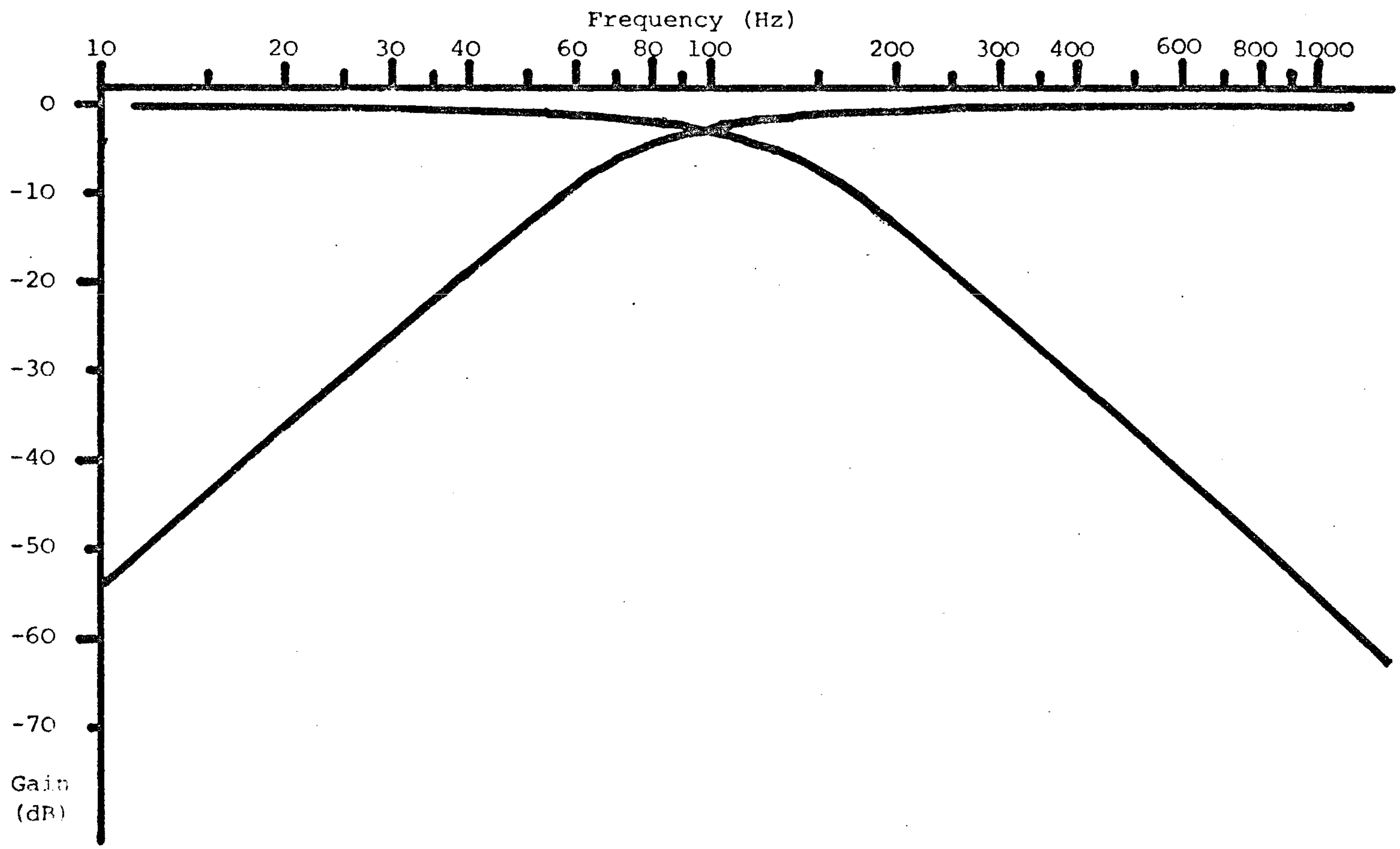


Figure 3. Gain-vs.-frequency responses of the 3-pole low-pass filters and the 3-pole high-pass filter (with 3 dB cut-off frequency set to 100 Hz).

