

THE COARTICULATION OF TONES: AN ACOUSTIC STUDY OF THAI

Arthur S. Abramson¹

The University of Connecticut and Haskins Laboratories

Phonemes occurring in a great number of phonetic environments are subject to so much coarticulatory perturbation that phoneticians and psychologists have speculated about the nature of the invariance to be found in any set of speech sounds said to manifest an underlying phonological entity (MacNeillage 1970). The question seems to have received relatively little attention in the matter of phonemic tones which have glottal repetition rate as their primary articulatory base. It is conceivable that in something as continuously variable as the pitch of the voice, the lexical tones undergo so much sandhi and other perturbations that the tone system is not fully preserved in running speech. Even if this is not so, are the ideal fundamental-frequency (f_0) contours, epitomized by citation forms, preserved in a wide variety of contexts? Very little instrumental work on this topic has appeared in the literature. Examples are an investigation of Vietnamese by Han and Kim (1974) and a brief study of Thai by Palmer (1969).

The language chosen for this study is Central Thai (Siamese), which has five phonologically distinctive tones. In earlier work (Abramson 1962) ideal f_0 contours for these tones in citation forms were derived instrumentally and synthesized for perceptual validation. Although recent work (Erickson 1974) lends general support to the curves found, there is considerable disagreement over the preservation of these forms, as well as the full system of tonal contrasts, in running speech (Henderson 1949; Gandour 1975). Other attributes of the tones, such as glottalization and amplitude shifts, may be contextually induced, but they are not examined here.

Procedure

All possible sequences of two tones from the five tones of Thai--mid, low, high, falling, and rising--were embedded, on monosyllabic words in a carrier sentence beginning and ending on the mid tone (/raw _____ dii/) to form 25 grammatical sentences. The mid tone provided as neutral a frame as possible.² The embedded words were all of the form CV, as were the two words of the carrier. No care was taken to control for the balanced occurrence of voiced and voiceless initial consonants, which themselves might give rise to brief f_0 perturbations

(Erickson 1974; Gandour 1974). Indeed, the exclusive use of voiced consonants would have allowed for the starting of coarticulatory effects during the occlusions or constrictions of the consonants. It seemed from a preliminary look at some productions that all these effects, in addition to possible others induced by the failure to use the same vowel in all the key words, were too small to hide serious tonal coarticulation. Finally, there is the syntactic and semantic difficulty or impossibility of composing 25 such sentences with few or none of the foregoing irregularities.

The set of sentences was recorded at normal conversational speed on different occasions by each of four native speakers, three women and one man, who were all university graduates. Given the usual artificiality of such recording sessions, the productions were played back to the speakers themselves and the experimenter, and rejected if they did not sound normal. In addition, careful listening revealed no loss of tonal contrast for the embedded key words. Citation forms of the tones on isolated words were also obtained from the four speakers.

Patterns of f_0 were extracted from all recordings by means of Lukatela's (1973) computer-implemented autocorrelation method. The resulting graphs permitted the examination of every tone in each left and right environment along the time axis.

Results

Although all the data will not be presented here, a full analysis is to be given in a subsequent report (Note 1). Selected environments for Speaker P.C., one of the women, will be shown. They are rather representative of all four speakers.

In Figure 1 we see P.C.'s citation forms. The f_0 contours are normalized in time and displayed as percentages of her total voice range. Such data are sometimes shown as semitones or on a logarithmic scale as an attempt at depicting physical f_0 measurements in terms of the auditory sensation of pitch. The speech used here was calm and dispassionate, so it seemed that presenting the tonal shapes as percentages of each speaker's voice range provided reasonable comparability across tones and speakers.

The mid, low, and high tones are often said to be static or level tones, while the falling and rising tones, which have larger movements through the voice range, are called dynamic or contour tones (Abramson 1962, 1975). The other speakers have similar citation forms except that for two of them the mid tone slopes downward a little more and, for one of them, the falling tone starts its sharp drop immediately.

