AERODYNAMIC INVOLVEMENT IN INTRINSIC F0 PERTURBATIONS - EVIDENCE FROM THAI-PHAKE

Phil Rose

Department of Linguistics (Arts)
Australian National University

ABSTRACT - Mean fundamental frequency and airflow data are presented for 17 acoustic allotones of the six tonemes of Thai Phake on syllables with [k] and [x] initial consonants. It is argued that observed F0 perturbations at syllable onset are caused by aerodynamic factors associated with the difference in initial consonants. Historical tonological implications are briefly discussed.

INTRODUCTION

It has long been recognised that the manner of articulation of a syllable-initial consonant can cause intrinsic perturbations in the fundamental frequency (F0) of a following vowel (Hombert 1978). The most frequent type of assocaition cited is between voicing and F0 height, with F0 after voiced consonants being lower than after voiceless. Both aerodynamic and physiological reasons have been proposed for this correlation, but the role of airflow has been explicitly discounted, both for stops differing in voicing (Ohala 1978, 26-28), and aspiration (Hombert 1978, 89,90). This paper presents data from a tone language which show, contrary to the received view, intrinsic F0 differences which are likely to have been caused by differences in airflow associated with the manner of syllable-initial consonant.

The tone language used is Thai-Phake (TP), one of the extreme north-western members of the Tai family, spoken in Assam. A description of the mean acoustic characteristics of the tones -- F0, duration and amplitude -- together with data on airflow and tonology can be found in Rose (1990). TP contrasts six tones on unstopped syllables: two falling, two level, one dipping, and one low falling with creak and glottal-stop offset. Examples are: T(one) 1 [xaa 335] *leg*; T2 [xaa 44] *trellis*; T3 [xaa 33] *stair*; T4 [xaa 552] *thatch grass*; T5 [xaa 331] *bold*; T6 [xaaʔ 21] *to kill*. Two tones (T2, T5) have additional allotones conditioned by a syllable-final stop and vowel length: T2 [xat 4̲] *to rub*, [xaat 44] *to lack*; T5 [xat 2̲1̲] *to extract* ; [xaat 21] *sp. tree*. Of importance in this paper is the fact that in TP [k] and [x] are are allophones of /k/ and /kh/. TP initial stops contrast in aspiration, with grave aspirated phonemes /ph/ /kh/ having fricated allophones in free-variation with stops.

PROCEDURE

Four repeats of each of the allotones described above were obtained from one female speaker from Namphake Gaon. Each allotone was said on syllables with [k] and [x] syllable-initial consonants, with the exception of stopped tone 5, for which only the forms [xaat 21] and [kat 2̲1̲] were available. Elicitation and mensuration details are in Rose (1990, 394-395). Measurements were made from a simultaneous mingographic record of the airflow (Af), full wave audio, and F0. Af was registered with an F/J electroaerometer, with the low pass filter switched off, so that the rapid AC changes in Af during phonation could be resolved. The audio signal, from which F0 was extracted using a F/J pitchmeter, was transduced by the microphone in the aerometer face mask.

Measurements of the acoustic and aerodynamic parameters were made in the following way. For each tone, a sampling base was defined with respect to its F0 shape. The F0 and Af were then sampled at percentage points of the duration of this base with a high enough sampling frequency to resolve the recurrent details in their mean time course. The sampling base was defined to extend between the observed onset of the F0 shape and an easily identifiable offset point that appeared to be comparable between tokens. Af was measured as the point half way between maximum and minimum airflow, and quantified in arbitrary units (cms).

RESULTS

Table 1 gives the means and standard deviations for corresponding F0 and Af, and duration, in the 17 allotones. Af values are quoted in (100 x cm.) units. The percent sampling point is given down the left. The actual value of the sampling point can be calculated from the duration value at the bottom of the table.

At the 60% sampling point in the high level [44] tone 2 with initial [x], for example, the mean F0 was 229 Hz, with a standard deviation of 6 Hz, and the corresponding Af was 78 (sd= 17). The 60% sampling point occurred at (60% x 48.2=) csec. 28.8 from onset of phonation.

F0 PERTURBATIONS

In all tones, F0 at phonation onset is higher after [x] than after [k]. (This was also found in a second set of tones recorded in the same session without the airflow mask (Rose 1990:395)). Higher F0 values after [x] persist for at least 15 csec. (unstopped tones) and 10 csec. (stopped tones) in all except creaked and low stopped tones. During this time, the difference between the two sets diminishes, from a mean difference of 30 Hz at about 5 msec. after phonation onset, through one of 8 Hz after 5 csec. Typical F0 perturbations as a function of the [k]/[x] difference can be seen in figure 1, which shows mean F0 for unstopped tones 2 and 3 with [k] and [x] initial consonants.