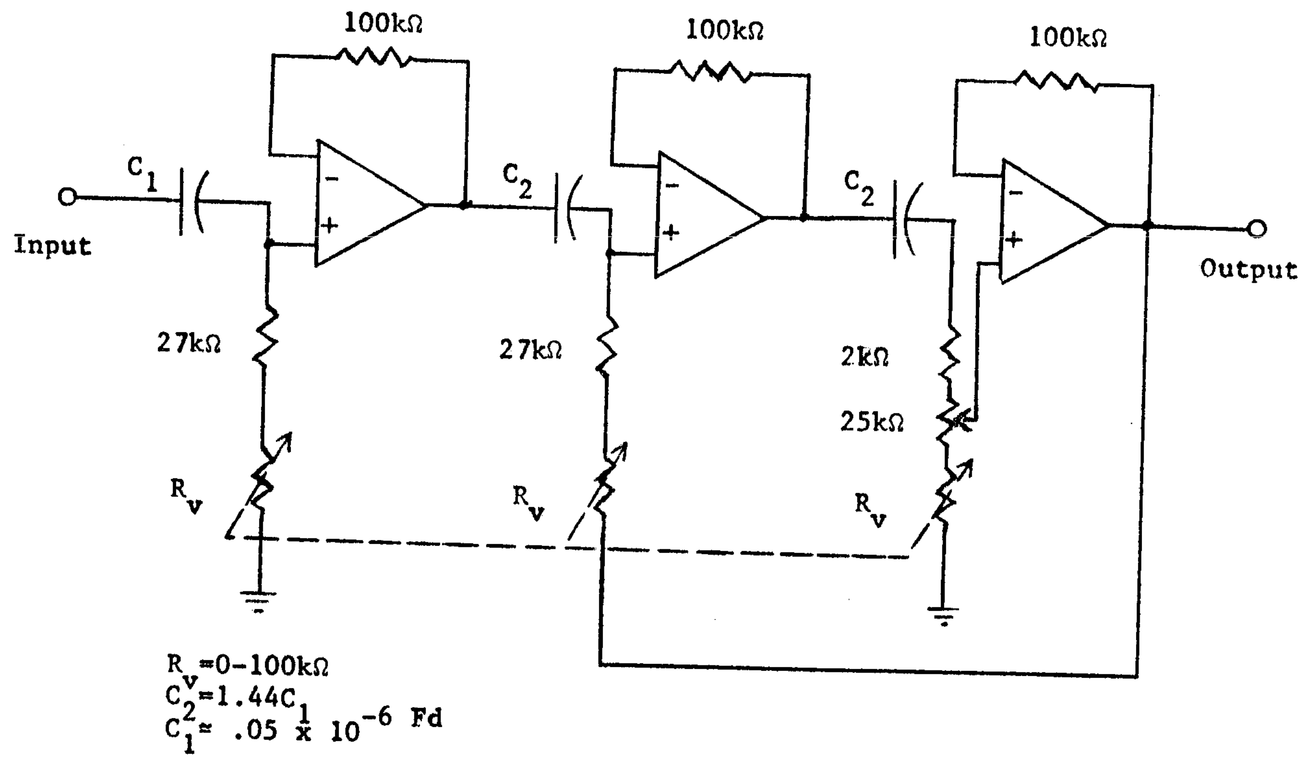


Figure 4. Circuit diagram of the 3-pole high-pass filter.

by him as standard at the Speech Transmission Laboratory:

$$p_1 = -0.85\ \omega_I$$
$$p_{2,3} = -0.7\ \omega_I \pm j\omega_I$$

The prototype is described as being an optimum phase designed LRC whole section low-pass filter, designed as an approximation to an idealized filter possessing a linearly rising step function response. The design of active low-pass filters with this pole configuration, and of active high-pass filters with a corresponding response characteristic, has been discussed by Fant and Liljencrants (1964).

