

# THE TONES OF CENTRAL THAI: SOME PERCEPTUAL EXPERIMENTS

Arthur S. Abramson

The University of Connecticut and Haskins Laboratories<sup>1</sup>

## Introduction

In recent years, research into prosodic features as part of the effort to understand the nature of speech communication has become an increasingly conspicuous aspect of experimental phonetics (Fry 1968, Lieberman 1974). Within prosody, such features as phonologically distinctive stress, segment length and tone are clearly central to the sound pattern of the language in that they differentiate linguistic expressions. These are usually of more immediate interest to the linguist than are other types of prosodic features, and it is usually easier to design experiments testing hypotheses concerning them.<sup>2</sup> Prosodic research of this kind should make contributions to general phonology and, more narrowly, to our understanding of the phonology of a particular language, in this instance, Thai.

The research to be reported here forms part of a larger project other aspects of which will be presented elsewhere. It is hoped that the several studies ultimately to appear will furnish much information on the perceptually tolerable ranges of the tones of Thai. The work takes as its starting point the usual assumption that the major phonetic features of phonemic tone are found in the domain of pitch. The primary acoustic correlate of pitch is, of course, frequency. In instrumental analyses of tonal features, then, we measure the fundamental frequency of the voice as determined by the repetition rate of glottal pulsing.

The language under analysis here is Central Thai (Siamese), the regional dialect which serves as the official language of Thailand. All native speakers used were from Bangkok or its close environs. The question of the tonal homogeneity of the central area of the kingdom has not been well explored<sup>3</sup> and thus cannot be categorically ruled out as a perturbing factor in some of the data to be presented; nevertheless, the group of test subjects gave an impression of general uniformity in tonal behavior.

One aim of the present study was to determine how well the five tones of the language could be identified in isolation. It is at least conceivable that the identifiability of one or more of the tones would suffer without the benefit of an immediate context. Another aim was to reaffirm earlier work (Abramson 1961, 1962) on the sufficiency of certain ideal fundamental frequency contours for the identification of the tones, using synthetic speech in which it would be possible to make frequency

contours the only variable. In the expectation that less than perfect identification would be achieved with fundamental frequency alone, it was also planned to learn whether the addition of variations in the amplitude of the speech signal would enhance the identifiability of the tones. Finally, it was proposed to test the strong hypothesis that absolute fundamental-frequency heights contribute nothing to the identification of the tones but that it is rather the shapes of the frequency contours that carry all the information.

### Background Work

Much of the present work is a continuation of earlier work done by the author (Abramson 1962) with fewer informants and test subjects. In that study it was shown (p. 128) that sets of tonally differentiated monosyllabic words, as produced by a single speaker, could be correctly identified nearly 100% of the time. In addition, fundamental-frequency ( $f_0$ ) measurements were taken from a large sampling of monosyllabic words with both short and long vowels, yielding average contours for the five tones (pp. 112-127). It is gratifying to note that in data obtained some fourteen years later<sup>4</sup> Erickson (1974) provides general verification of the old contours while adding important information on the perturbing effects of initial consonants.

These average  $f_0$  contours were then synthesized to see if Thai speakers could indeed identify each of the tones on the basis of  $f_0$  alone. The synthesizer used for the purpose was the Haskins Laboratories Intonator (Abramson 1962:20), which enabled the experimenter to analyze the spectrum of the speech signal and then resynthesize it on the machine's own "voice" source with new  $f_0$  contours. Thus, most of the phonetic features of the original utterance are kept even when a  $f_0$  contour is imposed. The five average tonal curves were thus imposed upon syllables that had originally carried the mid tone, namely /khaj/ 'dried sweat' and /loo/ 'unstable'.<sup>5</sup> The two perception tests prepared in this way exposed native Thai listeners to the tonal contours on both short and long vowels. The perceptual labeling of the randomized stimuli of these two tests showed clearly that the isolated  $f_0$  contours provided sufficient cues for identifying the tones (pp. 131-2). In another experimental condition, five monosyllabic words minimally differentiated by tone were manipulated with the Intonator to yield 25 new syllables. That is, five new syllables were derived from each spoken syllable by removing the original tonal contour and replacing it with the synthetic contours, one by one, on this syllable carrier. This was done to see whether the curves carried enough information to override the effects of other features found in association with the pitch movements of the tones.<sup>6</sup> The results (pp. 131-4) included a small number of confusions apparently attributable to such concomitant features as variations in amplitude and duration, which would be likely to survive the analysis and resynthesis, although by and large the  $f_0$  curves were heard as intended. In general, then, the data supported the conclusion that the  $f_0$  contours isolated by means of acoustic analysis furnished sufficient cues by themselves for the identification of the five tones of Central Thai.



