STATIC AND DYNAMIC ACOUSTIC CUES IN DISTINCTIVE TONES*

ARTHUR S. ABRAMSON
The University of Connecticut
and
Haskins Laboratories

It is conventional to classify phonemic tones into dynamic or contour tones and static or level tones. The perceptual relevance of this impressionistic dichotomy is considered here for Central Thai, which has two dynamic tones (falling and rising pitches) and three static tones (high, mid, and low). A fundamental-frequency range appropriate to an adult male voice was used to synthesize three series of tonal variants on a syllable type available for five tonally differentiated words: (1) 16 F_0 levels at intervals of 4 Hz, (2) 16 F_0 movements from a mid origin to end points ranging from top to bottom of the range in steps of 4 Hz, and (3) 17 variants rising from the bottom to end points from top to bottom in steps of 4 Hz. The stimuli were played to native speakers for identification. The results indicate that level variants contain sufficient cues for identification as static tones but with considerable overlap. Identification, however, is enhanced by slow F_0 movement. Rapid F_0 movement is required for dynamic tones. Although imprecise, the typological dichotomy is useful.

BACKGROUND

In a tone language, part of the specification of each morpheme or word is a distinctive pitch pattern. Although some tones may have additional phonetic features,[1] the major characteristics of a tone system are found in voice pitch as determined by fundamental-frequency states and movements.

Some linguists refer to level tones, which are heard as having no pitch movement, and gliding tones which audibly rise or fall (Pike, 1948). In phonological analysis, the question may arise as to whether glides should be treated simply as whole pitch movements or as movements between level tones that are otherwise present in the system

* *This is a revised and somewhat lengthened version of a paper presented at the 91st meeting of the Acoustical Society of America, Washington, D.C., 4-9 April 1976. The data 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. The stimuli were prepared and the data analyzed at Haskins Laboratories under Grant HD-01994 from the National Institute of Child Health and Human Development.*

1 *E.g., creaky voice and shifts in loudness.*

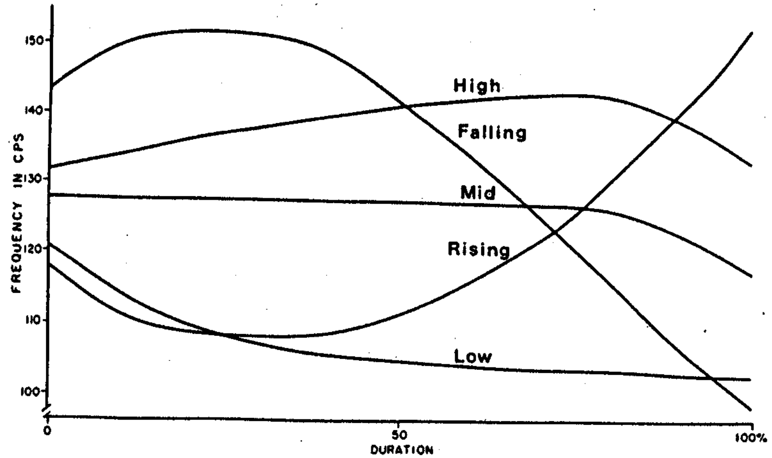

Fig. 1. Average F_0 contours of the tones of Thai on long vowels for a single male speaker (from Abramson, 1962, Fig. 3.6).

(Gandour, 1975, 1977). Here I am more interested in the validity or usefulness of the distinction between gliding or dynamic tones and level or static tones. The question is examined in Central Thai (Siamese), the official language of Thailand.

Some years ago I published typical fundamental-frequency contours of the five tones of Thai, as shown in Fig. 1 (Abramson, 1962). The tones are conventionally labeled high, falling, mid, rising, and low. Perceptual experiments with synthetic speech showed that these contours carried sufficient information for high intelligibility in the labeling of monosyllabic words. More recent experiments with the present formant synthesizer of Haskins Laboratories have again demonstrated the sufficiency of these contours (Abramson, 1975). Moreover, these findings provide a baseline for the experiments to be discussed here. Note that all the tones show at least some movement. The only one that may really be level is the mid tone because its final drop appears to be an intonational phenomenon before a pause.[2] In the experiments, I sought a basis for a division between dynamic and static tones in these curves. The falling and rising tones with their abrupt changes in frequency showed considerably more movement than the others. I labeled the falling and rising tone dynamic and the high, mid, and low tones, static.

