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 super-
imposed 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.

The secondary aim of the paper is to determine how close these percep-
tual 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 dis-
tances between stimulus objects based on judgments of similarity of dif-
ferent 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 addi-
tion 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 ana-
lysis of the subject space that we may determine to what extent dif-
ferences in perception, if any, may be attributed to particular individ-
uals 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. /khaa/ 'to
engage in trade', /khaa/ 'to be stuck', /khaa/ 'a kind of spice', /khaa/
'to kill', /khaa/ '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.
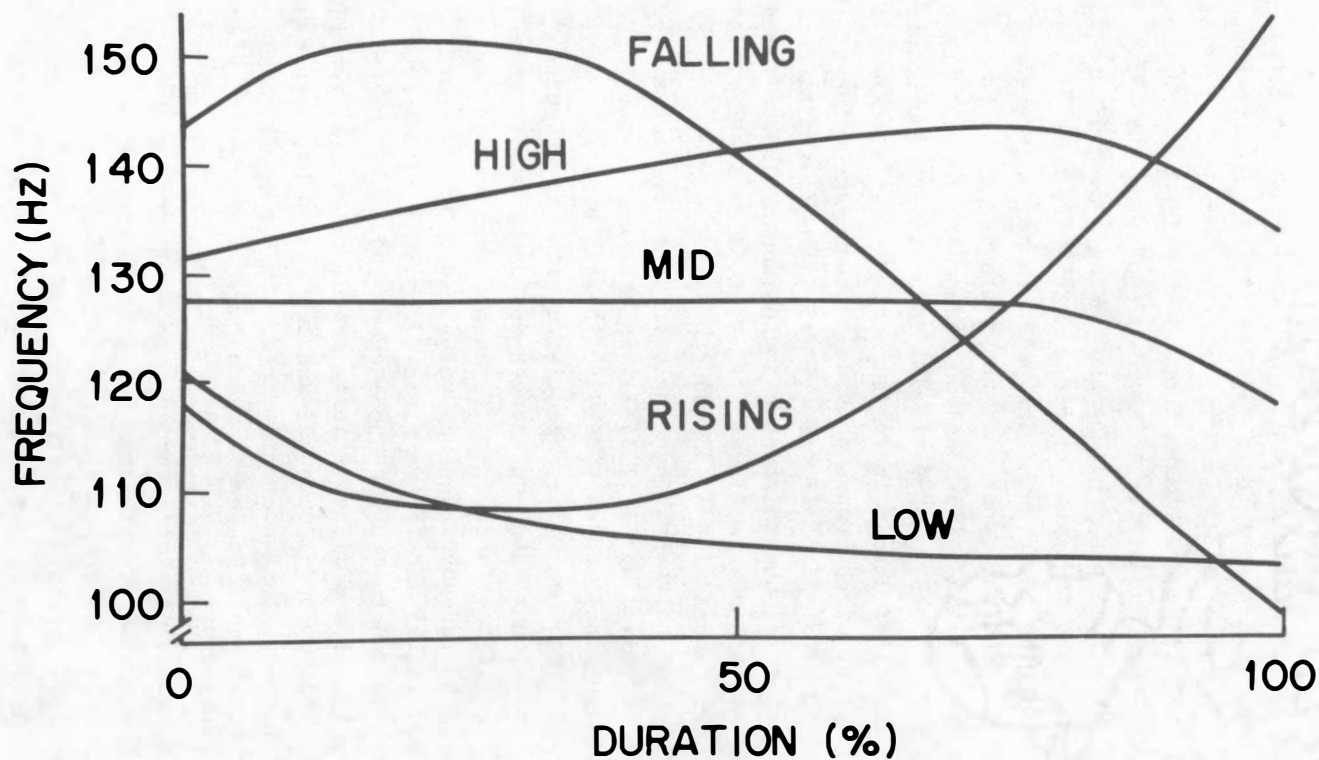
FIGURE 1

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 53 31*) and five rising tones (*35 13 15 35 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.
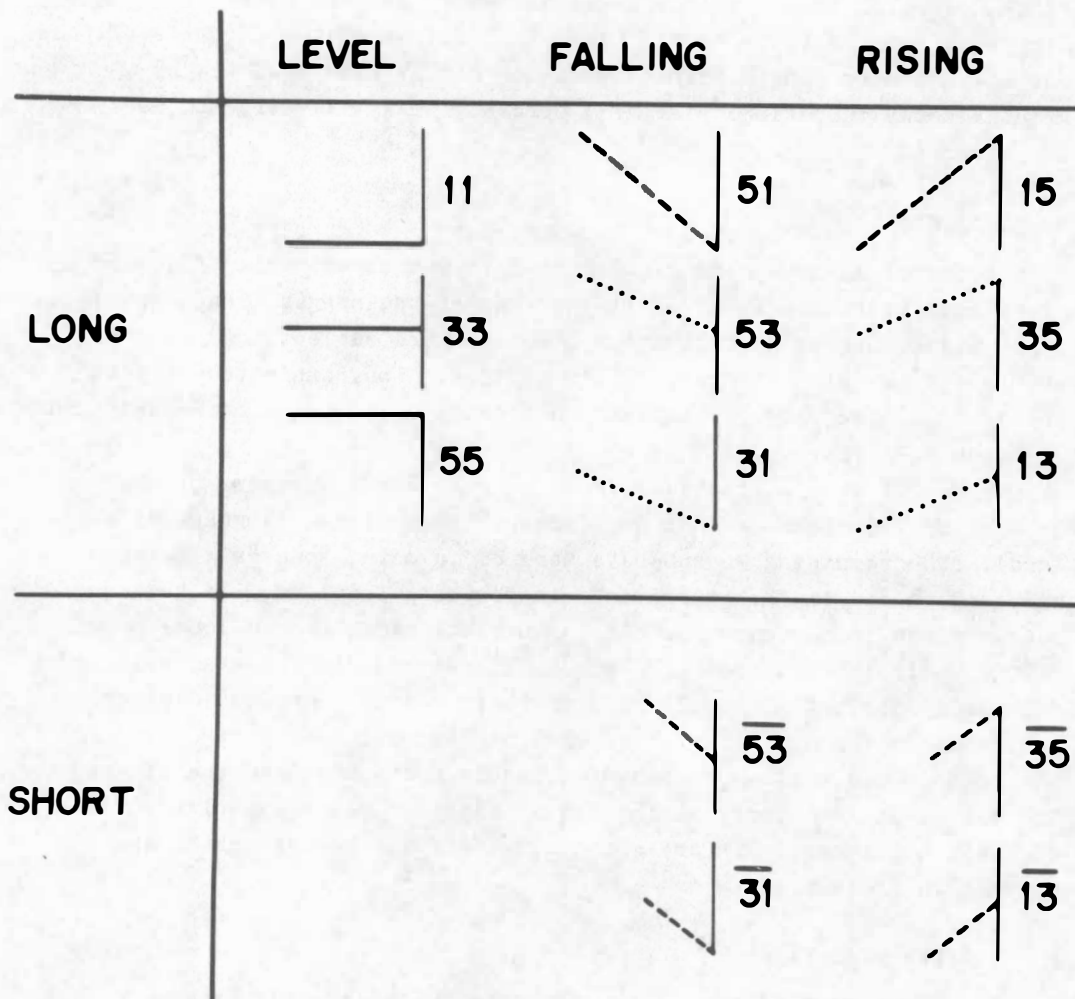
FIGURE 2