## Experiment 1: Isolated Monosyllabic Words

Earlier work (Abramson 1962:128) indicated that Thai listeners can identify the tones of monosyllabic utterances nearly perfectly. Although four sets of tonally differentiated words were used in that study, it is true that the productions of only one speaker furnished the stimuli. It has been argued by R.B. Noss (Personal communication) that generalizing from these results is, perhaps, not warranted and that normally the mid and low tones, at least, are difficult to identify unless they are embedded in a context. It was thought that this objection might be handled by replicating the experiment with a somewhat larger number of speakers and listeners. In addition, data from responses to real speech were needed to furnish a standard for the later evaluation of results obtained with synthetic speech.

The following set of words was chosen for all the experiments to be described:

Tone	Thai Script	Transcription	Gloss
Mid	คก	/khaa/	'a grass (Imperata Cylindrica)'
Low	ค่า	/khàa/	'galangal, a rhizome'
Falling	ข้า	/khâa/	'slave, servant'
High	ค้ำ	/kháa/	'to engage in trade'
Rising	ขา	/khǎa/	'leg'

The tone names are conventional but not fully descriptive. Ten native speakers of Central Thai recorded three or four randomizations each of the list with short pauses such that there were five tokens of each word in each randomization. The speakers, five men and five women, included nine university instructors and one clerk.

The tests were played to twenty-five native speakers of Central Thai through headphones in the language laboratory<sup>7</sup> of Ramkhamhaeng University, Bangkok twice a week for a month. Only one test order was used for each of the ten speakers. Here and in later experiments the subjects were instructed to write the numerals 0, 1, 2, 3 or 4 for the mid, low, falling, high and rising tones respectively. These numbers are appropriate to the nomenclature of traditional Thai grammar and facilitated later scoring. Familiar as the subjects were with this convention, they were nevertheless provided with a sheet at each session showing the five words in Thai script with their numerical equivalents.<sup>8</sup>

The results are shown in Table 1, which is arranged as a confusion matrix with the tonal stimuli in the first column and the perceptual responses to them in the next five columns. The sixth column shows the total number of responses to each stimulus word. A perfect response to the tones as intended would yield 100% in each cell along

% Responses

Stimuli	Lables:	Mid	Low	Falling	High	Rising	N
	Mid	97.9	2.1				1220
	Low	3.4	96.6				1220
	Falling		0.2	99.1	0.4	0.3	1220
	High				100.		1220
	Rising		0.1	0.4		99.5	1220

Total = 6100

Subjects = 25

% Correct = 98.6

Table 1. Confusion matrix of real-speech responses.

the diagonal from the upper left to the lower right. The overall intelligibility of 98.6%, 85 errors out of 6100 responses, is high. Inspection of the responses to the mid and low tones indicates that some confusion between them accounts for most of the small number of errors.

The data were also examined for individual differences among speakers and listeners. One of the ten speakers caused 45.9% of the errors, another speaker, 16.5%, and a third speaker, 12.9%. The recordings of only one speaker produced no errors at all. As for inter-subject differences, the worst listener made 12.9% of the errors, followed by two who each made 8.2% of the errors. Three of the 25 subjects made no errors at all.

This larger sampling of speakers and listeners supports the earlier conclusion that Thai words minimally differentiated by tones can nearly always be identified correctly even as isolated forms. Some speakers appear to provide minimal perceptual cues for words out of context, especially for the contrast between mid and low tones. Some listeners seem to require more than these minimal cues for the identification task. Finally, the data in Table 1 provide a baseline for the other experiments now to be discussed.