Our 3-pole low-pass filters are designed to have a variable cutoff frequency set by the operator by means of a three-section ganged potentiometer. The circuit is shown in Figure 2, and its response in Figure 3. The single-pole low-pass filters shown in Figure 1 have essentially the same design as the first section of the 3-pole low-pass filters.

The 3-pole high-pass filters are designed as a transformation of the low-pass filters. The circuit is shown in Figure 4.

Interval-to-voltage converter

The analog t-V converter produces at each moment an output voltage proportional to the inverse of the time interval between the last two positive-going zero crossings in its input signal, provided that the interval lies within the 2-20 msec operating range of the circuit. This range corresponds to a frequency range of 50-500 Hz, which is sufficient for almost all adult speech (McKinney, 1965, Chapter 3). The converter is intended to measure the duration of each cycle of the first harmonic component of the speech signal and to output a voltage proportional to the fundamental frequency, cycle by cycle.

The basic principles of operation of the t-V converter are those of the one we have been using (McKinney, 1965). They were adapted from those of an earlier vacuum tube circuit used at the Acoustics Laboratory, MIT (1957). The converter circuit sums three decaying exponential current functions of time

$$I(t) = \sum_{n=1}^{3} A_n \exp\ (-t/T_n)$$

as an approximation of the parabolic function $f(t) = K/t$, corresponding to the inverse relationship between frequency and period of a periodic wave. Each new positive zero-crossing at the input to the converter causes the voltage at the output of the summing amplifier to be sampled and held, then resets the exponential functions to their initial values, and finally releases them to decay again after a delay of nearly 2 msec.

Appropriate values of the time constants T_n and ratios of initial values $A_n \exp(-.002/T_n)$ for the exponential functions were determined by means of a computer program written in this Lab by Lloyd Rice. It searched for values of these parameters which would reduce a sum of the squares error function over the 2-20 msec operating range. After experimenting with different error function definitions, we selected the results given by the following circuit values:

R(Ohms)	C(mFd)
147700	.005892
99080	.03030
193300	.09651

The approximation I(t) thus obtained for the function f(t) = k/t is within ± .25% of f(t) over a frequency range of 56 Hz to 475 Hz. We considered that this degree of accuracy of approximation was commensurate with the accuracies with which we could conveniently select the corresponding circuit components and test the complete circuit, and that it was more than adequate for research on speech.

An interval-to-voltage converter such as ours has considerable advantage over a commonly used type of post-processor component which we will call a pulse rate converter. In the latter, all the pitch-frequency pulses are made to have the same amplitude and wave shape (and hence the same area) and the resulting pulse signal is low-pass filtered. In the design of the low-pass filter, one can only arrive at an unsatisfactory compromise between (1) a low cutoff frequency to reduce the pitch-frequency ripple in the output and (2) a rapid time response to allow the output to follow sudden changes in the pitch frequency. The sharp contrast in performance between these two types of converters has been illustrated previously (McKinney, 1965, p. 234).

Both the step-function output of our t-V converter and the interrupted sawtooth output of a Grutzmacher-Lottermoser type converter are "instantaneous" in the sense of providing a measure of each pulse interval independent of previous pulse intervals. However, the output of the t-V converter is more convenient both for our linguists to read on oscillograms and for our computer to read with its analog-to-digital converter. The basic differences between the two circuits are the linearizing of the output voltage vs input frequency function and the addition of a sample-and-hold circuit in the t-V circuit.

The interval to voltage converter Figure 6 is primarily a three stage device. A front end squares the filtered speech input and detects zero-crossings. A series of one-shots, when fired by a detected zero-crossing, enables a sample-and-hold of the previous result and then initiates a 1.8 msec recharging period in the 1/t RC circuits. At the end of this period the RC circuits have accurately settled to the reference voltage and the 1/t current decay begins. When another zero-crossing is detected the voltage generated 1/t stage is sampled and held as output and the cycle repeats. Boundary conditions are easily seen in the output since when the circuit first fires (or there have been no zero-crossings for longer than 20 msec) the output indicates a pitch of 50 Hz.

