THAI-PHAKE TONES: ACOUSTIC, AERODYNAMIC AND PERCEPTUAL DATA ON A TAI DIALECT WITH CONTRASTIVE CREAK

Phil Rose

Department of Linguistics (Faculties) A.N.U.

ABSTRACT: Mean fundamental frequency, amplitude, duration and air flow data are presented for the 5 tonemes of Thai Phake on syllables with [k] and [x] initial consonants and /aa/, /aat/ and /at/ rimes.

INTRODUCTION

Thai-Phake (TP), one of the extreme north-western members of the Tai family, is a rather inaccesible dialect spoken by fewer than 5000 speakers in the villages of Dibrugarh district in the Indian state of Assam (Diller, forthcoming). This paper presents instrumental data on the tones of one female speaker from Namphake Gaon.

As in most varieties of Tai, the number of surface tonal contrasts on citation monosyllables in TP depends on the structure of the Rime - specifically, the absence or presence of a syllable-final stop in the Coda. In unstopped syllables, TP contrasts 6 tones, while in stopped syllables only a two-way contrast obtains. The tones on stopped syllables can be analysed as versions of two of the phonological tones (i.e. either tonemes or autosegmental tones) which are realised on unstopped syllables. Five of the six tones on unstopped syllables realise a simple, unremarkable set of contrasts, with pairs of high and low level and falling pitches, and one rising pitch. However, perhaps the most interesting way in which TP differs phonologically from other Tai varieties is in the contrastive use of a creaky phonation type to distinguish the sixth tone. The auditory characteristics of the 6 tones on unstopped syllables as spoken by the informant used in this paper are as follows (historical tone category is indicated by *). **Tone 1** (*HA): slightly dipping pitch in the mid pitch range, followed by rise into high pitch range, e.g. [xa: 335] 'leg'. **Tone 2** (*HB, MB): level pitch in upper pitch range , e.g. [xa: 44] 'trellis', [ka: 44] 'to go'. **Tone 3** (*LB): level pitch in mid pitch range, e.g. [xa: 33] 'stair', [ka: 33] 'price'. **Tone 4** (*MA, LA): level component in high pitch range followed by a fall into the low pitch range, e.g. [xa: 552] 'thatch grass', [ka: 552] 'crow'. **Tone 5** (*LC): short initial level component in mid pitch range followed by a fall well into low pitch range, e.g. [xa: 331] 'bold', [ka: 331] 'to trade'. **Tone 6** (*HC MC): falling pitch in low pitch range. Phonation becomes creaky soon after onset of Rime, and ends in an audibly released glottal-stop, e.g. [xa̰:ʔ 21] 'to kill', [ka̰:ʔ 21] 'to dance'. In all tones except 6, the offset to phonation is gradual and breathy.

The two-way contrast on stopped syllables is realised in both pitch height and contour. One tone has a level, relatively high pitch; the other a low falling pitch. On the former (in syllables with [a] where a vocalic length contrast is possible) pitch height is inversely proportional to vowel length, as well as, of course, directly proportional to pitch length. Thus, e.g. (*HD, MD): [xat 4̲] 'to rub', [xa:t̚44] 'to lack'; [kat 4̲] 'to bite', [ka:t̚44] 'market'. No difference in pitch correlates with vowel length in the low falling stopped tone; however, long vowels tend to have creaky phonation at their end, e.g. (*LD): [xat 2̲1̲] 'to extract', [xa̰:t 21] 'sp. tree'; [kat 2̲1̲] 'cold'. On the basis of similarity in pitch features, the stopped tones can be regarded as realisations of (high level) tone 2 and (low falling) tone 5 respectively.

