PERCEPTUAL DIMENSIONS OF TONE: THAI

JACKSON T. GANDOUR*

ABSTRACT. One hundred and fourteen Thai subjects made direct ratings of dissimilarity between pairs of pitch patterns superimposed on a synthetic speech-like syllable. An INDSCAL analysis of the dissimilarities data revealed four dimensions which were interpreted as AVERAGE PITCH, LENGTH, DIRECTION and SLOPE. These interpretive labels were supported by results of a multiple linear regression analysis. No significant differences in tonal perception could be attributed to an individual subject's dialect background.

I. INTRODUCTION

A search for linguistic explanations that account for the nature of tonal systems must converge with fundamental processes associated with the production and perception of tones. Using a multidimensional scaling procedure, the primary aims of this paper are (1) to discover the fundamental dimensions underlying Thai individuals' perception of different kinds of pitch patterns superimposed on a synthetic speech-like syllable and (2) to determine the extent and kinds of individual differences in perception that may be attributed to dialect background.

---

*Postdoctoral Fellow. The research for this paper was supported in part by Social Science Research Council and National Science Foundation dissertation fellowships, and by NSF G537235X during my doctoral program in linguistics at the University of California, Los Angeles.

I wish to especially thank the following persons and institutions of Thailand for their assistance and cooperation: (Central Institute of English Language) Jimmy Harris and Kanda Sittachita; (Ramkhamhaeng University) Pat Noisaengsri and Udom Warotamasikkhadit; (Chiang Mai University) Tom Scovel; (Phuket Teacher Training College) faculty members of the English and Thai departments. I also wish to thank the Thai students at these various institutions who participated in the experiment.

Requests for reprints should be sent to:
Jackson T. Gandour, Department of Audiology and Speech Sciences, Purdue University, West Lafayette, IN 47907.

The secondary aim of the paper is to determine how close these perceptual dimensions correspond to earlier proposed phonological features of tone.

The mutlidimensional scaling procedure used in this investigation, INDSCAL (=INdividual Differences SCALing: Carroll and Chang 1970; Carroll and Wish 1974a,1974b; cf. Harshman 1970,1972), simultaneously analyses similarity (or dissimilarity) matrices for several individuals. The data values in the matrices generally correspond to subjective distances between stimulus objects based on judgments of similarity of different individuals. The similarities are assumed to be related to distances between stimuli in some latent psychological space. The INDSCAL procedure determines a statistically unique set of diemsnions that usually can be interpreted without rotation of the axes. In addition to coordinates for the stimuli on each dimension, INDSCAL also provides information about the relative weights or perceptual saliences of each dimension for every individual. The distances between the stimuli depend on the subjects' dimension weights as well as on the stimulus coordinates. The dimension weights for a particular individual subject indicate approximately how much each dimension should be stretched so that the distances between stimuli will correlate as highly as possible with that subject's similarity (or dissimilarity) ratings. The stimulus coordinates on each dimension for all subjects may be plotted in a "group(composite) stimulus space", the dimension weights for individuals in a "subject space". It is primarily by analysis of the subject space that we may determine to what extent differences in perception, if any, may be attributed to particular individuals or subgroups.

II. METHOD AND PROCEDURE

A. *LANGUAGE*

Thai, a member of the Tai branch of the Austro-Thai language family, is an example of a tone language with five contrastive tones - three level tones and two contour tones, traditionally labelled "high" (´), "mid" ( ), "low" (`), "falling" (^) and "rising" (ˇ), e.g. /khaá/ *'to engage in trade'*, /khaa/ *'to be stuck'*, /khaà/ *'a kind of spice'*, /khaâ/ *'to kill'*, /khaǎ/ *'leg'*. For phonological description of the Thai tones, see Henderson 1949, Abramson 1962 and Gandour 1975. Figure 1 presents average fundamental frequency trajectories of the high, mid, low, falling and rising tones of Thai in word-final position.