In the past few years, further acoustic analysis of the Thai tones (Erickson, 1974, 1976; Abramson 1979) has suggested that, especially in running speech, the static tones are not very different from the dynamic tones. The *high tone* can be described as a high rising tone, while the *rising tone* can be described as a low rising tone. The low tone,

[2] *Speakers of Thai may find a prepausal mid tone without a final drop abnormal, but they identify such a contour nearly as well as the normal one (Abramson, 1975).*

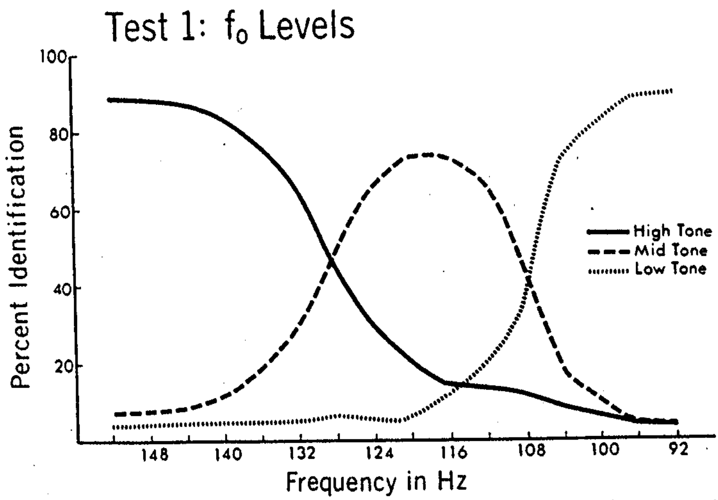

Fig. 2. Identification functions for fundamental-frequency levels as static tones. 37 subjects.

which could be viewed as a low falling tone, tends to fall to the bottom of the speaker's voice range and stay there, although this fall starts at a somewhat lower point than that of the *falling tone,* which has a high fall. It is only the *mid tone* that does not make extreme excursions into the high and low regions of the voice range, although it seldom has the ideal level shape of Fig. 1. The following three experiments are intended to shed light on the perceptual validity of the distinction between static and dynamic tones.

EXPERIMENTS

A syllable of the type [kha:] was prepared on the Haskins Laboratories formant synthesizer. Sixteen variants were made by superimposing 16 level fundamental-frequency trajectories ranging from 152 Hz down to 92 Hz in steps of 4 Hz. Each stimulus had a flat amplitude except for a slight rise at the beginning and a slight fall at the end. In Test 1, these were played in several randomizations to 37 native speakers of Thai for identification as one of five possible words.[3] The set of responses was presented to the subjects in Thai script, which provides a distinct spelling for each of the five tonally differentiated words. The question considered was the following: Do fundamental-frequency levels carry enough information for identification of the static tones, or must there be some movement for acceptability?

The results in Fig. 2 show that only the three static tones are used in Test 1 as response

[3] *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. The subjects were college students who were native speakers of the Central Thai dialect of Bangkok and its environs.*

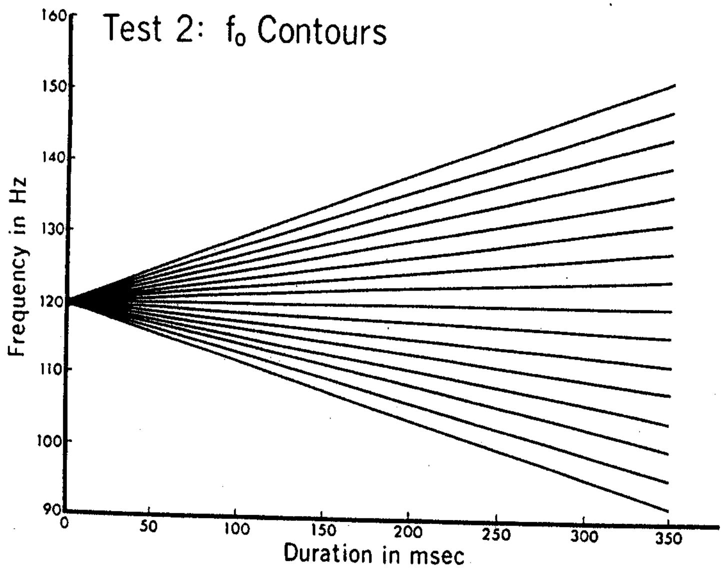

Fig. 3. Fundamental-frequency contours from a mid origin.

