The Coarticulation of Tones: An Acoustic Study of Thai*

Arthur S. Abramson+

ABSTRACT

Phonologically distinctive tones have glottal repetition rate as their primary articulatory base. In many tone languages, F0 contours derived from productions of isolated monosyllables are considered the ideal forms of the tones. These ideal contours are not fully realized in running speech. This may be ascribed to rhythmic conditions and to laryngeal coarticulation as a function of the immediate tonal environments. Ideal F0 contours for the five tones of Thai were obtained for four native speakers. Then, in a neutral frame, a mid tone at the beginning and a mid tone at the end, the four speakers produced all possible sequences of two tones in sentences. The data yield the following effects: (1) embedding a tone in running speech causes some departure from its ideal contour; (2) laryngeal coarticulation is governed by the specific tonal context; (3) tonal contrast is not lost.

Phones are subject to so much coarticulatory perturbation that phoneticians and psychologists (MacNeilage, 1970) have speculated about the nature of the invariance supposedly present in any set of speech sounds said to manifest an underlying phonological entity. The question seems to have received relatively little attention in the matter of phonemic tones that 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


+Also University of Connecticut, Storrs.

Acknowledgment: 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 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.

[HASKINS LABORATORIES: Status Report on Speech Research SR-44 (1975)]
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 about the preservation of these forms and the system of tonal contrasts in running speech (Henderson, 1949; Gandour, 1975).

**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 to form 25 grammatical sentences. The mid tone provided as neutral a frame as possible.\(^1\) The list of sentences was recorded at normal conversational speed on different occasions by each of four native speakers, three women and one man. Given the usual artificiality of such recording sessions, the productions were judged natural by the speakers themselves and the experimenter. 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, I hope to make a full analysis available in subsequent reports. 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. 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. 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.

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

\(^{1}\)The sentences were taken from Palmer (1969).
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 end points, 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 is 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
Figure 2: The five tones preceding the mid tone and the resulting variants of the mid-tone context.
Figure 3: The five tones preceding the falling tone and the resulting variants of the falling-tone context.
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 tones and at the end for some of the following tones. 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 fixed mid tone of the beginning of the carrier sentence. There may be rhythmic factors at work (Noss, 1972).

**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). Even the citation forms, let alone the F0 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.

**REFERENCES**


---

2 In unstressed syllables of compound words certain lexical tones are likely to undergo morphophonemic alternation resulting in the neutralization of some contrasts.


