Static and Dynamic Acoustic Cues in Distinctive Tones*

Arthur S. Abramson+

ABSTRACT

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) sixteen F0 levels at intervals of 4 Hz, (2) sixteen F0 movements from a mid origin to end points ranging from top to bottom of the range in steps of 4 Hz, and (3) seventeen 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 F0 movement. Rapid F0 movement is required for dynamic tones. Although imprecise, the typological dichotomy is useful.

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, the major characteristics of a tone system are 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

*This is a slightly revised version of a paper presented at the 91st meeting of the Acoustical Society of America, Washington, D.C., 4-9 April 1976.

+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.

1Example: creaky voice.

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 (Gandour, 1975). 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 Thai, the official language of Thailand.

Some years ago I published typical fundamental-frequency contours of the five tones of Thai, as shown in Figure 1 (Abramson, 1962). The tones are con-
more movement than the others. I labeled the falling and rising tones 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, 1975b) 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 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 of the falling tone. 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 Figure 1. The following three experiments are intended to shed light on the perceptual validity of the distinction between static and dynamic tones.

A syllable of the type [kha:] was prepared on the Haskins Laboratories formant synthesizer. Sixteen variants were made by superimposing sixteen 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 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 Figure 2 show that only the three static tones are used as response categories. Note that nowhere is 100 percent identification reached. A peak of 90 percent for the low tone is about the same as the peak shown in the baseline test (Abramson, 1975a) by the same subjects for the typical low tone displayed in Figure 1. The high tone at the left reaches a peak of only 88 percent compared with 98 percent for the typical high tone in the baseline test. The mid tone in the middle reaches 73 percent as compared with 82 percent in the baseline test.

It is also true that all three tones elicit responses throughout 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 trajectories can elicit static-tone responses. For this to happen in natural speech, there must be some auditory accommodation to the speaker's pitch range as well as to the immediate tonal context. At the time of this test, the subjects had become used to the voice and frequency range of the synthesizer. Lack of F\(_0\) 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.

In Figure 3 we see the tonal variants used in Test 2. 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

\(^3\) 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.
Test 1: $F_0$ Levels

![Graph showing percent identification vs frequency in Hz for High Tone, Mid Tone, and Low Tone.]

Figure 2: Identification functions for fundamental-frequency levels as static tones.

were played to 31 of the original subjects, and the results are shown in Figure 4. A few stimuli at either end do indeed yield dynamic responses, but no greater than a peak of almost 14 percent for the rising tone at the high end, and almost 5 percent for the falling tone at the low end. Otherwise, the static tones are again the predominant responses. Except for the low tone, there is somewhat better labeling here. The high tone goes from 88 percent in Test 1 to 94 percent in Test 2, and the mid tone improves from 73 percent to 84 percent. In fact, it is a slightly downward movement from 120 to 116 Hz that yields 84 percent, while the flat variant at 120 Hz yields only 72 percent.\footnote{This should be compared with the 82 percent for the mid tone of the baseline test (Abramson, 1975a). 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.} 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 start somewhat lower in the voice range.

In Figure 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

\footnote{Compare it with the flat variant at 120 Hz in Test 1 which yielded 73 percent.}
Figure 3: Fundamental-frequency contours from a mid origin.

Figure 4: Identification functions for the contours of Figure 3.
variant ending at 90 Hz. 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 Figure 6. With a peak at 91 percent, the rising tone is clearly favored. The low tone reaches a peak of 88 percent only at the very bottom of the range. It would be more convincing if it started higher and drifted downward. The third response category is the high tone which peaks at 38 percent. For this tone, a more appropriate movement would start higher. The mid tone which peaks at just under 12 percent, is negligible.

We may conclude that fundamental-frequency levels do carry much information on the static tones, although they improve with movement. For the dynamic tones, as exemplified here by the rising tone, a rather abrupt movement is required. Other continua that bear on this question have been tested but are not yet ready

---

Figure 5: Fundamental-frequency contours from a low origin.

---

6For 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.
Test 3: Slopes from Low Origin

Figure 6: Identification functions for the contours of Figure 5.

for presentation. Although the dichotomy between static and dynamic tones is imprecise and unstable, more so in production (Abramson, 1975b) 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 recognition of tones.

REFERENCES