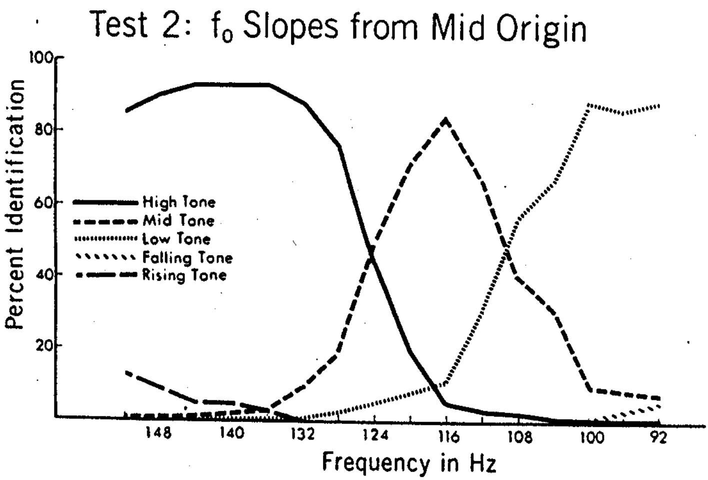

Fig. 4. Identification functions for the contours of Figure 3. 31 subjects.

categories. Note that nowhere is 100% identification reached. A peak of 90% for the low tone is about the same as the peak shown in the baseline test (Abramson, 1975) by the same subjects for the typical low tone displayed in Fig. 1. The high tone at the left reaches a peak of only 88% compared with 98% for the typical high tone in the baseline

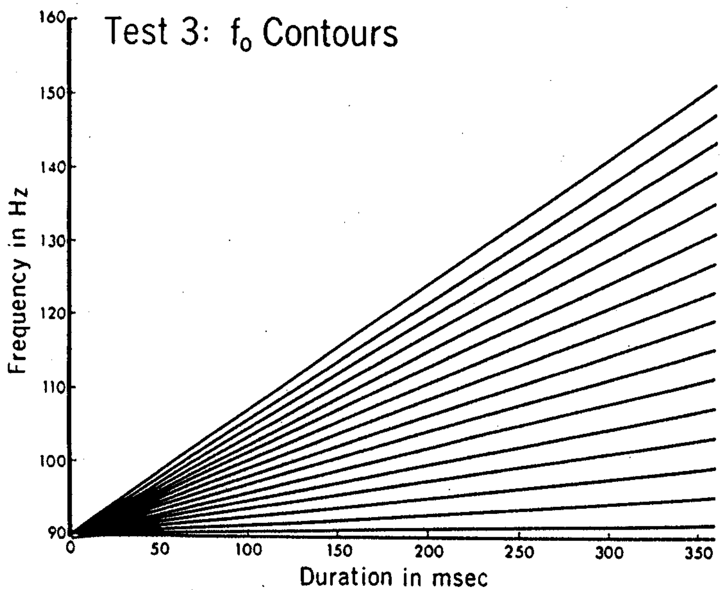

Fig. 5. Fundamental-frequency contours from a low origin.

test. The mid tone in the middle reaches 73% as compared with 82% in the baseline test. It is also true that all three tones elicit responses through the range. Most of the latter effect was caused by three subjects who used only two labeling categories, high and low or mid and low. Even in isolated monosyllables, then, flat fundamental-frequency paths can elicit static-tone responses. For this to happen in natural speech, there must be some auditory accommodation to the speaker's pitch range (Abramson, 1976) as well as to the immediate tonal context. At the time of this test, the subjects had served in other experiments and become used to the voice and frequency range of the synthesizer. Lack of F_o movement did cause some confusion for the subjects, and for three of them it was rather disrupting. It is not surprising that the dynamic tones were not used as response categories for these truly static fundamental frequencies.

In Fig. 3 we see the tonal variants used in Test 2 with the same syllable type. They all start from a common mid origin and end at the same points as in Test 1. I wondered whether the static-tone responses would be increased by the moderate amount of movement in most of these variants and, at the same time, whether at least the extreme values in the continuum would yield mainly dynamic responses. These stimuli were randomized and played to 31 of the original subjects, and the results are shown in Fig. 4. A few stimuli at either end do indeed yield dynamic responses, but no greater than a peak of almost 14% for the rising tone at the high end, and almost 5% for the falling tone at the low end. Otherwise, the static tones are again the predominant responses. Except for the low tone, there is somehwat better labeling here. The high tone goes from 88% in Test 1 to 94% in Test 2, and the mid tone improves from 73% to 84%. In fact, it is a slightly downward