## SYNTHETIC SPEECH

### Experiment 2: Perception of $f_0$ contours

To test once again the perceptual efficacy of the  $f_0$  contours derived from speech measurements in the earlier study (Abramson 1962: 112-27), a different speech synthesizer, the Haskins Laboratories formant synthesizer, was used under control of a computer. For the present experiment, the parameters specified were the frequency and amplitude values of the first three formants, the timing of source functions for voicing and voicelessness, the overall amplitude of the signal and the fundamental frequency ( $f_0$ ). Steady-state formant frequencies were chosen to yield a vowel acceptable as Thai /aa/; formant transitions (Liberman et al. 1954) appropriate to the velar place of articulation were used, and voiceless aspiration was simulated by providing a long voicing lag (Lisker and Abramson 1964, 1970) filled with turbulent noise in the regions of the formants. The overall amplitude was kept flat throughout the syllable except for a slight rise at the beginning and a slight fall at the end. These specifications yielded syllables of the type [kha:] which, it was hoped, with suitable  $f_0$  contours would be heard as the five Thai words listed in the section on Isolated Monosyllabic Words. The five  $f_0$  contours of the 1962 study (p. 127, Fig. 3.6) were retained with slight adaptations required by the nature of the computer program and imposed one by one upon this syllable. The contours, shown in the upper part of Figure 1, covered a range from 92 Hz to 152 Hz, which was reasonable for an adult male voice. Three tokens of each stimulus type thus produced were randomized into six test orders and played to 38 native speakers of Central Thai over a period of a month, together with other tests in the same sessions.

The results of these listening tests are shown in Table 2. Note that the tonal names in the first column are written with quotation marks. This is meant to convey that these  $f_0$  contours were intended as those tones but can be so labelled only to the extent that the subjects accept them as such. In the same spirit, the word *correct* at the bottom of the table is also printed with quotation marks. Otherwise, the form of the confusion matrix is the same as that of Table 1.

### Experiment 3: $f_0$ plus amplitude

Changes in the contraction of certain laryngeal muscles<sup>9</sup> and in subglottal air pressure can separately or together produce variations in the fundamental frequency of the voice. These mechanisms are also available for controlling the intensity of phonation and thus variations in the overall amplitude of the speech signal. To a certain extent, then, the two acoustic features,  $f_0$  and amplitude, may co-vary. Since the major psychological correlate of amplitude or intensity is loudness, just as that of  $f_0$  is pitch, it is not un-

reasonable to suppose that the ear may detect shifts in loudness in conjunction with large pitch excursions of one or more of the tones. If this is so, even though it has already been demonstrated with synthetic speech that certain ideal  $f_0$  contours carry sufficient information for the identification of the five tones, the perceptual processing of some of the tones in real speech may include awareness of relative amplitude as a concomitant feature which, under certain conditions, may actually contribute to tonal identification. In the present study the question arises as to whether the discrepancies in intelligibility between Experiments 1 and 2 will be removed simply by the addition of appropriate amplitude contours.

To answer this question, the amplitude contours of the lower part of Figure 1 were added to the corresponding  $f_0$  contours of the upper part. Amplitude is indicated in decibels (db) as decrements from the maximum output of the synthesizer at zero db. These amplitude contours are approximations derived from inspection of a small corpus of amplitude displays made with sound spectrograms of Thai words in isolation. The new synthetic stimuli were randomized five times each into three test orders and played to 40 native speakers of Central Thai over a period of a month, together with other tests. The results are shown in the confusion matrix of Table 3.

#### Comparison of the Three Conditions

Inspection of Tables 1-3 shows that the overall identifiability of the stimuli moves from 98.6% for real speech through 92.8% for fundamental frequency alone to 96.1% for  $f_0$  plus amplitude, thus

#### % Responses

Stimuli	Labels:	Mid	Low	Falling	High	Rising	N
	"Mid"	82.0	3.5	0.2	14.2		906
	"Low"	7.2	87.3	5.2	0.3		906
	"Falling"	0.4	0.7	97.8	0.6	0.6	906
	"High"	2.0		0.2	97.7	0.1	906
	"Rising"	0.2	0.2	0.3	0.1	99.1	906

Total = 4530      Subjects = 38      % "Correct" = 92.8

Table 2. Synthetic speech: responses to five  $f_0$  contours.