Numeric and corresponding graphic representations of the thirteen
stimulus tones used in the experiment (cf. Chao 1930). The ordinate
represents a 5-step pitch scale ranging from 5 (highest point) to 1
(lowest point), and the abscissa represents time: 55 "high level", 13
"low rising", 53 "high falling", etc. A suprascript horizontal bar
identifies the numeric representations of the "short" tones. Solid,
dashed and dotted lines along the abscissa indicate magnitude of slope
of the thirteen stimulus tones.

The duration of the syllable was 312 msecs for the nine "long" tones, 156 msecs for the four "short" tones (cf. Abramson 1962). The amplitude curve was controlled to display a rather abrupt rise followed by a gradual decay (duration: rise = 31 percent, peak = 11 percent, decay = 58 percent).

C. SUBJECTS

A total of 114 subjects participated in the experiment. All were in attendance at either a university or teacher-training college. Subjects were paid for their participation. None of the subjects had received any formal training in music or linguistics. They exhibited varying levels of proficiency in English. Subjects' ages ranged between 18 and 23, mean = 21 years.

Of the 114 Thai subjects, 38 were monodialectal speakers of the Central Thai dialect (= Thai or Siamese, the national language of Thailand). The remaining 76 subjects were bidialectal speakers: 65 were born and raised in Phuket province in southern Thailand, 11 in Chiang Mai province in northern Thailand. They all had attended local provincial schools through the tenth grade. The local dialect was their first "acquired" dialect; Central Thai was their second "learned" dialect. It is the latter dialect that is used as the medium of instruction in the schools and medium of communication in radio, TV, and the cinema. The bidialectal subjects having had at least 10 years experience with the national language in the schools, were fully conversant in the Central Thai dialect.

D. EXPERIMENTAL TASK

Subjects were told that they were going to hear words from a foreign language, and that these words all had the same sequence of consonant and vowel but different pitch patterns, as in the Thai words /naâ/ 'face' and /naˇa/ 'thick'.

The stimulus set was played twice to acquaint the subjects with the nature and range of pitch variations between the thirteen tones. They were then told that their task was to report their impression of how different the pitch patterns are between pairs of these words, by circling an appropriate number on an 11-point didimilarity scale (see Figure 3). They were also told to ignore any other differences between the words that they might hear.

THAI

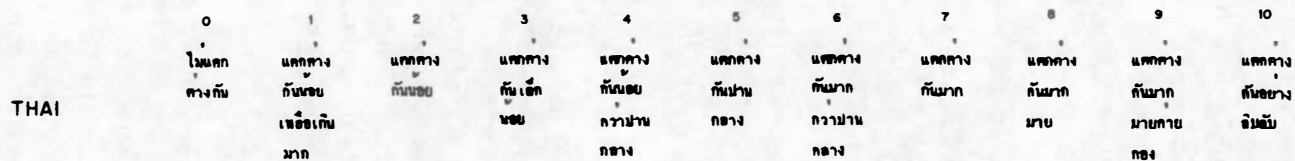| 0	1	2	3	4	5	6	7	8	9	10
ไม่แตก ต่างกัน	แตกต่าง กันน้อย เหลือเกิน มาก	แตกต่าง กันน้อย	แตกต่าง กันเล็ก น้อย	แตกต่าง กันน้อย กว่าปาน กลาง	แตกต่าง กันปาน กลาง	แตกต่าง กันมาก กว่าปาน กลาง	แตกต่าง กันมาก	แตกต่าง กันมาก มาย	แตกต่าง กันมาก มายกาย กอง	แตกต่าง กันอย่าง อินมับ

FIGURE 3

11-point scale (0 = no difference, 5 = medium difference, 10 = extreme difference) on which Thai subjects gave ratings of dissimilarity between pairs of stimulus tones.

The paired-stimuli were presented in a natural speech carrier frame
/ karunaa priàpthiâp siǎŋ khɔ̌ɔŋ kham tɔ̀ɔ pay níi /
'Compare the pitch of the following words _____'.
This natural speech context encouraged subjects to listen to the syn-
thetic stimuli in a speech mode of processing.

A trial consisted of the carrier frame, the paired-stimuli and the
response period. The time-interval between the carrier frame and the
paired-stimuli was .56 secs, between stimulus *a* and stimulus *b* of the
paired-stimuli (interstimulus interval) 0.2 secs and between the paired-
stimuli of one trial and the beginning of the carrier frame for the
next trial (intertrial interval) 4.0 secs.

Four blocks of trials were presented, each block consisting of 91
trials, yielding a total of 364 paired-comparison judgments, four judg-
ments for each stimulus-pair. Ten buffer trials were placed at the
beginning of block 1, three buffer trials at the beginning of each of
the remaining three blocks. The trials were presented under one pseudo-
random order in blocks 2 and 4. In blocks 1 and 2, the stimuli for each
stimulus-pair type were presented in order *a-b*; in blocks 3 and 4, the
order of presentation was reversed. A block of ten practice trials was
administered before proceeding into the actual experiment.