AF PERTURBATIONS

The difference between the initial consonants [k] and [x] correlates with relatively large differences in Af contour and height up to ca.
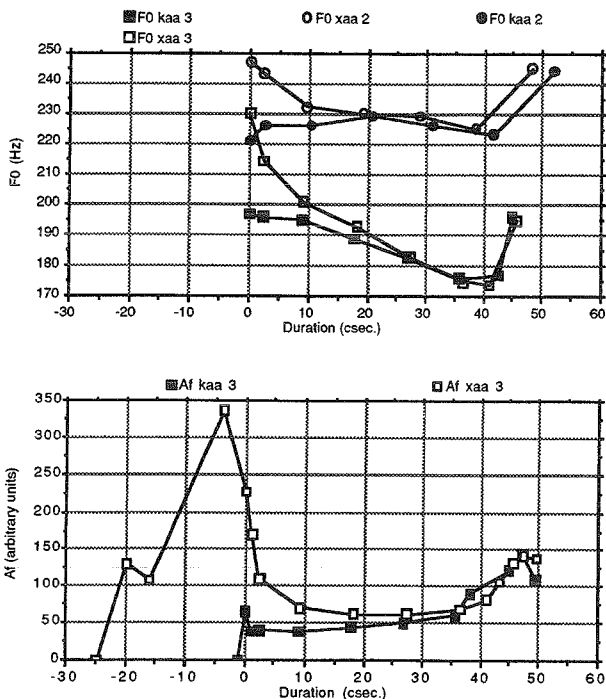
Figure 1. Mean F0 shapes for TP tones 2 and 3 after [k] and [x] (above), and corresponding Af shapes for tone 3 (below). All shapes are aligned at phonation onset (csec.0).

the first 15 csec. after phonation onset, and small differences in Af height thereafter. Typical Af profiles can be seen in figure 1, which shows mean Af for tone 3 with [k] and [x] inital consonants. Most [x] tokens present a for voiceless fricatives characteristically bimodal Af profile, with the second peak much higher than the first. This shape is interpretable in terms of the well-known effects of the twin resistance to flow at glottis and supraglottal constriction (Klatt et al.1968, 48). However, the much higher Af rate of the second peak is atypical for [x] or [X], and indicates that the articulation involves more than a voiceless velar/uvular fricative (pers. comm., Ian Maddieson). The details are as follows. Af is first registered about 25 csec. before phonation onset. It rises abruptly, but its rate soon slows down and after about 5 csec. may even change sign. Af values for some 4 to 5 csec. after this point probably reflect the supraglottal resistance at the dorso-velar/-uvular constriction. After about another 4- 5 csec., the rate again increases fairly sharply, and reaches a peak after 13 csec. Over the last 3 to 4 csec. before the Af becomes periodic, the flow rate falls because of increasing glottal resistance as the cords are adducted from their voiceless position to begin phonation.

The high flow rate achieved before cord adduction is noteworthy. The rapid increase in flow towards this peak reflects the decreasing oral resistance as the tongue body moves away from the velar/uvular

constriction towards the target position for the following [a] vowel. However, the very high rate of airflow attained at the second peak indicates that the tongue movement towards the target vowel is not accompanied by cord adduction as it would be in the case of the more normal, coincident VOT fricative. Rather the Af trace indicates that the cord adduction prior to phonation onset is delayed, resulting in a voiceless glottal fricative -- specifically, a voiceless [a] vowel -- with its consequent high Af rate.

Table 1. Means and standard deviations (X, sd) for F0 (Hz), duration (D) (csec.) and Airflow (Af) (arbitrary units) in Thai phake tones.
n = 4 unless otherwise shown.

	T2 [xaa 44]						T2 [kaa 44]						T1 [xaa 335]			
	F0		Af				F0		Af				F0		Af	
SP	X	sd	X	sd			X	sd	X	sd			X	sd	X	sd
0	247	4	253	20			221	6	69	28			238	17	264	51
5	243	6	135	21			226	7	67	28			228	7	165	56
20	232	3	85	18			226	4	62	17			209	5	56	19
40	230	5	79	17			229	6	68	15			207	4	56	20
60	229	6	78	17			226	6	62	18			215	3	60	25
80	225	4	84	13			223	5	69	18			244	14	81	34
90	-	-	107	22			-	-	-	-			-	-	-	-
95	-	-	127	29			-	-	103	9			-	-	110	27
100	245	10	161	39			244	7	145	10			330	32	136	24
D	48.2, 1.7						52.0, 2.6						41.3, 2.1			