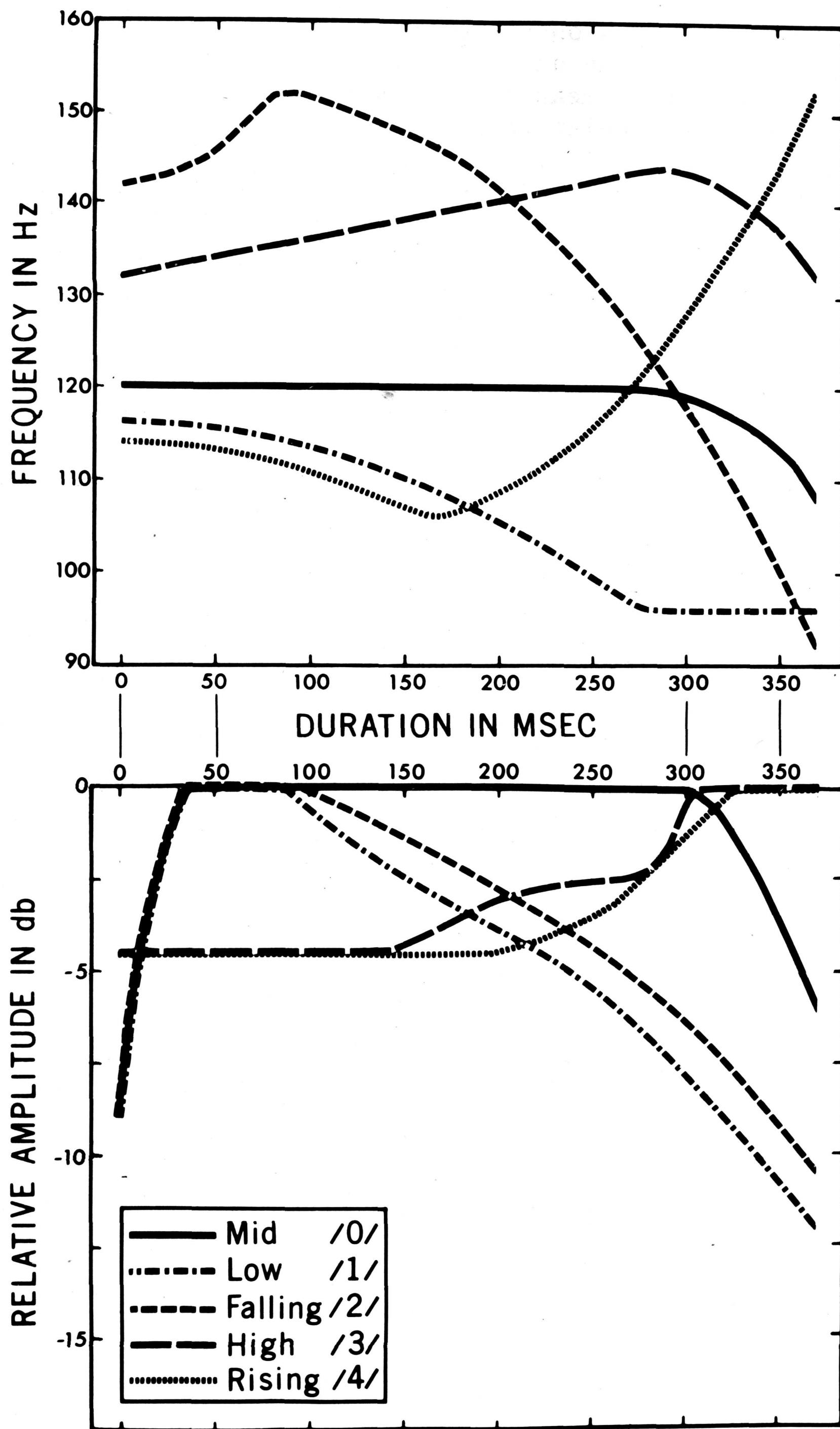


Fig. 1

Upper part:  $f_0$  contours used in Experiment 2.

Lower part: amplitude contours used in Experiment 3 in conjunction with the  $f_0$  contours.

suggesting that while  $f_0$  alone is by and large a sufficient cue, its efficacy is enhanced by the addition of amplitude information.<sup>10</sup> A comparison of the correct mean scores in the diagonals across the matrices is more to the point than just the overall scores. As we move from real speech (Table 1) to  $f_0$  alone (Table 2), the most striking changes are in the cells for the mid and low tones which lose, respectively, 15.9% and 9.3%. In Table 3 for  $f_0$  combined with amplitude information, the entries in the corresponding two cells move back in the direction of real speech, although the improvement is considerably greater for the mid tone.