Stimuli were presented on a Uher Model IC-4000 tape recorder over a
loudspeaker in a conventional classroom setting. This enabled more than
one subject to be tested during any one session, although no more than
twelve subjects were normally run at a single session. Each subject
was seated within a comfortable hearing range of the tape recorder.
Brief rest periods were scheduled between each of the four blocks of
trials. At the conclusion of the session, subjects were asked to com-
plete a biographical information sheet and questionnaire. The entire
experiment lasted approximately 1 hour and 15 minutes.

E. METHOD OF ANALYSIS

The input to INDSCAL consisted of 114 (individual subjects) thirteen
(stimulus tones) x thirteen (stimulus tones) symmetric data matrices.
Each of the 114 individual subject data matrices contained distance
estimates for each paired-comparison of the thirteen stimulus tones.
The distance estimate was obtained by averaging the four dissimilarity
scores for each pair of stimulus tones for each subject, converting
these averaged dissimilarity scores into distance estimates, and sub-
sequently into scalar products using standard procedures described in
Torgerson (1958). The aim of INDSCAL is to determine, by means of an
iterative least squares procedure, the stimulus coordinates and the

subject weights that account for as much total variance in all
subjects' data as possible. The output from INDSCAL, consisted of two
matrices, a thirteen (stimulus tones) x n (dimensions) matrix of co-
ordinates of the thirteen stimulus tones on n dimensions (plotted in a
"group stimulus space"), and a 114 x n matrix of weights of each of the
114 individual subjects on the same n dimensions (plotted in a "subject
space"). INDSCAL analyses of the subjects' input dissimilarity matrices
were performed in several n-dimensional spaces in order to determine
the correct number of dimensions.

III. RESULTS

A. *INTERPRETATION OF DIMENSIONS*

A 4-dimensional INDSCAL analysis of these dissimilarities data was
found to provide the best summary interpretation of the major perceptual
structures employed by subjects in making their paired-comparison
ratings of dissimilarity between the thirteen stimulus tones. This
4-dimensional solution accounted for 74% of the total variance in the
dissimilarities data for all subjects. Since only 4% more variance was
accounted for in five dimensions and 9% less was accounted for in three
dimensions, it was concluded that four dimensions were both necessary
and sufficient. The high correlations (.81 or above) obtained between
each individual subject's original dissimilarities data and the distances
between the stimulus tones in his associated appropriately-weighted
4-dimensional solution indicate that the private stimulus space for a
particular individual may be derived quite reasonably from the group
stimulus space by differentially stretching or shrinking the four dimen-
sions.

The four dimensions that emerged from the multidimensional scaling
analysis are numbered from 1 to 4 in decreasing order of variance ac-
counted for: dimension 1 = 26.5%, dimension 2 = 20.4%, dimension 3 =
19.8% and dimension 4 = 7.3%. The proportion of variance accounted for
by each of the four dimensions indicates that dimension 4 was somewhat
weaker in saliency than the other three dimensions. Although dimension 4
accounted for a smaller proportion of the variance, its statistical
reliability as well as the reliability of the composite 4-dimensional
group stimulus space was confirmed through the computation of correla-
tion coefficients between each of the four dimensions for separate in-
dependent split-half (57 subjects) solutions. If the stimulus co-
ordinates between corresponding dimensions of the split-half solutions
correlate highly, then we can reasonably conclude that the total group
solution is very reliable. The correlations between stimulus coordinates

on the corresponding unrotated dimensions of the split-half solutions
were as follows: dimension 1 = .98%, dimension 2 = .99%, dimension 3 =
.98% and dimension 4 = .90%. The moderately high replicability of
dimension 4 across split-half INDSCAL analyses establishes its statis-
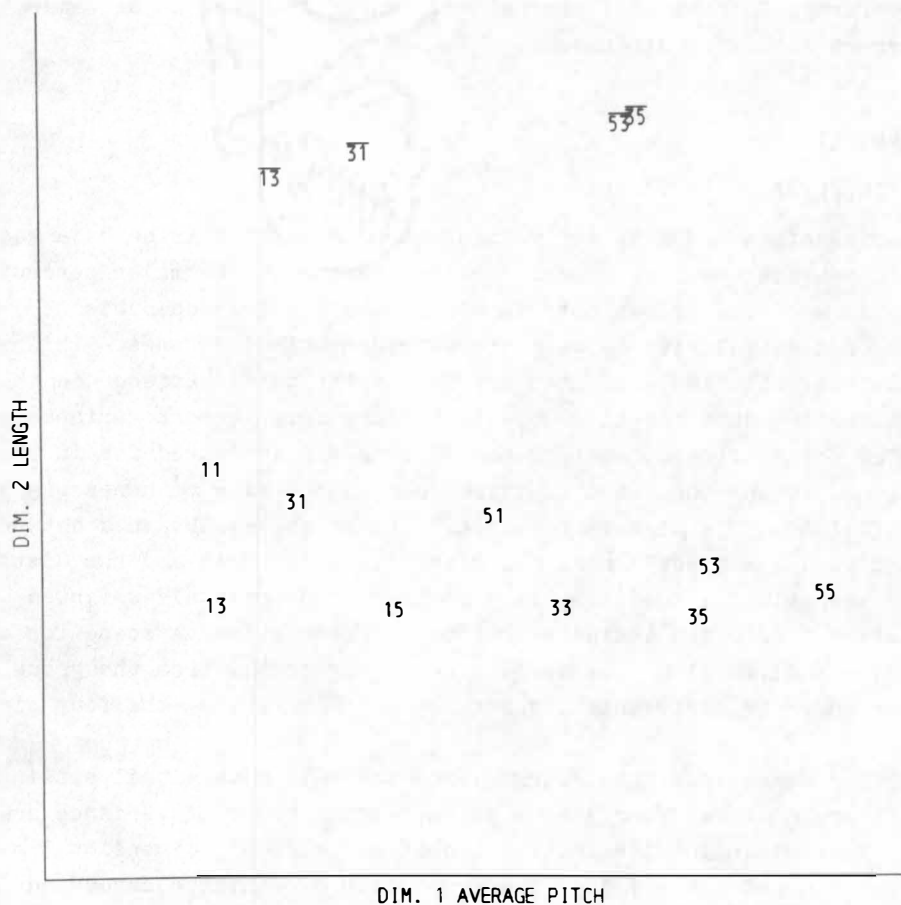tical reality.