PROCEDURE

The corpus consisted of the 18 TP monosyllabic forms exemplified above. A numbered list of the english glosses was prepared, and the TP word for each item was then read out by a female native speaker in response to a number uttered by a promptor. A pause of about 4 sec. separated the utterance of each word. Two types of instrumental record - one primarily for acoustic and one for aerodynamic measurement - were made on professional equipment in the phonetics laboratory of the A.N.U. Faculties' Linguistics Department. The first type, called "set A" below, provided detailed data on fundamental frequency (F0), radiated amplitude (Ar), and duration. For this the 18 item set was recorded 7 times on analog tape, with the 3rd and 5th times recorded in reverse sequence. Conventional oscillograms (125 mm/sec.) were made of the audio wave, and F0 and flat Ar as extracted by F/J fundamental frequency and intensity meters (the latter with 20 ms. integration time). These oscillograms were used to determine, by visual inspection, appropriate sampling bases for F0 and Ar. Ar was measured from them, as were also F0 perturbations at offset. F0 values at other points, however, were calculated directly from the period of the wave form in a set of high speed oscillograms (1cm: 5 msec). For the second type, called "set B" below, a simultaneous oscillographic record was made at 125 mm/sec of the airflow (Af), full wave audio, and F0 of 4 repeats of the 18 item list. Af was registered with an F/J electroaerometer. The low-pass filter was switched off, so that the rapid AC changes in Af during phonation could be resolved. The audio signal, from which F0 was extracted using an F/J fundamental frequency meter, was transduced by a microphone in the aerometer face mask. This resulted in a smoother F0 trace than the set A recordings, but tended not to resolve F0 below a certain amplitude, e.g. during the creaked tone, or at the end of syllables in general.

In the absence of a methodological precedent for sampling tonal Af, I adopted the following strategies. Firstly, because during phonation the mean Af and peak-to-peak amplitude of the Af (Af *ptp*) seemed at least partially independent, phonatory Af was quantified in terms of maxima and minima at each sampling point. It was then easy to derive both mean Af and Af *ptp* separately from these values. Secondly, in order to facilitate comparison between Af and F0, phonatory Af was sampled as a function of the sampling base of the concomitant F0, rather than of a separate base based on Af landmarks alone. Thirdly, as detailed a profile as possible of the mean time course of the Af throughout the production of the syllable was obtained by sampling at points where the derivative changed sign, and by having a high sampling rate where derivatives were high.

RESULTS

In reading the 7 repeats of set A, the informant made 2 systematic errors: the low falling stopped xat 'to extract' was read with a high tone, i.e. as xat 'to rub', 5/7 times; and it also sounded as if the low level unstopped forms (xa: 'stair' & ka: 'price') were read as the corresponding low falling forms xa: 'bold' 2/7 & ka: 'to trade' 4/7 times. 2 tokens of the low falling tone also sounded as if read as low level. Possible errors in set B could only be detected by visual inspection of the oscillographic trace. This showed a repeat of only the first pattern of confusion in set A, in that all tokens of xat 'to extract', and one of kat 'cold' were read as 'to rub' & 'to bite' respectively.

Because some of the tokens of the low falling and low level tones in set A were auditorily similar, and because there was obviously some kind of confusion between them, it seemed advisable to take precautions to ensure that mean acoustic values were based on samples containing perceptually unambiguous tokens of these tones. I therefore determined with an identification test using tokens edited from set A which of the recorded tokens of the low level and falling tones could be correctly identified. Only these were used for measurement.

Out of what I considered 18 unambiguous examples of the low level tone, the informant identified 4 correctly, 11 as high level, 1 as low level, and 2 as 'unsure'. Of 30 unambiguous tokens of the low falling tone, 17 were correctly identified, 5 were heard as high fall, 4 as low level, and 4 as 'unsure'. I took this to indicate only a slight perceptual confusion between low fall and low level, but a rather greater perceptual similarity between high and low versions of the same contour.

Mean F0 and Ar for set A, and mean F0 and Af for set B are shown in fig.1. Solid and dotted lines indicate values for syllables with [x] & [k] initial consonants respectively. For each tone, set A data are on the left. F0 (Hz) is shown at the top, Ar (dB) at the bottom left. Af & Af *ptp*, quantified in arbitrary units (cms), are shown at the bottom right (Af *ptp* from d.c to 1). On the non-stopped level and falling tones in set A quite drastic F0 perturbations are visible in the 10 csec or so before phonation offset (comparable perturbations also occur on set B, but not all have been shown). The partial congruence of the airflow traces suggests that the F0 perturbations occur intrinsically as the result of aerodynamic factors associated with changing mean glottal area at phonation offset.