Turning to the confusions in the matrices, we note a much greater scattering of errors in the two synthesis experiments as compared with real speech. Specifically, in Table 2 there is some confusion between the mid and high tones. The intended mid tone is called high 14.2% of the time. There is also some confusion in the other direction, high heard as mid, but only 2% of the time. Table 3 shows that the addition of amplitude information eliminates most of the confusion between these two tones. The only notable confusion in the real-speech test of Table 1 is between the mid and low tones. This is even worse in Table 2. Under both conditions hearing the intended low tone as mid accounts for most of the confusion. This latter effect is not eliminated by the addition of amplitude information in Experiment 3, although the identification of the intended mid tone itself is now improved to the level of real speech or slightly better. The intended low tone is heard as falling 5.2% of the time in Table 2 but is nearly back to the level of real speech in Table 3. One small but puzzling distortion apparently caused by the addition of amplitude information is the hearing of the intended falling tone as high 2.5% of the time in Table 3.<sup>11</sup>

It seems safe to infer from the results of the preceding experiments that  $f_0$  contours carry most of the information for tonal identification in Thai; that is, they carry sufficient information most of the time to identify words that are minimally distinguished by tone. A concomitant feature of some relevance for at least part of the tone system is contained in changes in the overall amplitude of the utterance. The confusions in the various matrices indicate the need for further information to improve the synthesis of Thai tones by rule. The improvements needed may be small refinements of the fundamental-frequency contours and better amplitude specification. In addition, simulating glottal tension in the voice source might make the high tone more natural and acceptable in utterance-final position.

#### Experiment 4: Perception of $f_0$ levels

The five tones of Central Thai can be viewed as falling into two groups, the dynamic tones and the static tones (Abramson 1962: 9-11). In this scheme, the sharp upward and downward movements of the rising and falling tones place them in the dynamic category.



% Responses

Stimuli	Labels:	Mid	Low	Falling	High	Rising	N
	"Mid"	95.7	1.0	0.2	3.0	0.1	1555
	"Low"	7.3	91.2	1.2	0.1	0.2	1555
	"Falling"	0.1	0.5	96.6	2.5	0.3	1555
	"High"	0.2	0.2	0.8	98.5	0.4	1555
	"Rising"	0.2	0.4	0.5	0.2	98.7	1555

Total = 7775

Subjects = 40

% "Correct" = 96.1

Table 3. Synthetic speech: responses to  $f_0$  + amplitude.

Since the high, mid and low tones often sound as if they simply occupy three levels, they are classified as static. Of course, the acoustical measurements, as reflected in the upper part of Figure 1, show that even the static tones undergo some  $f_0$  movement. Kenneth Pike (1948:5) speaks of level tonemes and gliding tonemes: "a LEVEL toneme is one in which, within the limits of perception, the pitch of the syllable does not rise or fall during its production. A GLIDING toneme is one in which during the pronunciation of the syllable in which it occurs there is a perceptible rise or fall, or some combination of rise and fall, such as rising-falling or falling-rising." It may be the case, then, that speakers of the language, not to mention field phoneticians, hear the static tones of Thai as simple levels. A lengthy quotation from Pike (1948:4) is of considerable interest here:

Tone languages have a major characteristic in common: it is the relative height of their tonemes, not their actual pitch, which is pertinent to their linguistic analysis. It is immaterial to know the number of vibrations per second of a certain syllable. The important feature is the relative height of a syllable in relation to preceding and following syllables. It is even immaterial, on this level of analysis (but not in the analysis of the linguistic expression of emotion), to know the height of a specific syllable in proportion to the general average pitch which the speaker uses. Rather, one must know the relationship of one specific syllable to the other syllables in the specific context in the particular utterance. A man and a woman may both use the same tonemes, even though they speak on different general levels of pitch. Either of them may retain the same tonemes while lowering or raising the voice in general, since it is the relative pitch of syllables within the immediate context that constitutes the essence of tonemic contrast.