FIGURE 4

Dimensions 1 and 2 from the INDSCAL 4-dimensional group stimulus space.

Figure 4 shows a plot of the coordinates or projections of the
stimulus tones on the first and second dimensions in the group stimulus
space. Interpretation of the dimensions was guided by inspection of the

order and position of the tones along each dimension, particularly how
the tones at one extreme differ from those at the other. It should be
pointed out, however, that the locations of the stimulus tones in the
4-dimensional group stimulus space came directly from the INDSCAL ana-
lysis, and preceded any particular interpretation assigned to the dimen-
sions.

Dimension 1 is interpreted as AVERAGE PITCH. Stimulus tones *11* and
55 are found at either extreme; *33* near the middle. The clustering of
15 and *51* with *33*, and the clustering of ($\overline{13}$ $\overline{31}$ *31 13*) and ($\overline{53}$ $\overline{35}$ *53 35*)
lead to the following assignment of average pitch values: *55* = 5,
($\overline{53}$ $\overline{35}$ *53 35*) = 4, (*15 51 33*) = 3, ($\overline{13}$ $\overline{31}$ *31 13*) = 2, *11* = 1. The
second dimension, interpreted as LENGTH, contrasts the nine long tones
(*11 13 31 15 51 33 35 53 55*) with the four short tones ($\overline{13}$ $\overline{31}$ $\overline{53}$ $\overline{35}$).
This dimension is based on the duration, rather than pitch, characteris-
tics of the thirteen synthetic tonal stimuli.

Figure 5 (overleaf) shows a plot of the third and fourth dimensions.
Both of these dimensions are based on the pitch characteristics of the
stimulus tones. Along the third dimension, interpretively named DIREC-
TION, rising tones (*15 35 13* $\overline{35}$ $\overline{13}$) are found at one extreme, falling
tones (*51 31* $\overline{31}$ *53* $\overline{53}$) at the other and level tones (*33 11 55*) near the
middle. The fourth dimension, labelled SLOPE, appears to separate the
level tones (*33 11 55*) from the contour tones (*51 31 53* $\overline{31}$ $\overline{53}$ $\overline{13}$ $\overline{35}$ *13*
35 15). Along this dimension, tones tend to be ordered by the size of
the pitch range as determined by the difference in fundamental frequency
values between the beginning and ending points of the stimulus tones:
(*51 15*) = 50 Hz, (*31* $\overline{31}$ *53* $\overline{53}$ $\overline{13}$ $\overline{35}$ *13 35*) = 25 Hz, (*33 11 55*) = 0 Hz.
This dimension also appears to be related to the magnitude of slope of
the stimulus tones.

The interpretation of the four dimensions was supported by the re-
sults of multiple linear regression analysis. Table I (page 287) pre-
sents the eight properties that were treated as dependent variables;
the stimulus coordinates of the four dimensions in the group stimulus
space were treated as independent variables. The regression analysis
locates directions in the group stimulus space best corresponding to
these properties. Of the eight properties, 1-4 are of a binary nature,
6-8 tertiary and property 5 quinary. Properties 1-3 correspond to
Wang's (1967) binary phonological features of tone [±CONTOUR], [±RISING]
and [±FALLING], respectively. Property 5 corresponds to duration
characteristics of the stimulus tones, properties 5-8 various pitch
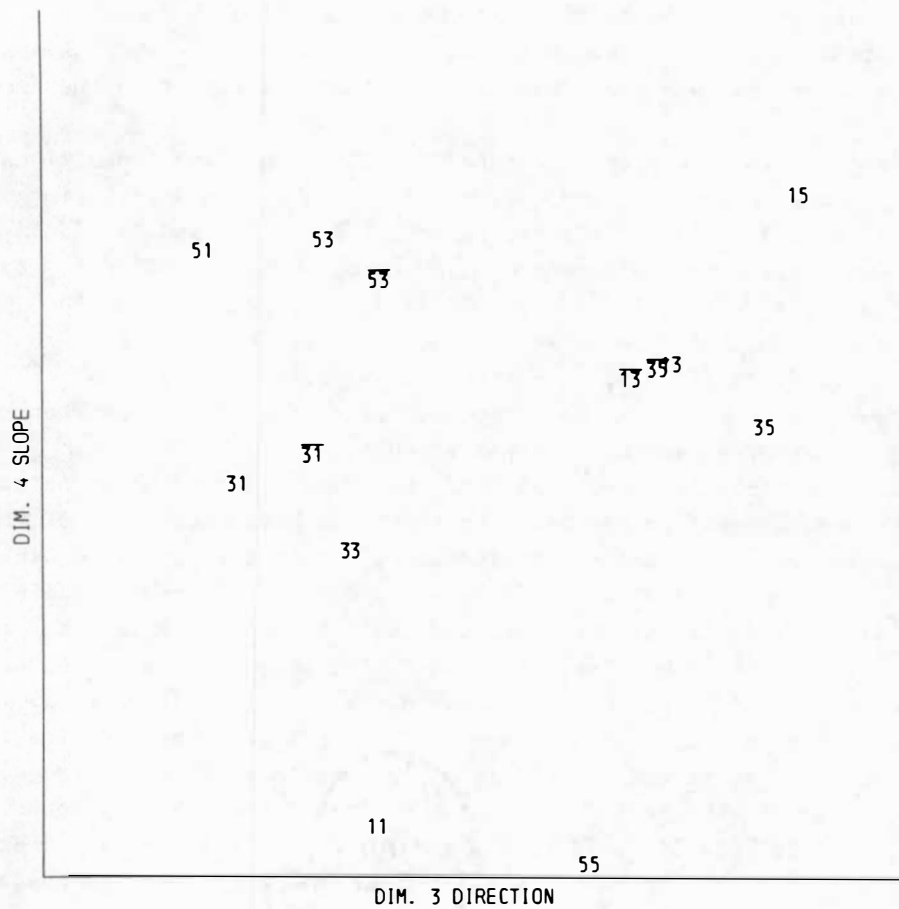characteristics (cf. Section II.8).

FIGURE 5

Dimensions 3 and 4 from the INDSCAL 4-dimensional group stimulus space.