CITATION FORMS

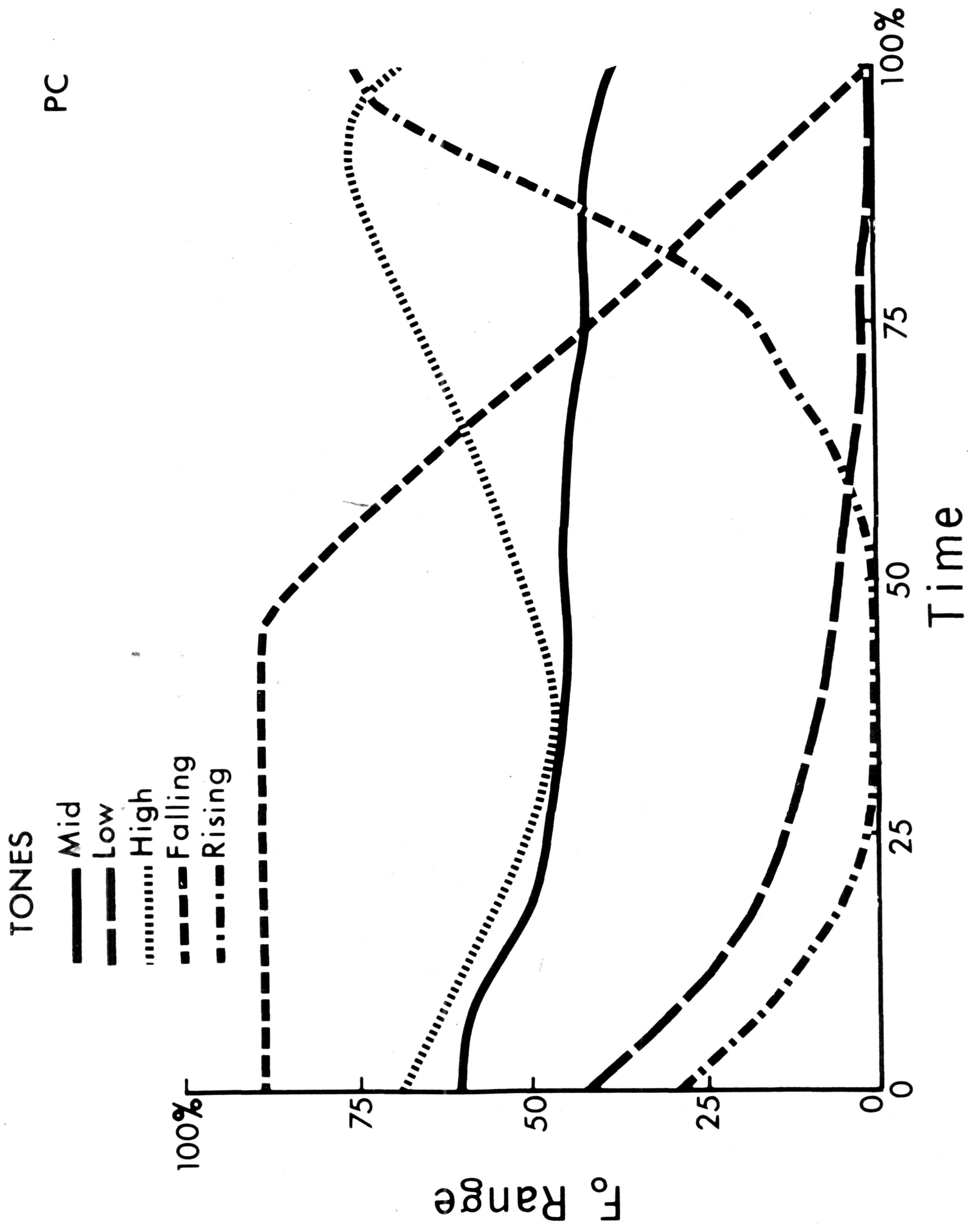


Figure 1

Turning to an examination of the tones and their contexts in the sentences, we find average curves for all of P.C.'s tones in the environment of the following mid tone in Figure 2. The left-hand portion of the figure shows the full set of tones in its environment, while the right-hand portion shows, as indicated by the coded lines, what happens to the context itself. P.C.'s tones in the environment of the following falling tone appear in Figure 3.

The arrays of f_0 curves in the left-hand portions of Figures 2 and 3 clearly support the auditory impression that the tonal system is preserved in these environments. Indeed, inspection of similar displays for the remaining three following contexts, as well as all five preceding contexts for this speaker and the other three speakers, leads to the same conclusion. If we compare the citation forms with the sets of tones in context, the f_0 ranges of the latter are somewhat compressed. Also, the curves are perturbed, especially at their endpoints, by the preceding and following contexts. Note, for example, that P.C.'s high tone in Figure 2 dips at the end before the following mid tone, while in Figure 3 it stays up before the high beginning of the following falling tone. The falling tone in both figures, coming as it does after the mid tone of the carrier frame, is different from the citation form in that it rises considerably before it falls. In many contexts, as in Figure 2, the high tone and the rising tone might be called high rising and low rising, respectively, as they are almost mirror images of each other with U-shapes that differ mainly in absolute frequency height. It should be noted, however, that the dip of the high tone of this type is smaller than that of the rising tone and does not give so obvious an auditory impression of rising as does the rising tone itself.