Ilse Lehiste (1970:Chap. 3) provides a useful survey of these matters and related questions.<sup>12</sup>

How likely is it that absolute values of  $f_0$  levels provide sufficient cues for the recognition of the tones of a given speaker of a tone language in a certain context? The relativity of the pitches of tones (Cook 1972) is usually taken for granted. Indeed, the few studies that have yielded acoustical measurements of tones, e.g. for Mandarin Chinese (Howie 1972) and Thai (Abramson 1962, Erickson 1974), show that the tones are characterized by at least some movement and, in some cases, much movement. That is, they tend not to be uttered with a flat, unchanging  $f_0$ . The one tone in Thai that appears most likely to have a flat  $f_0$  in certain non-final environments is the mid tone. In an interesting experiment, Victor Zue (discussed in Klatt 1973) demonstrated that Howie's Mandarin contours still showed rather high intelligibility even when the range of absolute  $f_0$  is severely compressed, thus indicating that the pitch movement still available to the listeners carried sufficient information.

Of course, most of the foregoing observations are derived from tones in isolated words or at least in very short utterances. It seems likely that at least for some of the tones of Thai, presumably the high, mid and low, the perturbations of their  $f_0$  contours occasioned by the many coarticulations of running speech should produce flat variants here and there (Abramson, in preparation). If so, are such words understood by virtue of contextual redundancy or, to return to the question raised in the preceding paragraph, do the absolute levels furnish sufficient cues? That is, are "level" pitches assigned to tones only when they are the perceptual responses to small  $f_0$  movements, or will true acoustic levels suffice? The present experiment was designed to provide some answers to the question.



As shown in Figure 1, the "voice" of the synthesizer in the experiments reported so far was set to range from 152 Hz down to 92 Hz. In this experiment the range was divided to produce 16 flat fundamental frequencies in steps of 4 Hz; the amplitudes were flat too except for a slight rise at the beginning and a slight fall at the end. These variants were imposed upon the same basic syllable, randomized and played to 37 native speakers of Thai for identification as members of the same set of five words as before. Table 4 reveals that the falling and rising tones, which are characterized by very dynamic movements, elicited practically no responses. Indeed, the 0.1% response to the 100 Hz-level as the rising tone is so improbable as to suggest momentary inattention. From top to bottom, we have here a gradual crossover from the high tone through the mid tone to the low tone. Nowhere is 100% identification as a particular tone achieved. The closest is a peak of 90.1% for the low tone, which is comparable with the peak of 87.3% for the low tone in Experiment 2 (Table 2). The other two peaks here, 73% for the mid tone and 87.7% for the high tone, compare somewhat less well with their counterparts, 82% and 97.7% respectively, among the ideal contours of Experiment 2. Despite these peaks, it is important to note that all three tones persist in eliciting responses throughout their ranges. Most of this is accounted for not by the sporadic responses of the whole group of subjects but rather the deviant behavior of three of them. One of them agrees with the main group in calling the upper part of the range the high tone, but at 120 Hz she starts crossing over to the low tone and remains there the rest of the way down with only scattered responses in the mid-tone column. The second person of the three calls the upper part of the range the mid tone and crosses over to the low tone at 108 Hz. The third subject deviates in the most surprising way from the performance of the main group. She assigns the upper part of the range to the low tone and, crossing over at about 120 Hz, the rest of the range to the high tone!<sup>13</sup> Apparently, however, her psychological set shifts from time to time, because in some of the test sessions she assigned variations in the lower part of the range to the mid tone.

We may infer from the results of Experiment 4 that even in isolated monosyllabic words unchanging levels of fundamental frequency can carry considerable information as to the identity of the high, mid and low tones. It is to be supposed that in such a situation some accommodation to the speaker's pitch range is necessary. The subjects who took these tests were quite used to the "voice" of the synthesizer, and care was taken to confine the absolute levels of  $f_0$  to the range already in use in other tonal experiments. At the same time, the fact that there is no gliding movement at all in the stimuli seems to cause a certain amount of confusion across the three categories that the subjects accepted. Indeed, for three of the 37 subjects this factor is perceptually very disrupting. They may represent a population of Thai speakers for whom the small gliding movements normally found in productions of the static tones in isolation are essential for correct identification. In the absence of any  $f_0$  movement at all, it is not surprising that the falling and rising tones were not used as response categories.