Not including the computer, the system can be built for about $250 in parts. The sample-and-hold circuit and analog gates were manufactured by Teledyne Crystalonics. Newer analog gates can be used to reduce the package count to five "chips" per t-V channel. The full wave recitifier

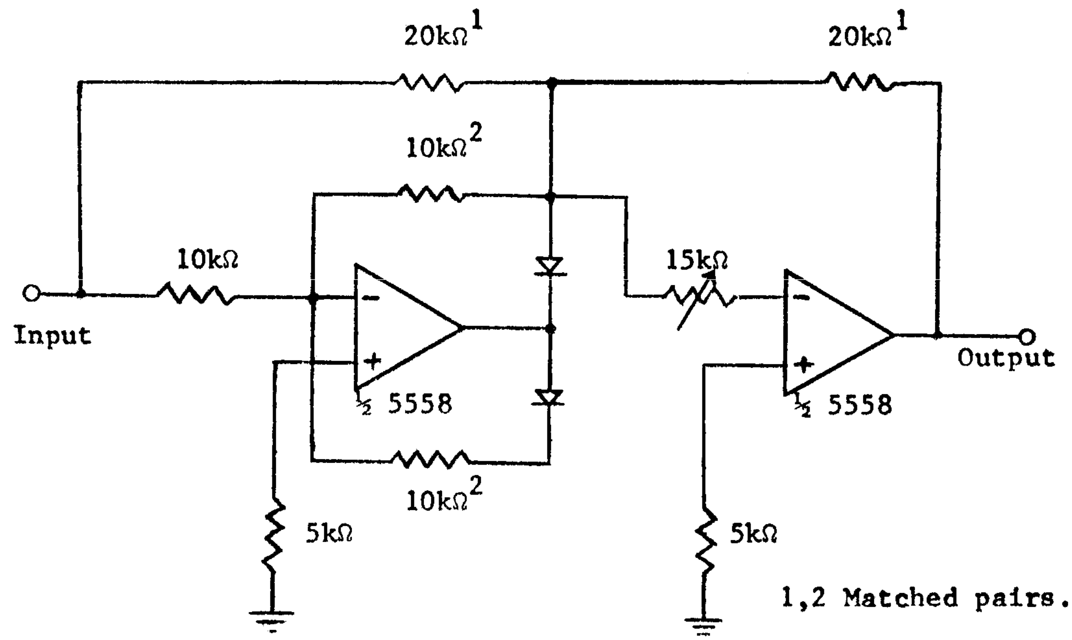


Figure 5. Full wave rectifier.