The families of curves representing the mid-tone and the falling-tone contexts of Figures 2 and 3 form remarkably tight clusters. For all such contexts produced by the four speakers there is very little difficulty in assigning any member of each family of curves to the tone that it is supposed to represent. This is so even though one can see perturbations at the beginning for some of the preceding contexts and at the end for some of the following contexts.³ In Figure 2 note the considerable variation in the beginnings of the five variants of the mid tone. In Figure 3 only the variant of the falling tone following the mid tone shows serious initial perturbation, while the others all start in a similar fashion. The other three speakers, however, do not show even this perturbation but rather seem to reset the larynx for a high beginning of the falling tone no matter which of the five tones comes before. They do, nevertheless, show the initial rise of the falling tone after the mid tone of the beginning of the carrier sentence. There may be rhythmic factors at work (Noss 1972).

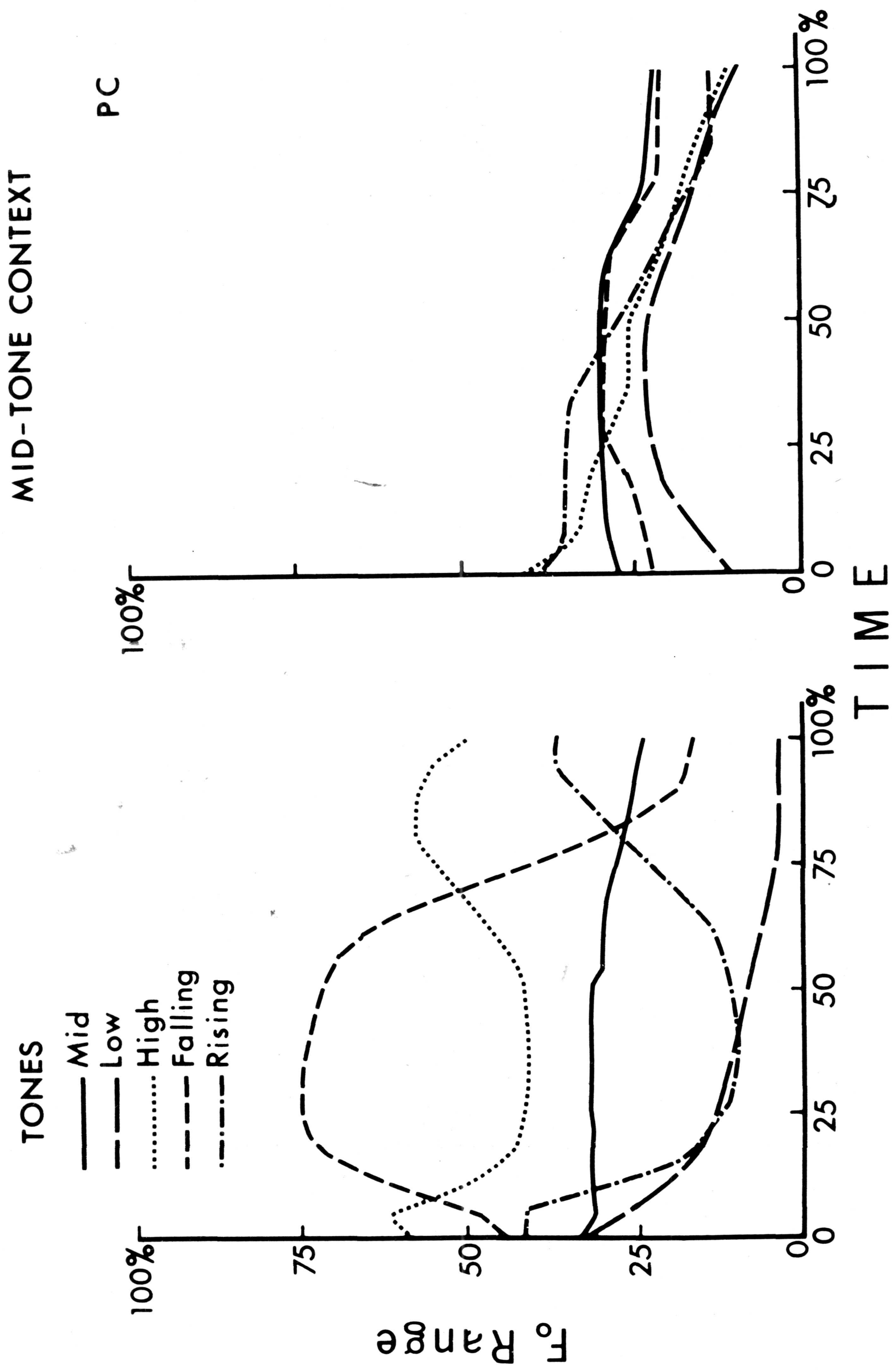


Figure 2

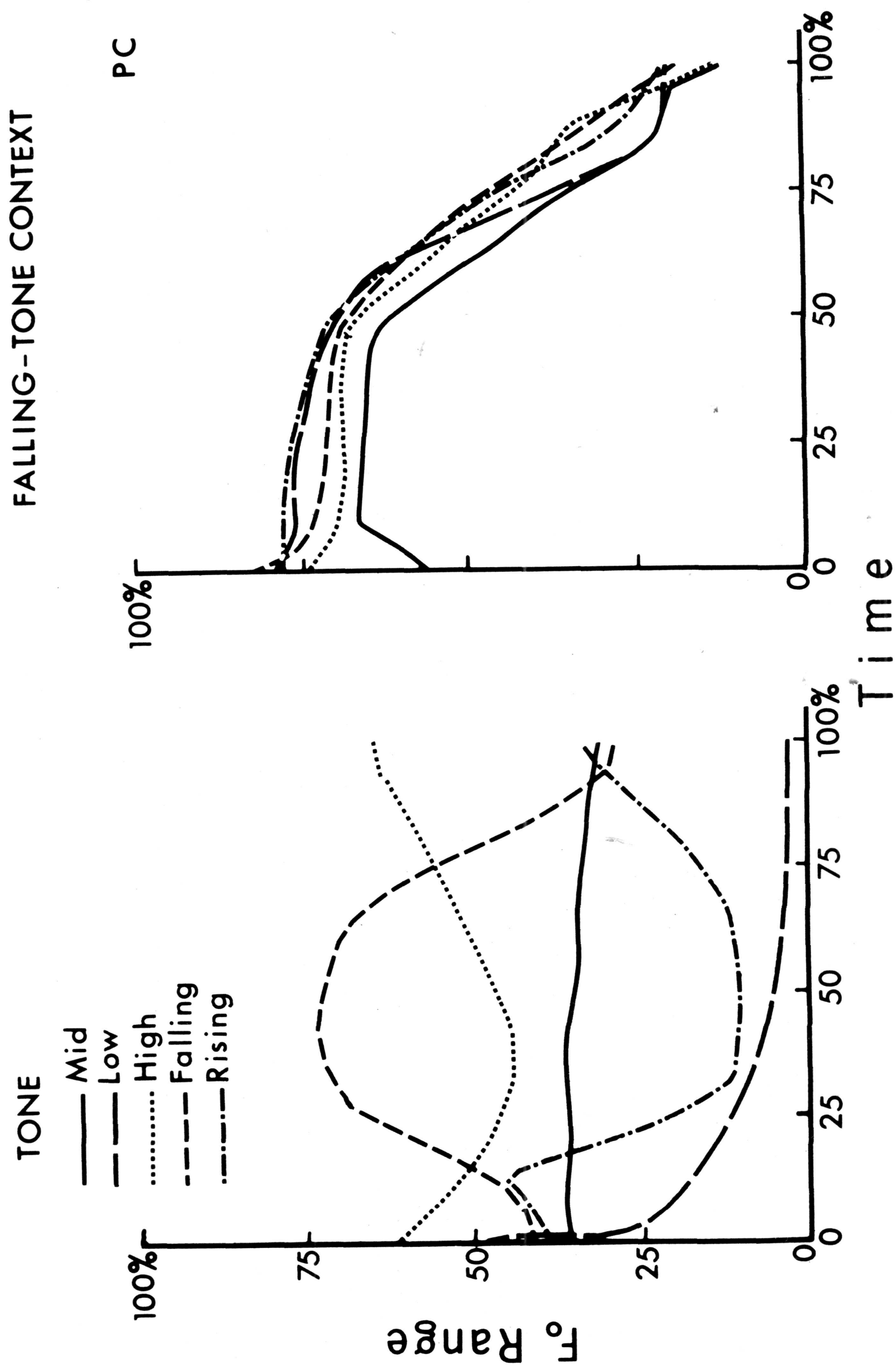


Figure 3

Discussion

The data can be understood as reflecting three related effects.

