A Seven-Tone Dialect in Southern Thai with Super-High: Pakphanang Tonal Acoustics and Physiological Inferences

Phil Rose

Australian National University

INTRODUCTION

Some varieties of Southern Thai are notable for the large number of surface tonal contrasts they exhibit. As one progresses south in this 500 kilometer long dialect continuum, the number of tones increases from five to six to seven (Diller, 1979b, p. 119). The latter varieties are thus members of the small number—15% — of the world's tone languages with more than six tones (Hombert, 1977, p. 21). This paper describes part of the tonal acoustics in one of these seven-tone varieties: Pakphanang (PPhN). This variety is tonally interesting not only for its large number of surface contrasts, but also for its super-high tone. It is an honour to offer this contribution to Dr. Vichin Panupong, and to be part, however modestly, of the tradition of Southern Thai phonetic studies she founded.

TONOLOGICAL DESCRIPTION

As in most varieties of Tai, the number of surface tonal contrasts on citation monosyllables in PPhN depends on the structure of the Rhyme, specifically the absence or presence of a syllable-final stop [p t k] in the Coda. Like the neighbouring dialect of Nakhon Sithammarat (Haas, 1958), PPhN contrasts seven tones on non-stopped syllables with phonologically at least two sonorants in the Rhyme. Examples of minimal and subminimal contrasts, from the field notes of A. Diller, are given in Table 1 in a conventional transcription for Tai varieties, with the long vowels [aː] and [uː] represented as a sequence.

As in Nakhon dialect (Diller, 1987, p. 148), PPhN tones 1, 2, 3, and 4 are in complementary distribution with respect to syllable-initial obstruents; tones 1 and 2 co-occur with voiceless aspirated stops and fricatives, and tones 3 and 4 with voiceless unaspirated stops. But as the examples in Table 1 show, there are minimal and sub-minimal pairs on syllables with the initial palatal glide /y/ to demonstrate a tonemic contrast between tones 1 and 3, and 2 and 4.

The auditory characteristics of the seven unstopped Pakphanang citation tones are as follows. These are based on my transcriptions of recorded utterances of one male and one female speaker. *Tone 1* is the super-high tone. It has convex pitch very high in the speaker's pitch range, with the falling portion a little more salient than the rising. The voice quality is often falsetto, which is possibly a consequence of the very high pitch target. This tone is auditorily very salient in running speech, and its extreme pitch characteristics are used by speakers of other, more distant Southern dialects to stereotype PPhN (A. Diller, p.c.). *Tone 2* has level pitch in the upper mid pitch range. *Tone 3* has convex pitch in the lower half of the pitch range, with the rising portion longer and more salient. In a second recording of the same speaker several years later, this tone had only a low rising pitch, which suggests that the falling part is optional. *Tone 4* has level pitch in mid pitch range. *Tone 5* has a relatively short level component in mid pitch range followed by a fall. *Tone 6* has a fairly low dipping pitch rising just into the mid pitch range. *Tone 7* has falling pitch in the low pitch range, and sounds longer than the other tones.

Table 1. *Examples of Tonal Contrast in Pakphanang Unstopped Syllables. (Bold face indicates forms acoustically analysed in this paper; the second row shows historical tonal category and consonant class).*

Tone	1	2	3	4	5	6	7
	*A H *B H	*C H	*A M *B M	*C M	*A L	*B L	*C L
	khaa *leg*	**khaa** *kill*	**kaa** *crow*	**kaa** *mark*	**khaa** *thatch grass*	**khaa** *value*	**khaa** *trade*
	khuu *threaten*	**phuu** *male*	**kuu** *I*	**kuu** *borrow money*	**khuu** *ditch*	**phuu** *wasp*	**ruu** *know*
	naa *thick*	naa *face*	paa *jungle*	paa *parent's older sister*	naa *rice field*	naa *worth doing*	naa *mother's younger sibling*
	yaa *fish sp.*	yaa *grass*	yaa *medicine*	yaaŋ *roast*	yaa *don't*	yaa *father's mother*	

The above description suggests that the PPhN pitch range is neither optimally nor neatly exploited. Thus the convex pitches of tones 1 and 3 lie more or less at extremes of the pitch range, but the level pitches of tones 2 and 4 are concentrated in the mid pitch range, and the two falling pitches of tones 5 and 7 lie in the lower half of the pitch range. A reasonable characterisation of the PPhN tonal system described above would be that it comprises four distinctive pitch contours—convex, level, falling, and concave—the first three of which each have a contrasting higher and lower value. Phonetic effects other than pitch, e.g., falsetto voice with tone 1, longer duration with tone 7, reinforce the pitch contrasts.

Hombert (1977, p. 29, f.n. 9) notes that languages with more than six tones usually recruit an additional distinctive phonetic parameter like phonation type into their tone system. Typologically, then, PPhN appears unusual. However, genetically it is not atypical, since the *distinctive* use of such parameters is not common in Tai. One documented occurrence is in the N.W. Tai dialect of Thai-Phake, which has tonally contrastive creak (see Rose (1990) for acoustic and aerodynamic data on this contrast).

PROCEDURE

In this paper, the acoustic properties of the seven unstopped tones will be described. Apart from fundamental frequency (F0), which is the major acoustical dimension in which tonal contrasts are cued perceptually, radiated amplitude (Ar), and Formant-pattern (F-pattern) were also examined for correlation with tone.