% Responses

Stimuli

Hz	Mid	Low	Falling	High	Rising	N
152	8.0	4.1	0.2	87.7		903
148	7.9	4.2	0.3	87.6		903
144	8.6	4.1	0.2	87.0		903
140	12.1	4.3	0.1	83.5		903
136	18.6	5.4	0.1	75.9		903
132	29.2	5.4		65.3		903
128	49.3	6.2		44.5		903
124	65.3	5.5	0.1	29.0		903
120	72.6	6.3	0.1	20.9		903
116	73.0	12.2		14.7		903
112	66.4	19.7		13.8		903
108	42.1	45.7		12.2		903
104	18.4	73.8		7.9		903
100	11.0	81.5	0.3	7.1	0.1	903
96	5.5	88.7	0.1	5.6		903
92	4.8	90.1		5.1		903

Total = 14,448

Subjects = 37

Table 4. Synthetic speech: responses to 16 level  $f_0$ 's.



## Conclusion

The experiments described here lead to a few general conclusions about some aspects of the perception of the tones of Thai. First of all, it is more evident than heretofore that the intelligibility of tonally differentiated monosyllables presented in isolation is quite high. In addition, the average fundamental frequency contours obtained some years ago (Abramson 1961, 1962), when applied as instructions to the parameters of a different synthesizer and therefore presumably other synthesizers, still carry enough information for acceptable synthesis of Thai words. A slight reduction of the discrepancy in intelligibility between these contours and those of real speech can be effected by adding rough approximations of natural amplitude movements often found in correlation with shifts in fundamental frequency. To eliminate the small remainder of the discrepancy further work is needed. Finally, a continuum of level fundamental frequencies can be divided perceptually by native speakers of Thai into the high, mid and low tones but with very gradual transitions and with rather aberrant response patterns on the part of some test subjects; these data suggest that while levels, even in isolated syllables, can carry much information about these tones, there is still a fair amount of interference from the abnormal lack of change in fundamental frequency.

Given the paucity of similar perceptual data for other tone languages,<sup>14</sup> it is hard to say how we may generalize these findings beyond the Thai language. In fact, even within Thai there is little to be said in this regard about other major regional dialects.<sup>15</sup> Such phonetic features as "creaky voice" or "glottal tension" prominent in some tonal systems will probably require parameters other than simple control of fundamental frequency or amplitude for experimental investigation. Glottal tension is found in the high tone of Central Thai but appears to be unstable. Reports on more complicated manipulations of fundamental frequency to elucidate further the nature of tonal perception in Thai will be forthcoming.

## NOTES

- <sup>1</sup> This research was conducted 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. The hospitality of the Faculty of Humanities, Ramkhamhaeng University and the Central Institute of English Language, both in Bangkok, is acknowledged gratefully. The resources of the Haskins Laboratories, New Haven, Connecticut, U.S.A. were indispensable.
- <sup>2</sup> This is not to assert that other realms of prosody, such as sentence intonation and rhythm, are irrelevant to the enterprise of phonetic research. Indeed, some important psycholinguistic questions have been raised in work on intonation; see, e.g. Lieberman (1967) and Studdert-Kennedy and Hadding (1973).