	11	33	55	53	31	51	35	13	15	5̄3̄	3̄1̄	3̄5̄	1̄3̄
1. Level *v.* Contour	0	0	0	1	1	1	1	1	1	1	1	1	1
2. Rising *v.* Nonrising	0	0	0	0	0	0	1	1	1	0	0	1	1
3. Falling *v.* Nonfalling	0	0	0	1	1	1	0	0	0	1	1	0	0
4. Long *v.* Short	0	0	0	0	0	0	0	0	0	1	1	1	1
5. Average Fundamental Frequency	100	125	150	137.5	112.5	125	137.5	112.5	125	137.5	112.5	137.5	112.5
6. Magnitude of Slope	0	0	0	1	1	2	1	1	2	2	2	2	2
7. Amount of Change in Fundamental Frequency (Hz)	0	0	0	25	25	50	25	25	50	25	25	25	25
8. Rising *v.* Level *v.* Falling	0	0	0	1	1	1	2	2	2	1	1	2	2

TABLE I

Values associated with the thirteen stimulus tones for each of the eight properties used in the linear regression analysis.

The associated linear regression results are shown in Table II. The multiple correlations in the last column show how well each of the properties can be predicted from the form INDSCAL dimensions. The normalised regression weights show the relative importance of each dimensions for predicting order and position of stimulus tones for a particular property.

Regression Weights

	Dim. 1	Dim. 2	Dim. 3	Dim. 4	Multiple Correlation
1. Level v. Contour	−.124	−.235	.115	.957*	.882
2. Rising v. Nonrising	−.240	−.070	.925*	.288	.958
3. Falling v. Nonfalling	.151	−.114	−.868#	.459	.901
4. Long v. Short	−.055	.988*	−.135	−.051	.986
5. Average Fundamental Frequency	.981*	−.008	.157	.111	.980
6. Magnitude of Slope	−.086	−.584	.131	.796#	.933
7. Amount of Change in Fundamental Frequency	−.158	.094	.053	.982*	.864
8. Rising v. Level v. Falling	−.224	−.167	.677#	.681#	.936

TABLE II

Linear regression of eight properties on 4-dimensional space for Thai subjects' dissimilarity ratings data.   * = regression weight .925 and above; # = regression weight between .675 and .925.

Figures 6 and 7 show the geometric representation of the linear re-
gression results in Table II. In this representation, the direction
and extent of the straight (dashed) line corresponding to a property
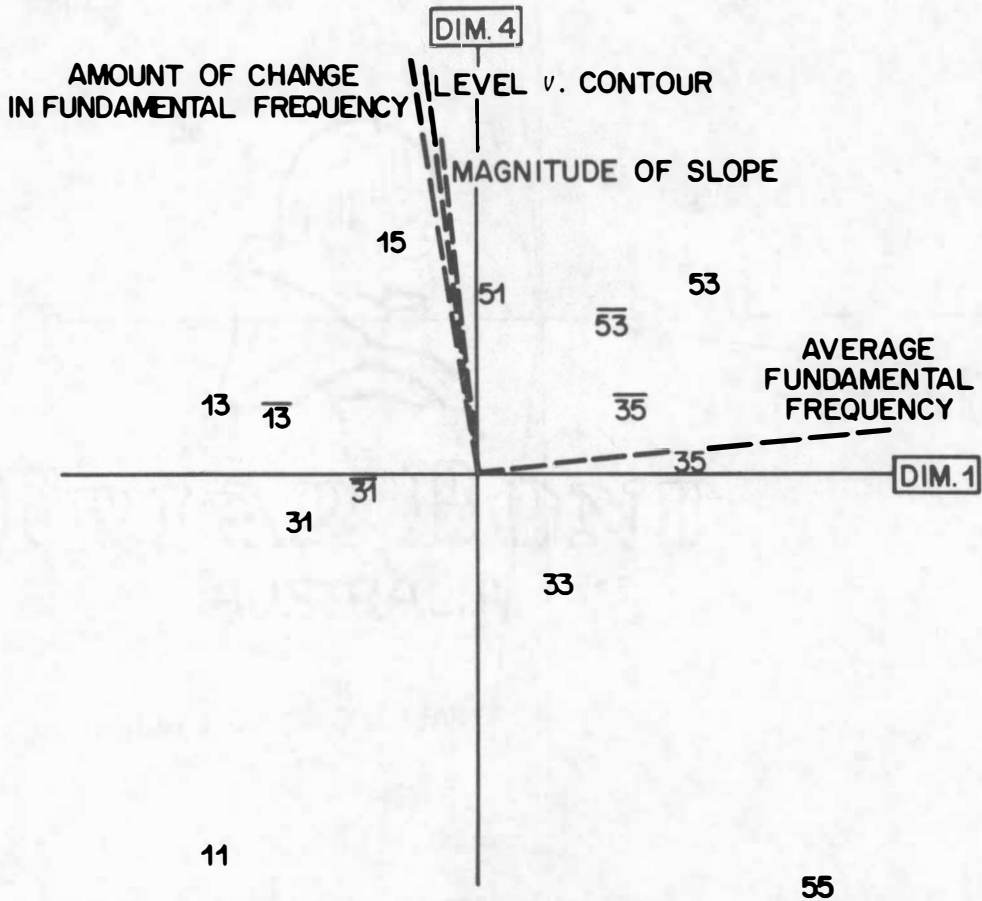indicates which of the INDSCAL dimensions most nearly corresponds to
that property.