Fig. 1 shows that there are regular differences associated with the contrast between syllable-initial [k] and [x] in all parameters examined. For all tones at about 5 msec after phonation onset, F0 is a mean of 30 Hz higher after [x] than [k]. Higher F0 values after [x] persist for at least 15 csec (unstopped) and 10 csec (stopped) for all except creaked and low stopped tones. The two sets differ in the following

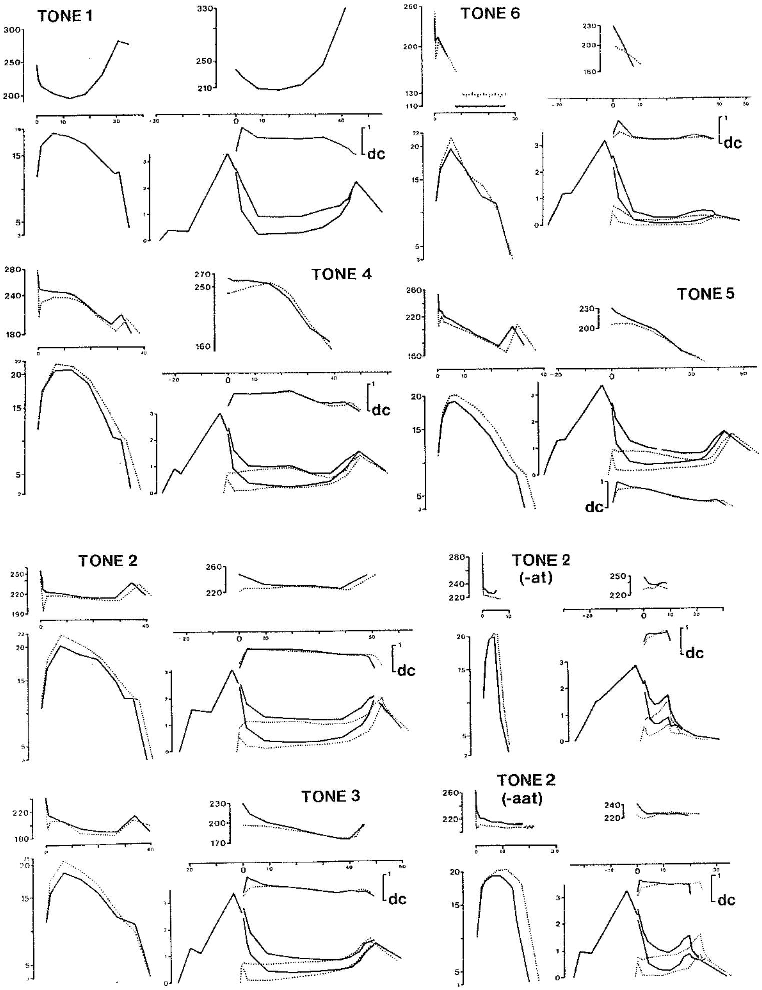
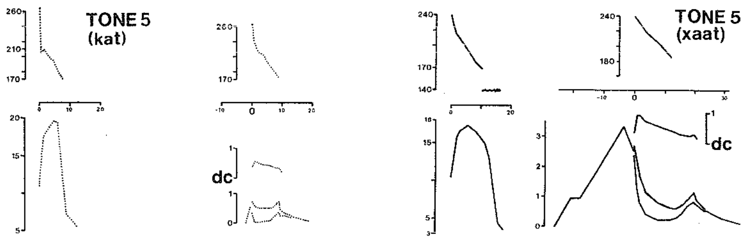


Figure 1. Mean fundamental frequency, radiatied amplitude, and airflow for Thai-Phake tones.

Figure 1 (cont.)



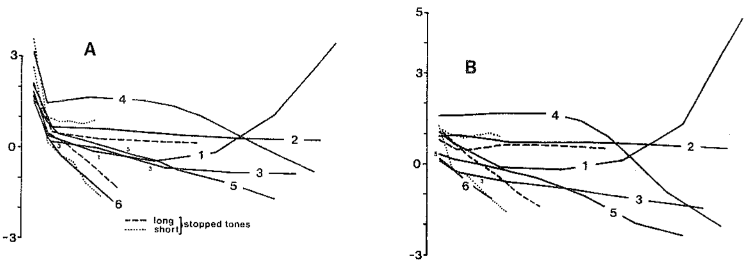


Figure 2. Normalised F0 shapes for the two sets of Thai-Phake tones

ways with respect to the F0 difference. *Persistence*: In set A, higher values after [x] last for the whole duration in all tones; in set B this only occurs in high stopped short tones. For unstopped, non-creaked tones, the F0 difference generally disappears between 15 & 20 csec after onset of phonation, and for long stopped tones at 9 csec. *Sensitivity to the final stop*: The F0 difference in set B is smaller on stopped tones. Mean values for F0 differences at onset, 2.5 and 5 csec are 27, 17 & 13 Hz for unstopped, but 19, 6 & 3 Hz for stopped tones. *Timing of minimisation*: In set A, where mean values for F0 differences at onset, 2.5, 5, 10 & 15 csec are 31, 9, 6, 5 & 4 Hz, F0 differences are minimised earlier than set B, with comparable values (for unstopped tones) of 27, 17, 13, 6 & 3 Hz. *Onset perturbation* : Set A shows no systematic difference between [k] & [x] syllables in the period of the first glottal pulse but the periods of subsequent pulses show large differences. After [k], the 2nd pulse shows a mean drop of 50 Hz, followed by a mean rise of 20 Hz in unstopped, and 6 Hz in stopped syllables. No such perturbation occurs in [x] syllables. In set B, the perturbation associated with [k] in set A is usually absent.