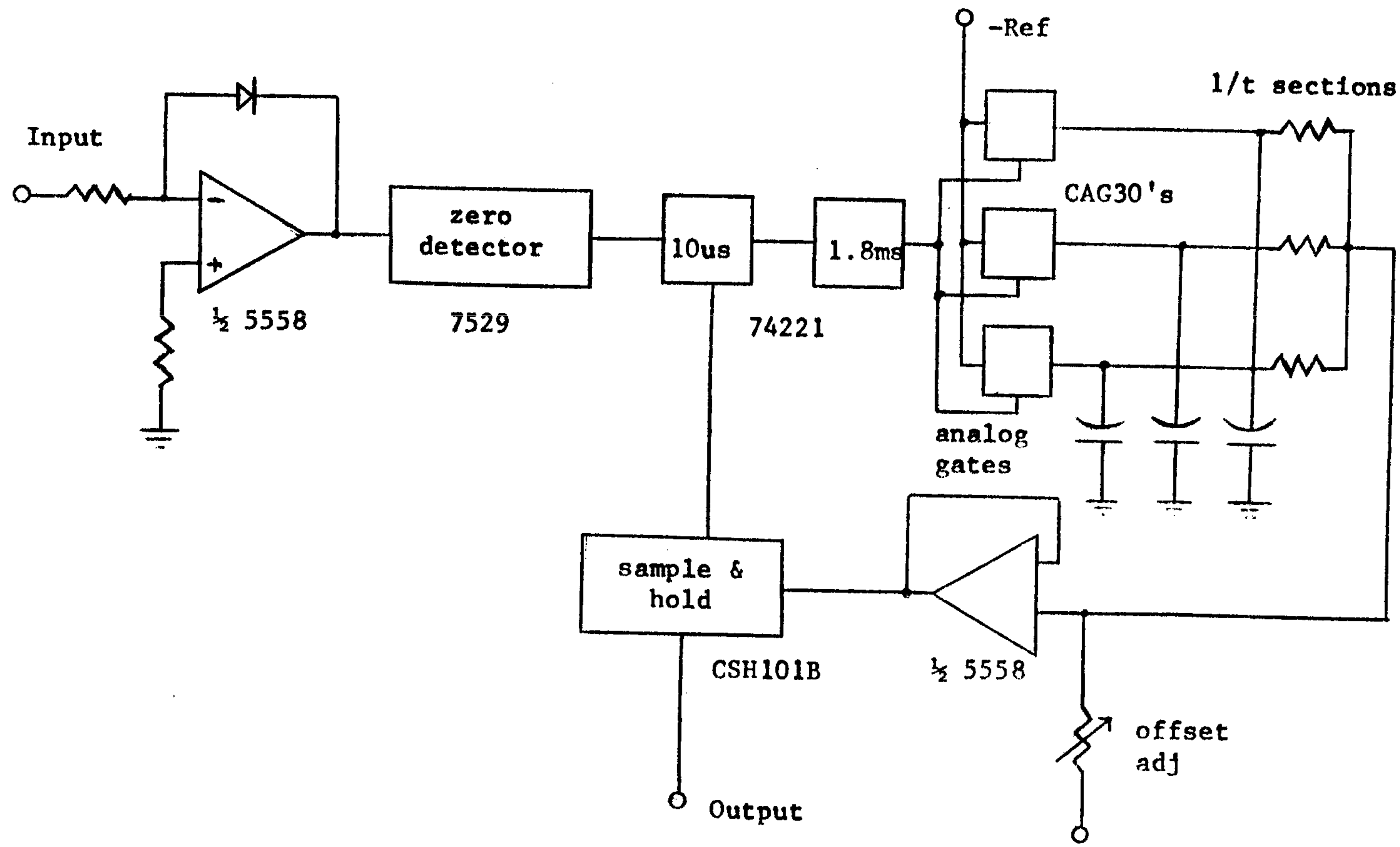


Figure 6. **Interval to voltage converter.**

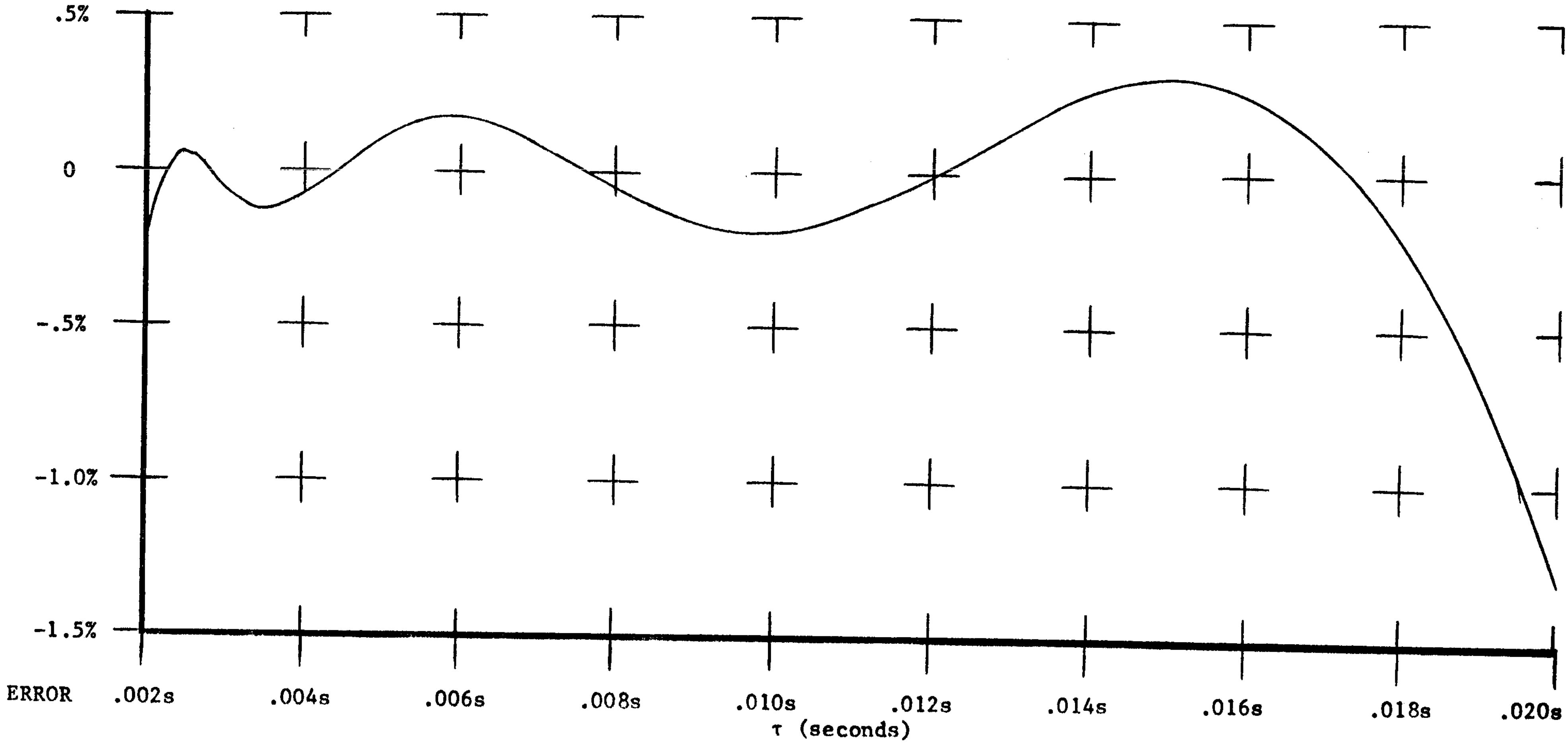


Figure 7. Error function of 1/t approximation.

is active to ensure excellent operation below 50 mv. The front end wave shaper is an open loop op-amp (with protective negative gain limiting at -.4v) to improve the accuracy of the zero crossing detector.

To ensure stability of .25% for at least a year the following steps are recommended:

1) Select the wave shaper op-amp for positive offset to ensure an accurate zero-crossing point.

2) "Burn in" end cycle resistors and capacitors until stable and seal with varnish.

3) Use temperature coefficients of less than 100 ppm.

4) Match RC coefficients to counteract.

5) Exercise care in power distribution, TTL noise cannot be allowed in the analog circuits.

6) Shorter sampling windows will provide more accurate 1/t acquisition at the expense of noise. 10 μsec seemed to be an adequate tradeoff.

7) Use rubber band regulation of power (+5v,-5v,+15v,-15v) to reduce offset variations and increase common mode rejection.

Bibliography

Acoustics Laboratory (1957) *Speech Compression Research: Final Report* Contract Number AF 19(604)-626. Cambridge, Massachusetts: Acoustics Laboratory, MIT. February 1957.

Dudley, H. W. *Signaling System* United States Patent Number 2,098,956, November 16, 1937; Application, December 2, 1936.

Fant, C. G. M. (1959) Acoustic Analysis and Syntehsis of Speech with Applications to Swedish, Ericsson Technics, Number 1.

Fant, C. G. M. and Liljencrants, J. (1964) Encla filter med aktiva RC- och LRC-nät (Simple filters from active R-C or L-R-C circuits) *Elektronik* 1.