FIGURE 6

The dimension 1 - dimension 4 plane from the INDSCAL 4-dimensional
group stimulus space. Projections of straight (dashed) lines op-
timally corresponding to amount of change in fundamental frequency,
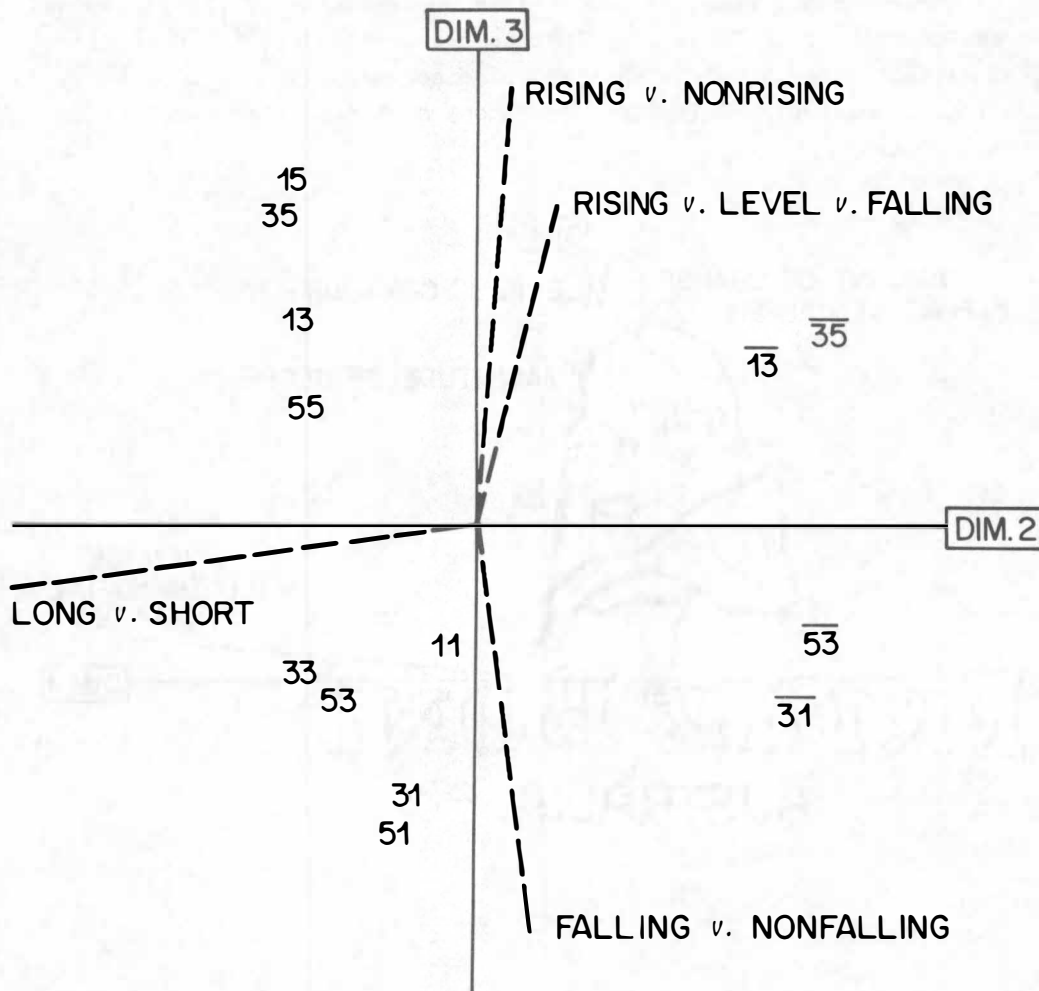level versus contour, magnitude of slope and average fundamental
frequency.

FIGURE 7

The dimension 2 - dimension 3 plane from the INDSCAL 4-dimen-
sional group stimulus space. Projections of straight (dashed)
lines optimally corresponding to rising versus nonrising, rising
versus level versus falling, long versus short and falling
versus nonfalling.

The direction and extent of the straight lines optimally correspond-
ing to these properties in the multidimensional space support our inter-
pretation of the four dimensions. In Figure 6, dimension 1 closely cor-
responds to average fundamental frequency, dimension 4 to amount of
change in fundamental frequency, level versus contour and magnitude of
slope; in Figure 7, dimension 2 closely corresponds to long versus short,
dimension 3 to rising versus nonrising, falling versus nonfalling and
rising versus level versus falling. The third dimension, however, does

suggest that rising (property 2) pitch movement is represented more
directly in terms of subjects' perceptual dimensions of tone than fal-
ling pitch movement. It is of further interest to note that the clus-
tering of stimulus tones on the third and fourth dimensions lead to a
close correspondence to Wang's binary phonological features of tone
(properties 1-3).

8. SUBGROUP DIFFERENCES IN WEIGHTS FOR DIMENSIONS

Subjects' weights on the AVERAGE PITCH and LENGTH dimensions are
indicated in Figure 8. The relative salience of AVERAGE PITCH is
greater than LENGTH for most of individual Thai subjects as illustrated
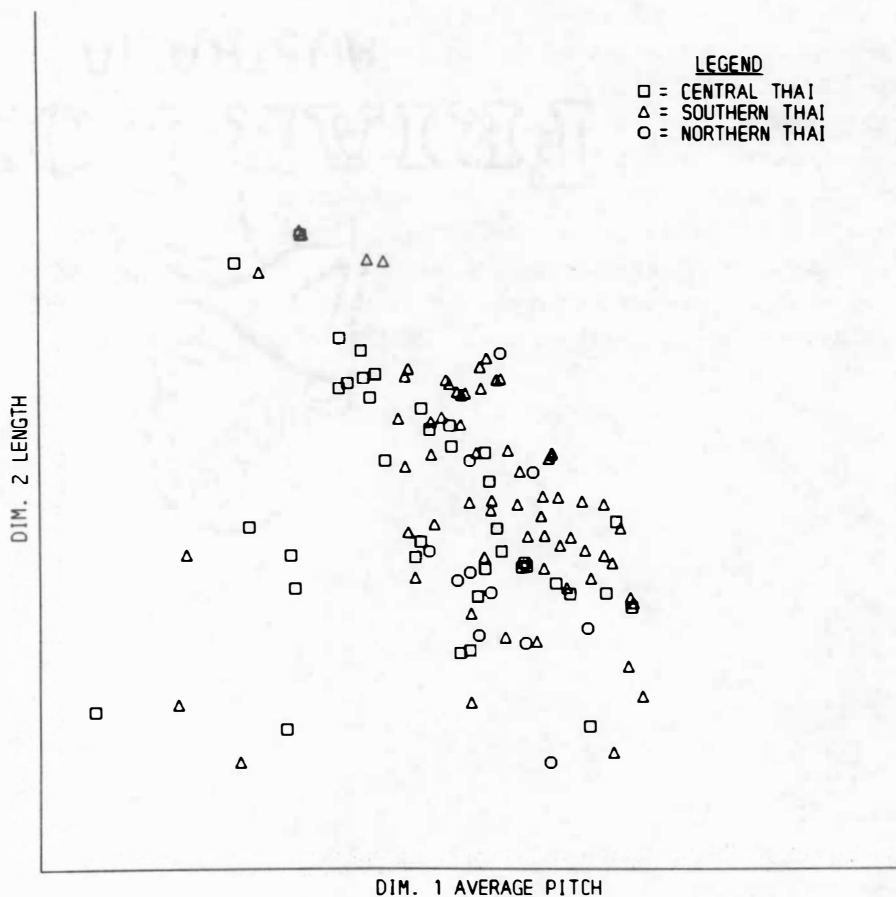by the tendency for subjects to cluster in the middle-right area of the
plot.