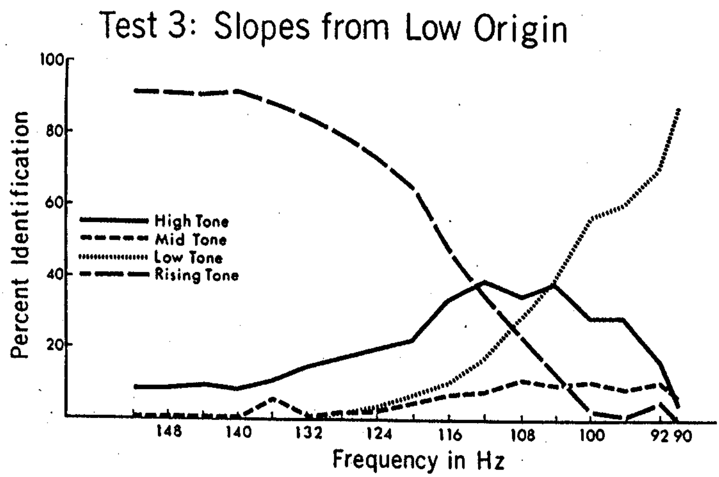

Fig. 6. Identification functions for the contours of Figure 5. 31 subjects.

movement from 120 to 116 Hz that yields 84%,[4] while the flat variant at 120 Hz yields only 72%![5] It seems safe to say that fundamental-frequency movements increase the acceptability of synthesized syllables as static tones. For the low tone, a more appropriate movement would reach a lower point in the voice range somewhat earlier (cf. Fig. 1).

In Fig. 5 we see the variants for Test 3. All the variants start from a low origin at 90 Hz and reach the same end points as before except for a flat variant ending at 90 Hz.[6] In the test these 17 stimuli were played to the 31 subjects for identification. It was expected that the sharply rising variants would be heard as a dynamic tone, namely the rising tone, with the others divided among the static tones with some preference for the low tone. The results are shown in Fig. 6. With a peak at 91%, the rising tone is clearly favoured. The low tone reaches a peak of 88% only at the very bottom of the range. It would be more convincing if it started higher and drifted downward to form a low fall (cf. Fig. 1). The third response category is the high tone which peaks at 38%. For this tone, a more appropriate movement would start higher. The mid tone which peaks at just under 12%, is negligible.

[4] *This should be compared with the 82% for the mid tone of the baseline test (Abramson, 1975). That stimulus did not slope downward from its onset as does the one described for Test 2 here, but it did have a final drop.*

[5] *Compare it with the flat variant at 120 Hz in Test 1 which yielded 73%.*

[6] *For a reason that is hard to reconstruct, possibly no more than an oversight, the low point was set at 90 Hz instead of 92 Hz as in Test 2. It is not likely that the downward shift of 2 Hz has any bearing on the outcome.*

CONCLUSION

We may conclude that fundamental-frequency levels do carry much information on the static tones, although they improve with movement. A rather abrupt movement is required for the dynamic tones; this is shown negatively by the weak falling-tone responses in Fig. 4 and positively by the strong rising-tone responses in Fig. 6. Other continua that bear on this question have been tested but are not yet ready for presentation. Although the dichotomy between static and dynamic tones is imprecise and unstable, more so in production (Abramson, 1979) than perception, it is still useful as a rough classification of tone production and as an index to the types of acoustic cues used in the recognition of tones.

REFERENCES

ABRAMSON, A.S. (1962). *The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments* (Bloomington, Ind.). [= *Int. J. Amer. Ling.*, **18**, No. 2, Part II].

ABRAMSON, A.S. (1975). The tones of Central Thai: some perceptual experiments. In J.G. Harris and J.R. Chamberlain (eds.), *Studies in Tai Linguistics in Honor of William J. Gedney* (Bangkok), pp. 1-16.

ABRAMSON, A.S. (1976). Thai tones as a reference system. In T.W. Gething, J.G. Harris and P. Kullavanijaya (eds.), *Tai Linguistics in Honor of Fang-Kuei Li* (Bangkok), pp. 1-12.

ABRAMSON, A.S. (1979). The coarticulation of tones: an acoustic study of Thai. In V. Panupong, P. Kullavanijaya, K. Tingsabadh and T. Luangthongkum (eds.), *Studies in Tai and Mon-Khmer Phonetics in Honour of Eugenie J.A. Henderson* (Bangkok, in press).

ERICKSON, D. (1974). Fundamental frequency contours of the tones of Standard Thai. *Pasaa*, **4**, 1-25.

ERICKSON, D. (1976). *A Physiological Analysis of the Tones of Thai*. Ph.D. dissertation, University of Connecticut.

GANDOUR, J. (1975). On the representation of tone in Siamese. In J.G. Harris and J.R. Chamberlain (eds.), *Studies in Tai Linguistics in Honor of William J. Gedney* (Bangkok), pp. 170-95.

GANDOUR, J. (1977). On the interaction between tone and vowel length: evidence from Thai dialects. *Phonetica*, **34**, 54-65.

PIKE, K. (1948). *Tone Languages* (Ann Arbor).