Grützmacher, M. and Lottermoser, W. (1937) Über ein Verfahren zur trägheitsfreien Aufzeichnung von Melodiekurven. *Akustische Zeitschrift* Füftes Heft, Band 2, Zweiter Jahrgang, pp. 242-248, Sept. 1937.

Hunt, F. V. (1935) A direct-reading frequency meter suitable for high speed recording. *The Review of Scientific Instruments* 6:43-46, February, 1937.

McKinney, N. P. (1965) *Laryngeal Frequency Analysis for Linguistic Research* Contract Nonr 1224(22) NR 949-122. Ann Arbor: Communication Sciences Laboratory, September, 1965.

Voss, H. H. (1954) Realisierbare Tiefpässe und Bandpässe minimaler Phase mid geebneter Laufzeit und aperiodischem Einschwingverhalten. *Frequenz* nr. 8.

Vincent van Heuven:
25 oct, 1978

Summary of recently completed, current, and planned experiments on the phonetic function of rise and decay time of Dutch and English speech sounds

(1) JND has been established for 500 msec (in a second experiment 250 msec) 1000 Hz sine waves and white noise bursts for rise and decay times. Average JND was found to be in the order of 25%. Results in the 250 and 500 msec duration conditions turned out to be highly stable. Accuracy of adjustment of rise or decay time (depending on the experimental condition) of a subject controled signal to a reference signal (as a measure of JND) was intermediate for noise-onset and sine-offset (cf. figure 1), slightly superior for sine-onset, and distinctly better for noise-offset in the upper half of the stimulus range. This asymmetry was tentatively explained on the basis of auditory differences due to well-definedness versus neural response behavior of narrow and wide band signals.

(2) JND was established for two artificial Dutch vowels [a] and [ɑ] of ambiguous total duration with 0 - 100 msec rise or decay times. Results indicate inferior performance on both onset conditions (equal to or even worse than on the non-speech stimuli used in the previous experiments), but distinctly better performance for the offset conditions (figure 2). Two explanations may be suggested:
- The asymmetry is an inherent effect due to auditory backward masking in complex periodic signals (suggested by Steve Greenberg),
- The asymmetry is caused by prolonged exposure to Dutch, a language in which vowel decay time (rather than rise time, steady state or total duration) has been claimed as a cue for a ternary opposition (lax, intermediate and tense).

Marcel van den Broecke will run an experiment with non-speech complex periodic signals (saw tooth or square wave) to check possibility #1. He will bring data along at the end of november). Personally I favor possibility #2.

(3) An experiment concerning certain production aspects of Dutch vowel rise and decay time is being carried out. 50 stimulus words embedded in an identical carrier phrase were read by two speakers, three times at normal rate of speech, and three times fast.

zei je	" ({s, z, t, d})	a:	({s, z, t, d})	-en,"	matie ?
zɛɪ jə		ɑ		ən	ma:rI: ?
	5	x 2	x 5		= 50 types

I attempt to relate the duration parameters rise time, decay time, steady state and total duration, to the variables implicit in the above table:
- the phonological vowel category [±tense],
- stop/fricative character of a voiced/voiceless preceding or following or absent consonant environment,
- rate of speech

There are phonotactic constraints against a pure vowel followed by another

Vincent van Heuven:
25 oct, 1978

Summary of recently completed, current, and planned experiments on the phonetic function of rise and decay tame of Dutch and English speech sounds

(1) JND has been established for 500 msec (in a second experiment 250 msec) 1000 Hz sine waves and white noise bursts for rise and decay times. Average JND was found to be in the order of 25%. Results in the 250 and 500 msec duration conditions turned out to be highly stable. Accuracy of adjustment of rise or decay time (depending on the experimental condition) of a subject controled signal to a reference signal (as a measure of JND) was intermediate for noise-onset and sine-offset (cf. figure 1), slightly superior for sine-onset, and distinctly better for noise-offset in the upper half of the stimulus range. This asymmetry was tentatively explained on the basis of auditory differences due to well-definedness versus neural response behavior of narrow and wide band signals.