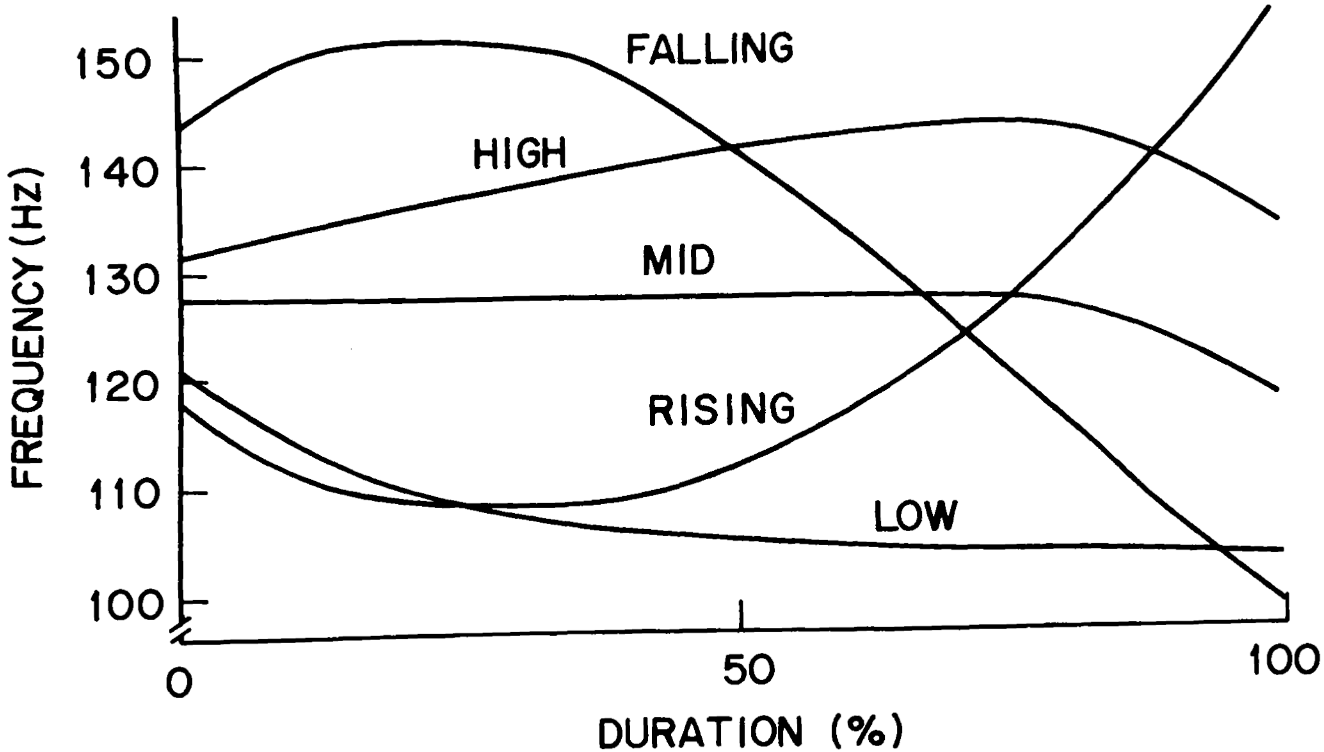

FIGURE 1

Average fundamental frequency contours of Thai tones on double vowels in word-final position (one speaker, adapted from Abramson 1962, by permission of Indiana University Research Center in Anthropology, Folklore and Linguistics).

B. STIMULI

The stimulus set included thirteen different tonal patterns (see Figure 2). There were three level tones (*11 33 55*); ten contour tones, five falling tones (*53 31 51* $\overline{53}$ $\overline{31}$) and five rising tones (*35 13 15* $\overline{35}$ $\overline{13}$). Within each subset of falling and rising tones, two pairs of tones traverse the same pitch range but differ in magnitude of slope (*53*-$\overline{53}$, *31*-$\overline{31}$:*35*-$\overline{35}$, *13*-$\overline{13}$); the remaining falling and rising tones, *51* and *15*, have the same direction and magnitude of slope as ($\overline{53}$ $\overline{31}$) and ($\overline{35}$ $\overline{13}$), respectively, but differ in the size of the pitch range. Differences in begin and end point divided the tones into the following six subsets: (*11 13 15* $\overline{13}$), (*33 31 35* $\overline{31}$ $\overline{35}$), (*55 53 51* $\overline{53}$) according to begin point; (*11 31 53* $\overline{31}$), (*33 53 13* $\overline{53}$ $\overline{13}$), (*55 35 15* $\overline{35}$) according to end point. Differences in the amount of change in fundamental frequency between the begin point and end point of the tones provide a ternary grouping of the tones: (*51 15*), (*53* $\overline{53}$ *31* $\overline{31}$ *35* $\overline{35}$ *13* $\overline{13}$), (*11 33 55*). In addition to these pitch characteristics, the thirteen stimulus tones also differed in length. Nine of the tones were "long" (*11 33 55 53 31 51 35 13 15*), four of the tones were "short" ($\overline{53}$ $\overline{31}$ $\overline{35}$ $\overline{13}$).

This particular set of thirteen tones included seven of Wang's (1967: 99) phonological tones (*11 33 55 53 31 35 13*); his other two level tones (*22 44*) and four bidirectional tones (*535 313 353 131*) were excluded.

Actual fundamental frequency values associated with the stimulus tones were intended to approximate real-speech measurements of tones (cf. Abramson 1962). All the fundamental frequency trajectories were linear: *11*, *33* and *55* had steady frequency values at 100, 125 and 150 Hz, respectively; *35*, *13*, *15*, $\overline{35}$ and $\overline{13}$ had rising frequency values from 125 to 150, 100 to 125, 100 to 150, 125 to 150 and 100 to 125 Hz, respectively; *53*, *31*, *51*, $\overline{53}$ and $\overline{31}$ had falling frequency values from 150 to 125, 125 to 100, 150 to 100, 150 to 125 and 125 to 100 Hz, respectively. The rate of change in fundamental frequency for the linear rising and falling tones was 1 Hz per 12 msecs for (*53 31 35 13*), 2 Hz per 12 msecs for ($\overline{53}$ $\overline{31}$ $\overline{35}$ $\overline{13}$ *51 15*).

These fundamental frequency trajectories were superimposed on a synthetic speech-like syllable that phonetically approximated [wa], using a line analog speech synthesiser on the PDP-12 computer at the Phonetics Laboratory, University of California, Los Angeles (for description of speech synthesiser, see Rice 1971). In this synthetic syllable, both the first and second formants displayed rising transitions into the vowel; the steady-state portion of the vowel constituted about 62 per cent of the total duration, with spectral peaks at 630, 1130 and 3300 Hz for the first three formants.