03182

- <sup>3</sup> The topic is being studied by Dr. Udom Warotamasikkhadit.
- <sup>4</sup> Most of the tonal data for the 1962 publication were collected before 1960.
- <sup>5</sup> In /loolee/.
- <sup>6</sup> This kind of information can preserve the tonal distinctions to a severely limited extent in whispered speech in which no  $f_0$  is present (Abramson 1972).
- <sup>7</sup> Tandberg Teaching System IS 6 with Beyer DT98B Dynamic Headphones.
- <sup>8</sup> The capable and efficient selection and supervision of the test subjects by Miss Panit Chotibut of the Faculty of Humanities, Ramkhamhaeng University is much appreciated.
- <sup>9</sup> A University of Connecticut Ph.D. dissertation by Donna Erickson soon to be completed, Laryngeal Mechanisms and Co-articulation Effects in the Tones of Thai, explores the role of intrinsic and extrinsic laryngeal muscles in the production of the tones of Thai.
- <sup>10</sup> An experiment not performed as part of this research would be to try amplitude contours alone and then supplement them with  $f_0$  information. Previous research (Abramson 1972) implied that amplitude alone would be nowhere near sufficient for perception of the tones.
- <sup>11</sup> The effects set forth here seem obvious from simple inspection of the tables. Indeed statistical analysis (t-tests for differences between correlated means) performed with the kind help of Dr. Lyle Bachman of the Central Institute of English Language and the Ford Foundation, Bangkok shows them by the large to be significant at the 5% level of confidence or better. It is possible that more refined statistical procedures would lead to further observations, but the view taken here is that such effects would be so subtle as to be uninteresting for our understanding of the perception of tones.
- <sup>12</sup> An important background article for the phonological treatment of tone is Wang 1967.
- <sup>13</sup> This subject's correct use of labeling conventions in other tests taken during the same sessions shows that she is not guilty of misuse of labels here.
- <sup>14</sup> The most comparable study that comes to mind is Howie's work (1972) on Mandarin; for more detail see his Ph.D. dissertation (1970). See also Chan 1971.
- <sup>15</sup> Field work about to be started in Thailand by Jack Gandour of the University of California at Los Angeles should yield information along these lines.



## REFERENCES

- Abramson, A. S. 1961. "Identification and Discrimination of Phonemic Tones." (Abstract) *J. Acoust. Soc. Amer.* 33.
- \_\_\_\_\_. 1962. *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments.* Bloomington.
- \_\_\_\_\_. 1972. "Tonal Experiments with Whispered Thai" in *Papers in Linguistics and Phonetics to the Memory of Pierre Delattre.* Edited by A. Valdman. The Hague.
- \_\_\_\_\_. Forthcoming. "The Coarticulation of Tones: An Acoustic Study of Thai." *Proc. Eighth Intl. Cong. of Phonetic Sciences.* Leeds.
- Chan Yuen Yuen, Angela. 1971. *A Perceptual Study of Tones in Cantonese.* Ph.D. dissertation, University of London.
- Cook, E-D. 1972. "on the Relativity of Tones." *Lingua* 29.
- Erickson, D. 1974. "Fundamental Frequency Contours of the Tones of Standard Thai." *Pasaa* 4.
- \_\_\_\_\_. Forthcoming. *Laryngeal Mechanisms and Coarticulation Effects in the Tones of Thai.* Ph.D. dissertation, University of Connecticut.
- Fry, D. B. 1968. "Prosodic Phenomena" in *Manual of Phonetics.* Edited by B. Malmberg. Amsterdam.
- Howie, J. M. 1970. *The Vowels and Tones of Mandarin Chinese: Acoustical Measurements and Experiments.* Ph.D. dissertation, Indiana University.
- \_\_\_\_\_. 1972. "Some Experiments on the Perception of Mandarin Tones" in *Proc. Seventh Intl. Cong. Phonetic Sciences.* Edited by A. Rigault and R. Charbonneau. The Hague.
- Klatt, D. 1973. "Discrimination of Fundamental Frequency Contours in Synthetic Speech: Implications for Models of Pitch Perception." *J. Acoust. Soc. Amer.* 53.
- Lehiste, I. 1970. *Suprasegmentals.* Cambridge, Mass.
- Liberman, A. M. et al. 1954. "The Role of Consonant-Vowel Transitions in the Perception of the Stop and Nasal Consonants." *Psychol. Monogr.* 68, No. 8 (Whole No. 379).
- Lieberman, P. 1967. *Intonation, Perception, and Language.* Cambridge, Mass.

\_\_\_\_\_. 1974. "A Study of Prosodic Features" in Current Trends in Linguistics. Edited by T. Sebeok et al. 12. The Hague.

Lisker, L. and Abramson, A. S. 1964. "A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements." Word 20.

\_\_\_\_\_. 1967. "The Voicing Dimension: Some Experiments in Comparative Phonetics." Proc. Sixth Intl. Cong. of Phonetic Sciences. Prague.

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

Studdert-Kennedy, M. and Hadding, K. 1973. "Auditory and Linguistic Processes in the Perception of Intonation Contours." Language and Speech 16.

Wang, W. S-Y. 1967. "Phonological Features of Tone." IJAL 33.