(2) JND was established for two artificial Dutch vowels [a] and [ɑ] of ambiguous total duration with 0 - 100 msec rise or decay times. Results indicate inferior performance on both onset conditions (equal to or even worse than on the non-speech stimuli used in the previous experiments), but distinctly better performance for the offset conditions (figure 2). Two explanations may be suggested:
- The asymmetry is an inherent effect due to auditory backward masking in complex periodic signals (suggested by Steve Greenberg),
- The asymmetry is caused by prolonged exposure to Dutch, a language in which vowel decay time (rather than rise time, steady state or total duration) has been claimed as a cue for a ternary opposition (lax, intermediate and tense).

Marcel van den Broecke will run an experiment with non-speech cpmplex periodic signals (saw tooth or square wave) to check possibility #1. He will bring data along at the end of november). Personally I favor possibility #2.

(3) An experiment concerning certain production aspects of Dutch vowel rise and decay time is being carried out. 50 stimulus words embedded in an identical carrier phrase were read by two speakers, three times at normal rate of speech, and three times fast.

zei je	" ({s, z, t, d})	a:	({s, z, t, d})	-en,"	marie ?
zεɪ jə		ɑ		ən	ma:ri: ?
	5	x 2	x 5		= 50 types

I attempt to relate the duration parameters rise time, decay time, steady state and total duration, to the variables implicit in the above table:
- the phonological vowel category [±tense],
- stop/fricative character of a voiced/voiceless preceding or following or absent consonant environment,
- rate of speech

There are phonotactic constraints against a pure vowel followed by another

Vincent van Heuven:
25 oct, 1978

Summary of recently completed, current, and planned experiments on the phonetic function of rise and decay tóme óf Dutch and English speech sounds

(1) JND has been established for 500 msec (in a secend experiment 250 msec) 1000 Hz sine waves and white noise bursts for rise and decay times. Average JND was found to be in the order of 25%. Results in the 250 and 500 msec duration conditions turned out to be highly stable. Accuracy of adjustment of rise or decay time (depending on the experimental condition) of a subject controled signãl to a reference signal (as a measure of JND) was intermediate for noise-onset and sine-offset (cf. figure 1), slightly superior for sine-onset, and distinctly better for noise-offset in the upper half of the stimulus range. This asymmetry was tentatively explained on the basis of auditory differences due to well-definedness versus neural response behavior of narrow and wide band signals.

(2) JND was established for two artificial Dutch vowels [a] and [ɑ] of ambiguous total duration with 0 - 100 msec rise or decay times. Results indicate inferior performance on both onset conditions (equal to or even worse than on the non-speech stimuli used in the previous experiments), but distinctly better performance for the offset conditions (figure 2). Two explanations may be suggested:
- The asymmetry is an inherent effect due to auditory backward masking in complex periodic signals (suggested by Steve Greenberg),
- The asymmetry is caused by prolonged exposure to Dutch, a language in which vowel decay time (rather than rise time, steady state or total duration) has been claimed as a cue for a ternary opposition (lax, intermediate and tense).

Marcel van den Broecke will run an experiment with non-speech cpmplex periodic signals (saw tooth or square wave) to check possibility #1. He will bring data along at the end of november). Personally I favor possibility #2.

(3) An experiment concerning certain production aspects of Dutch vowel rise and decay time is being carried out. 50 stimulus words embedded in an identical carrier phrase were read by two speakers, three times at normal rate of speech, and three times fast.

zei je	" (s / z / t / d)	a: / ɑ	(s / z / t / d)	-en,"	marie ?
zɛɪ jə				ən	ma:ri: ?
	5 x	2 x	5		= 50 types

I attempt to relate the duration parameters rise time, decay time, steady state and total duration, to the variables implicit in the above table:
- the phonological vowel category [±tense],
- stop/fricative character of a voiced/voiceless preceding or following or absent consonant environment,
- rate of speech

There are phonotactic constraints against a pure vowel followed by another