FIGURE 8

Dimensions 1 and 2 from the INDSCAL 4-dimensional subject space.

A 1-way analysis of variance of the *mean* subject weights on each of the
dimensions across the three Thai dialect subgroups, however, does indi-
cate a significant difference on dimension 1 - AVERAGE PITCH between
the Central Thai and Southern Thai subgroups (F = 7.08, d f = 1,
p < .01). No significant differences were obtained between Thai dialect
subgroups on the second, third and fourth dimensions. Though the mean
weights for dimension 1 differ significantly between Central Thai and
Southern Thai, the considerable overlap between dialect subgroups calls
for caution in interpreting this apparent difference in the perceptual
spaces of these two dialect subgroups (cf. Brown 1965:59). The essen-
tial homogenity of the perceptual space for the Thai group can be at-
tributed, at least in part, to the fact that both Northern Thai and
Southern Thai subjects were educated bidialectals who had had several
years exposure to Central Thai; and the possibility that the differences
in the phonological inventory of tones between the three dialect sub-
groups were simply not great enough to induce differences in the percep-
tual space (for description of tones of Thai dialects, see Brown 1965).
For similar apparent reasons, Fox's (1974) multidimensional scaling
analysis of vowel perception failed to demonstrate a difference in the
perceptual spaces utilised by two dialect subgroups of American English.

The essential homogeneity of the perceptual space for the Thai group
is further shown in Figure 9. Subjects' weights on the third and fourth
dimensions tend to cluster in the lower-right area of the plot indi-
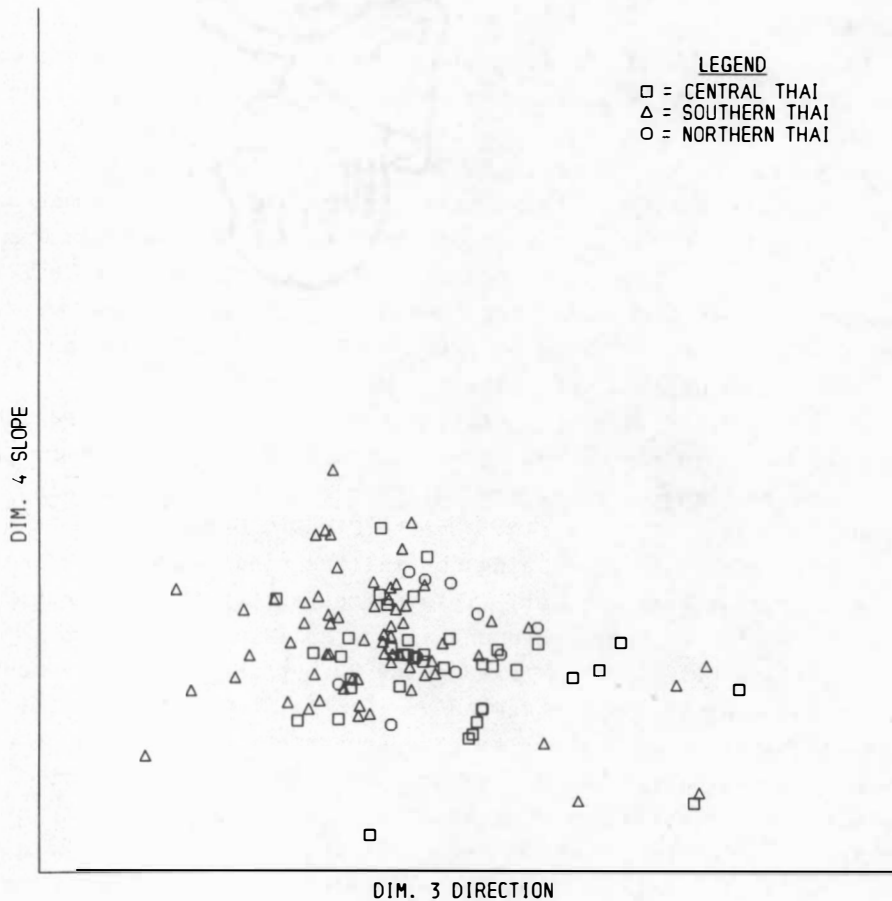cating relatively greater salience for DIRECTION than for SLOPE.

FIGURE 9

Dimensions 3 and 4 from the INDSCAL 4-dimensional subject space.

IV. DISCUSSION AND CONCLUSIONS

Our data indicate that four dimensions underlie the subjects' dis-
similarity ratings of the thirteen synthetic stimulus tones. Of the
four dimensions, AVERAGE PITCH was the most important factor that sub-
jects incorporated into their perceptual judgments. This dimension
would appear to be a lower-level auditory dimension; its utilisation at
higher levels of linguistic processing is certainly not indicated by the

phonological inventory of Thai tones, or for that matter, the lexical
tones of any tone language. Hombert's (1976) multidimensional scaling
analysis of the perception of tones of bisyllabic nouns in Yoruba (tone
language spoken in West Africa), interestingly enough, also extracted a
dimension that was, in part, related to the average fundamental frequency
value of the vowel in the second syllable.

The emergence of the nontonal LENGTH dimension as a relatively im-
portant factor in Thai subjects' dissimilarity judgments is not at all
surprising in view of the structure of the stimulus set, and perhaps
more importantly, the linguistic function of vowel duration in Thai.
The contrast between long and short vowels (e.g. /hàat/ 'shoal' v. /hàt/
'to practice') more than likely heightened the Thai subjects' sensitivity
to this particular physical property of the stimuli.