Due to well-known assumed intrinsic effects from concomitant segmental articulation, tones can never be acoustically observed independent of their segmental realisation. Because these effects are differential, it is not possible to infer a representative tonal shape from segmentally invariant data: to look at the F0 shape of a particular tone on a [u] vowel, for example, and assume this is also the shape representative of the tone. A corpus was therefore chosen using two vowels—[a] and [u]—which might be expected to elicit maximum difference in intrinsic effect of vowel quality on F0, Ar, and duration (D). It was further assumed that a better idea of the PPhN tonal F0 could be achieved by calculation of the mean value of these two extremes. This assumption needs to be tested, of course; one cannot be sure that the intrinsic effect of an [a] represents the same degree of divergence from the underlying, presumably invariant tone command as that of [u], or if perhaps the [a] interferes less with the realisation of the tone than the [u]. The corpus thus comprised the seven PPhN tones demonstrated on unstopped syllables with [a:] and [u:] vowels. The words used are given in bold face in Table 1. Note that the corpus reflects the above-mentioned complementary distribution of tone with respect to aspiration on obstruents. The words were elicited using Standard Thai orthography. Diller (1987, p. 15) notes that Thai orthography is capable of indicating Southern segmentals directly; that it represents either the Southern or Central tonal system with about equivalent degrees of fit; and that a text written in Thai can be given a seven-tone Southern Thai reading quite effortlessly. The words were recorded under laboratory conditions by a 46 year-old female native speaker in 1981. The speaker paused after each item. The [a:] forms were read first. Three repeats of the corpus were recorded (the informant did not know how many repeats were going to be requested). The aspiration on the obstruents was noticeably weak (this was also noted with another speaker's utterances); the unaspirated /k/ in tone 4 was noticeably fortis (but not ejected); and all examples of tone four were observed to register a much higher amplitude peak on the recorder's Vu meter. Since the informant was also typically fluent in Standard Thai, she was also recorded speaking corresponding Standard Thai forms in order to investigate how the Standard tones relate to those of the Southern varieties. Some initial results of this comparison were reported in Rose (1985).

The procedure used to extract F0, Ar, and D was essentially that described in Rose (1982, pp. 7–10). The duration of the Rhyme was first determined from wide band spectrograms with their good time-domain resolution. F0 was then sampled at percentage points of this duration from aligned expanded narrow-band spectrograms. Slightly different duration bases were used for tones 3 and 6; these are described below. This method of F0 measurement has an accuracy of +/- 4Hz at the 90% confidence level. Ar was sampled from the oscillographic trace from an F/J intensity meter, using 20 ms integration time and flat response. Alignment with the F0 curve was achieved by an audio wave on the oscillogram. This method is accurate to +/- 0.5 dB at the 90% confidence level. The high sampling frequency (mean value ca. 17 Hz, or once every 6 csec.) was chosen to permit detailed resolution of both F0 and Ar contours. \Measurements of possible F-pattern correlations of the tones were carried out as follows. Tokens were processed digitally, using the API and SGM commands of the ILS signal processing package (API did not reliably resolve two clear poles in the expected frequency range for F1 and F2 in the [u:] tokens, so these were not further analysed). The first three formants of the [a:] tokens were identified and sampled at three points each at approximately 25%, 50%, and 75% of the duration of the Rhyme. This gave nine tokens per formant per tone, over which means and standard deviations were then calculated.

Arithmetical mean and standard deviation values for F0, Ar, and D for the seven tones with [a:] and [u:] vowels, as well as VOT of initial obstruents and F-pattern for [a] vowels, are given in Table 2. The percentage points at which F0 and Ar were sampled are also shown under "SP." So, for example, in tone 2 words with [a:] vowels, the mean F0 at the 50% sampling point was 199 Hz, with a standard deviation of 8 Hz. The 50 % sampling point occurred at (56.8 csec * 0.5=) csec 28.4. For the low convex tone 3, visual inspection of the between-token variation in F0, and the above-mentioned free variation with low rising pitch, suggested that it would be better to treat F0 and Ar as splines, and sample them as functions of duration to F0 peak, and duration from F0 peak to phonation offset. Thus the sampling point at 60% of duration to F0 peak in tone 3 words with [u:] vowels occurs at (31.8 csec.* 0.6 =) csec. 19.1, and the 60% of duration to F0 offset at ((31.8 csec + (17.1 csec * 0.6))=) csec 42.1. In the low dipping tone 6, F0 and Ar were sampled as functions of duration to F0 peak, and also at phonation offset ("F0off").

INTRINSIC EFFECTS

In all except the convex tones 1 and 3 the [u:] allotype has, as expected, a higher F0. In the second half of the convex tones however (Figure 1), the intrinsic relationship is either reversed (tone 1) or cancelled (tone 3). It is easy to see that, because of the expected intrinsic differences in duration between [a:] and [u:] allotypes (the duration of the [a:] allotone is greater than the [u:] for all tones except 5, and some of the differences are quite big (e.g., 7 and 5 csec for tones 1 and 2) the nature of the intrinsic relationship on tones with appreciable F0 movement depends crucially on what F0 points are considered comparable between the two allotypes. If F0 values are compared as functions of absolute duration, then in the convex tones the intrinsic relationships are as just described. However, if F0 values are compared as functions of equalised duration (as will be done below), then for both convex tones the [u:] allotype has an intrinsically higher F0 for most of its duration. Moreover, for all tones the magnitude of an intrinsic difference thus established appears to correlate directly with its position in the speaker's F0 range. This relationship was quantified by linearly regressing intrinsic F0 differences at a given sampling point on their mean at that point.