As the result of the intrinsically high Af associated with the initial consonant, over about the first 8 csec after phonation onset, tones with [x] have a high, rapidly falling mean Af, as opposed to the low, relatively stable Af after [k]. Small differences in Af *ptp* are also observable between the allotypes in this period. Whereas the Af *ptp* differences rapidly equalise, however, the mean airflow after [k] tends to remain slightly lower than that after [x] through the whole of the voiced part of the syllable. There is good correlation between the F0 and mean Af values within each tone over the first 10 csec after phonation onset, which suggests that during this period the F0 difference between [k] and [x] allotypes is an intrinsic function of the difference in mean Af. However, although the general shape of the function is similar for each tone, (possibly with F0 proportional to the natural log of the mean Af), they have differing intercepts, which indicates that the absolute F0 values are determined by different vocal cord tension settings. Syllables with [x] also have 1 - 1.5 dB lower Ar over most of the vowel; and 2 - 4 csec shorter vocalic duration than with [k].

The minimal contrast between the creaked and low falling tone is cued acoustically primarily in F0, although the Ar of the former shows a more abrupt decay after peak. Although the F0 onsets at about the same value in both tones, the creaked tone shows a more rapid rate of fall. The high speed oscillogram showed that in set A, creak starts soon - after 5 to 8 csec after phonation onset - with a subsequent mean period of between 110 & 130 Hz. Interestingly, although the set B F0 (largely unresolved by the F0 meter) is clearly lower than in the low falling tone, there is no evidence in the airflow of higher jitter values typical of creak until late - 5 to 10 csec before phonation offset. As might be expected, the airflow of the two tones is highly contrastive, with the creaked tone showing both considerably smaller Af *ptp* and lower Af values.

F0 is also clearly the most contrastive acoustic parameter as far as the other tones are concerned. Fig. 2 shows z-score normalised F0 values for the two sets as functions of duration normalised with respect to the longest tone (tone 1). The perturbations at F0 offset have been discarded as tonally irrelevant. Those at F0 onset in set A , also tonally irrelevant, are retained to show the high value of the first glottal pulse in all tones relative to the rest of the confguration. The F0 configuration is very similar for both sets, except that set B has relatively more extreme offsets to the dynamic tones. The low falling and low level tones have very similar F0 values over the first half of their duration, which may account for the slight perceptual confusion noted above. Typically for Tai, there is very little differentiation between the Ar shapes. They all show, irrespective of F0 contour, an abrupt rise to an early peak followed by a more gradual decay. This points to predominant involvement of vocal cord tension in the production of the F0 contours (cf. Rose 1984). Duration relationships are constant across the two sets, with small differences correlating with F0 height and contour. The syllable-final stop considerably truncates the duration of the F0: phonologically the same nucleus (/aa/) in tones 2 & 5 has about half the phonetic duration when followed by a stop; and a nucleus linked to a single vowel before a stop has between one third and one fifth the duration of the corresponding unstopped tone. Finally, apart from the creaked tone with its very low mean Af and, possibly, the upper level tone with high mean Af, there is no difference between the tones in either mean Af level or contour. However, the tones do clearly show different Af *ptp* profiles, which also correlate positively with their F0 contours, and to a certain extent with their overall F0 level.

ACKNOWLEDGEMENTS

Many thanks are due to Mrs. Yehan Buragohain for being an ideal, co-operative informant. Thanks also to Dr. Tony Diller for help with preparation of the corpus, elicitation of data, and administration of the perceptual test.

REFERENCES

Diller, Anthony. (forthcoming) "Tai Languages in Assam: Daughters or Ghosts?" to appear in a festschrift for Professor William Gedney.

Rose, Phil. (1984) :The role of sub-glottal pressure and vocal cord tension in the production of tones in a Chinese dialect, in B. Hong (ed.) *New Papers on Chinese Language Use,* (Australian National University: Canberra), 133-168.