The tonal third and fourth dimensions too seem to be of a more lin-
guistic nature. Both may be related to pitch characteristics that are
not only used to signal phonological distinctions in Thai, but also
other tone languages of the world. On the DIRECTION dimension, it may
be observed that neither the rising nor falling tones cluster according
to the beginning or ending point. This suggests that the Thai subjects
perceived direction of pitch movement in the stimulus tones, not as move-
ment *from* a fixed point A *to* a fixed point B, but instead as movement
in a direction *away from* A and *toward* B. On the SLOPE dimension, it may
be observed that the level tones tend to separate from the contour tones.
This dimension may be related to Abramson's (1962) proposed "static"
versus "dynamic" classification of the Thai tones. In his Yoruba study,
Hombert too found dimensions that were principally related to the direc-
tion and amount of change in fundamental frequency. The convergence in
results from these two studies clearly strengthens the claim for the
psychological reality of these dimensions.

The results of a linear regression analysis supported our interpreta-
tion of the dimensions. Of particular interest is the close correspon-
dence obtained between the third and fourth dimensions and Wang's unit-
contour tone features. There is no physical property of the stimulus
set that would bias the obtained result; indeed, one might have pre-
dicted the emergence of dimensions on which the stimulus tones clustered
according to begin point and/or end point. The fact that the decomposi-
tion of the contour stimulus tones into begin point and end point appar-
ently had little influence on Thai listeners' perceptual judgments
converges nicely with the traditional classification of Thai as a
"contour tone language" (Pike 1948, Ladefoged 1975).

Contrary to expectation, individuals did not differ much across
dialect subgroups. This, of course, does not mean that dialect member-

ship may not have an influence on one's tone perception. The dominant
influence of the national language could be effectively eliminated or
restricted through the use of monodialectal speakers of Northern Thai
and Southern Thai dialects. With monodialectal speakers, we might then
be able to furnish more direct experimental evidence bearing on Brown's
(1965:59) rather appealing suggestion that "contour" is more important
than "register" for Central Thai and Northern Thai dialects, but that
"register" is more important than "contour" for Southern Thai dialects.
Individual differences in tone perception have been demonstrated across
language groups - Thai, Yoruba and American English (Gandour and
Harshman 1978).

The perception of tone deals with how a listener transforms, organises
and structures the pitch information arising from the speech signal.
The nature of the dimensions that emerged from the INDSCAL analysis are
clearly consistent with a proposal that information-bearing aspects of
pitch are organised in terms of simple oppositions (binary or n-ary)
along a number of independent dimensions rather than in terms of a
number of steps along one or a few dimensions. Of the four dimensions,
the second, third and fourth are apparently utilised in signalling lin-
guistic categories in Thai, and consequently, show a greater tendency
toward binarity. These dimensions, though obviously not exhaustive of
all possibilities, represent a beginning toward a definition of a
necessary and sufficient set of perceptually-based features of tone.

REFERENCES

ABRAMSON, A.S.

1962   'The Vowels and Tones of Standard Thai: Acoustical Measure-
ments and Experiments'. *International Journal of American
Linguistics* 28/2, Part II.

BROWN, J.M.

1965   *From Ancient Thai to Modern Dialects*. Bangkok: Social
Science Association Press of Thailand.

CARROLL, J.D. and J.J. CHANG

1970   'Analysis of Individual Differences in Multidimensional
Scaling via a n-way Generalization of "Eckart-Young" Decom-
position'. *Psychometrika* 35/3:283-319.

CARROLL, J.D. and M. WISH

1974a   'Multidimensional Perceptual Models and Measurement Methods'.
In: E.C. Carterette and M.P. Friedman, eds. *Handbook of Per-
ception*, vol. II: *Psychophysical Judgment and Measurement*.
391-447. New York: Academic Press.

1974b   'Models and Methods for Three-way Multidimensional Scaling'.
In: D.H. Drantz *et al.*, eds. *Contemporary Developments in
Mathematical Psychology*, vol. II. 57-105. New York:
W.H. Freeman.

CHANG, J.J. and J.D. CARROLL

1972   How to use PROFIT, a computer program for PROperty FITing by
optimizing nonlinear or linear correlation. Bell Labora-
tories, Murray Hill, New Jersey.

CHAO, Y.R.
   1930   'A System of Tone Letters'. *La Maitre Phonetique* 45/24-7.

FOX, R.A.
   1974   An Experiment in Cross-dialect Vowel Perception. Papers from the 10th Regional Meeting of the Chicago Linguistic Society. 178-85.

GANDOUR, J.
   1975   'On the Representation of Tone in Siamese'. In: J.G. Harris and J.R. Chamberlain, eds. *Studies in Tai Linguistics*. 170-95. Bangkok: Central Institute of English Language.

GANDOUR, J. and R. HARSHMAN
   1978   'Crosslanguage Differences in Tone Perception: a Multi-dimensional Scaling Investigation'. *Language and Speech* 21/1-33.

HARSHMAN, R.A.
   1970   'Foundations of the PARAFAC Procedure: Models and Conditions for an "Explanatory" Multidimensional Factor Analysis'. *UCLA Working Papers in Phonetics* 16.

   1972   'Determination and Proof of Minimum Uniqueness Conditions for PARAFAC 1'. *UCLA Working Papers in Phonetics* 22/11-117.

HENDERSON, E.J.A.
   1949   'Prosodies in Siamese, a Study in Synthesis'. *Asia Major* (new series) 1/189-215. Reprinted in: F.R. Palmer, ed. *Prosodic Analysis*. 1970:27-53. London: Oxford University Press.

HOMBERT, J.M.
   1976   'Perception of Tones of Bisyllabic Nouns in Yoruba'. *Studies in African Linguistics*, Supplement 6. 109-21.

LADEFOGED, P.
   1975   *A Course in Phonetics*. New York: Harcourt Brace Jovanovich.

PIKE, K.L.
   1948   *Tone Languages*. Ann Arbor: University of Michigan Press.

RICE, L.
1971   'A New Line Analog Speech Synthesizer for the PDP-12'. *UCLA Working Papers in Phonetics* 17/58-75.

TORGERSON, W.S.
1958   *Theory and Methods of Scaling*'. New York: Wiley.

WANG, W. S-Y.
1967   'Phonological Features of Tone'. *International Journal of American Linguistics* 33/93-105.

WISH, M. and J.D. CARROLL
1974   'Applications of Individual Differences Scaling to Studies of Human Perception and Judgment'. In: E.C. Carterette and M.P. Friedman, eds. *Handbook of Perception* vol. II: *Psychophysical Judgment and Measurement*. 449-91. New York: Academic Press.