- (1) The full system of five tones is preserved on monosyllabic Thai words embedded in all possible preceding and following tonal contexts.
- (2) With citation forms taken as the standard, embedding a tone in running speech causes some perturbation, but generally not enough to damage its identifiability.
- (3) Laryngeal coarticulation is governed by the specific tonal context; nevertheless, in some contexts coarticulation does not occur for some speakers, and we may suppose that in such instances the larynx receives instructions to reset itself to produce an approximation of the ideal contour for the tone. Even then, rhythmic conditions may prevent the production of the ideal tone as found in citation forms.

In conclusion, it is interesting to note that the data lend no phonetic plausibility to arguments for the specification of the dynamic tones as temporal sequences of features that underlie two of the static tones. The argument, based on purely linguistic reasoning, specifies the falling tone as a sequence of high and low features, and the rising tone as a sequence of low and high features (Leben 1973; Gandour 1975). For phonetic support of the argument one would expect to be able to devise a formula by which the dynamic tones were obviously to be derived from the shapes of the static tones (Abramson Ms.). Even the citation forms, let alone the f_0 curves of running speech, provide no acoustic basis for such a claim. It seems psychologically far more reasonable to suppose that the speaker of Thai stores a suitable tonal shape as part of his internal representation of each monosyllabic lexical item.

NOTES

- ¹ This is a revised and somewhat lengthened version of an oral paper given at the Eighth International Congress of Phonetic Sciences, Leeds, England, 17-23 August 1975. While most of the analysis of data was performed at Haskins Laboratories, the data themselves were collected while the author was on sabbatical leave in Thailand (1973-74) on research fellowships from the American Council of Learned Societies and the Ford Foundation Southeast Asia Fellowship Program. I gratefully acknowledge the hospitality of Dr. Udom Warotamasikkhadit, Dean, the Faculty of Humanities, Ramkhamhaeng University and Mrs. Mayuri Sukwiwat, Director, the Central Institute of English Language, both in Bangkok.

- 2 The sentences were taken from Palmer (1969) because of some doubt about the replicability of his findings. This will be discussed in a more extensive treatment of the present subject now being done with Thomas Baer.
- 3 In unstressed syllables of compound words certain lexical tones are likely to undergo morphophonemic alternation resulting in the neutralization of some contrasts.

REFERENCES

- Abramson, Arthur S. 1962. "The vowels and tones of standard Thai: acoustical measurements and experiments." *Indiana University Research Center in Anthropology, Folklore, and Linguistics* 20. (=IJAL 28.2, Part III).
- _____. 1975. "The tones of Central Thai: some perceptual experiments." *Studies in Tai Linguistics in Honor of William J. Gedney*, ed. by Jimmy G. Harris and James R. Chamberlain. Bangkok: Central Institute of English Language.
- _____. Ms. "The phonetic plausibility of the segmentation of tones in Thai phonology." *Proceedings of the Twelfth International Congress of Linguists*.
- Erickson, Donna. 1974. "Fundamental frequency contours of the tones of standard Thai." *Pasaa* 4: 1-25.
- Gandour, Jack. 1974. "Consonant types and tone in Siamese." *J. Phonetics* 2: 337-50.
- _____. 1975. "On the representation of tone in Siamese." *Studies in Tai Linguistics in Honor of William J. Gedney*, ed. by Jimmy G. Harris and James R. Chamberlain. Bangkok: Central Institute of English Language.
- Han, Mieko S. and Kong-On Kim. 1974. "Phonetic variation of Vietnamese tones in disyllabic utterances." *J. Phonetics* 2: 223-32.
- Henderson, Eugénie J.A. 1949. "Prosodies in Siamese: a study in synthesis." *Asia Major*, N.S. 1: 189-215. (Reprinted in *Phonetics in Linguistics: a Book of Readings*, ed. by W. E. Jones and Laver. London: Longman, 1973)

- Leben, William R. 1973. "The role of tone in segmental phonology."
Consonant Types and Tone, ed. by Larry M. Hyman (Southern
California Occasional Papers in Linguistics No. 1).
- Lukatela, Georgije. 1973. "Pitch determination by adaptive autocorrela-
tion method." *Haskins Laboratories Status Report on Speech
Research SR-23*: 179-208.
- MacNeilage, Peter F. 1970. "Motor control of serial ordering in speech."
Psych. Review 77: 182-96.
- Noss, Richard B. 1972. "Rhythm in Thai." *Tai Phonetics and Phonology*,
ed. by Jimmy G. Harris and Richard B. Noss, 33-42. Bangkok:
Central Institute of English Language.
- Palmer, Adrian. 1969. "Thai tone variants and the language teacher."
Language Learning 19: 287-99.