	T3 [xaa 33]						T3 [kaa 33]			
	F0		Af				F0		Af	
SP	X	sd	X	sd			X	sd	X	sd
0	230	12	227	42			197	5	64	26
2.5	-	-	170	37			-	-	39	14
5	214	6	111	47			196	6	40	13
20	201	7	70	20			195	5	38	9
40	193	6	63	20			189	3	45	9
60	183	2	62	21			183	4	50	12
80	175	2	68	25			176	6	60	24
90	174	2	82	26			-	-	-	-
95	-	-	107	32			177	5	91	26
100	195	3	133	34			196	5	122	24
D	45.5, 1.3						44.8, 1.3			

	T4 [xaa 552]						T4 [kaa 552]			
	F0		Af				F0		Af	
SP	X	sd	X	sd			X	sd	X	sd
0	263	6	233	10			240	6	62	23
5	260	7	127	23			242	5	42	20
20	260	3	70	10			250	6	48	12
40	255	5	63	8			256	5	57	12
50	245	7	-	-			250	7	58	11
60	231	7	63	8			237	5	56	13
80	187	8	52	6			192	7	46	15
90	-	-	50	4			-	-	44	13
100	165	1	57	10			155	7	42	17
D	38.6, 2.7						39.0, 2.0			

In most [k] tokens, Af is registered for about 1 csec. before phonation begins. The abrupt rise in flow during this time reflects the rapid increase at release of the dorso-velar occlusion of the stop. The cords are presumably already so close together at release of supraglottal occlusion that the increase in glottal resistance upon adduction has negligable effect on the Af. For the first 2 - 3 csec. after phonation onset, Af after [k] is associated with a small but abrupt fall, after which the profile may rise, fall, or remain level, depending on the tone.

The difference in Af characteristics associated with the two consonants means that at the onset of phonation, the mean Af for "[x]" is about 4 times greater than for [k]. Thereafter this difference quickly diminishes, and there is some evidence that the difference is minimised earlier on tones with phonetically shorter nuclei. In the majority of cases, a difference in Af as a function of the initial consonant appears to obtain not just for the period during which the [k] and [x] trajectories converge, but to persevere throughout the syllable. (Only tone 3 and long stopped tone 2 show effectively the same Af for both allotypes towards the end of the tone.) Since it is reasonable to assume that other parameters which could effect Af -- subglottal pressure, supraglottal area function -- are equal for the same tone, the slightly higher Af values after [x] are presumably caused by perseverative difference in phonatory glottal resistance. Possibly the mean glottal area is slightly greater after [x], reflecting the original difference in interarytenoidal distance during the production of [x] and [k].

It will be recalled that the TP syllable-initial "[x]" is described above as being in free variation with [kh], both of which are allophones of /kh/. The duration from the release of the [x] target to the onset of phonation is about 17 csec. This would correspond fairly well to the duration of the hold phase plus VOT for an

aspirated velar stop in citation form. The Af indicates therefore that the relative timing of the glottal and supraglottal mechanisms for an aspirated velar stop is being preserved even when there is no supraglottal occulsion. On the basis of the instrumental phonetic evidence, the contrast between "[x]" and [k] appears to be still primarily one of aspiration *qua* VOT lag rather than frication.

Table 1 (con't)

SP	T5 [xaa 331] F0 X	sd	Af X	sd	T5 [kaa 331] F0 X	sd	Af X	sd
0	231	6	262	33	206	4	83	19
5	224	6	164	42	207	3	53	18
20	214	4	89	17	207	3	53	21
40	203	7	72	26	197	3	50	22
50	197	2	-	-	189	2	-	-
60	188	2	67	25	181	1	47	18
80	165	3	63	26	161	3	41	18
100	154	4	67	35	149	4	44	14
D	32.7, 1.0				34.9, 2.5			

SP	T6 [xaa? 21] F0 X	sd	Af X	sd	T6 [kaa? 21] F0 X	sd	Af X	sd
0	230	3	241	21	203	7	56	
5	217	11	149	27	197	4	35	
10	202	10	-	-	194	6	-	
20	182	11	36	14	183	7	13	5
30	167	9	-	-	169	8	-	-
40	161	3	20	7	160	10	11	4
50	154	1	-	-	142	9	-	-
60	146	11	21	7	131	13	13	5
65	-	-	22	12	-	-	-	-
70	-	-	-	-	113	13	-	
80	-	-	31	11	-	-	14	6
90	-	-	39	11	-	-	20	9
100	-	-	43	19	-	-	28	14
D	37.8, 2.9				37.0, 2.6			

SP	[xat 4] n=6 F0 X	sd	Af X	sd	[kat 4] F0 X	sd	Af X	sd
0	248	12	204	32	229	2	63	28
10	-	-	153	29	-	-	55	10
20	238	12	124	30	231	15	52	13
40	236	10	101	27	230	12	62	13
60	236	7	102	30	235	12	72	18
80	239	9	117	28	232	10	90	28
100	238	12	130	27	230	15	104	25
D	8.7, 1.2				8.5, 0.7			

period under examination.)

DISCUSSION

The following considerations make it likely that the observed F0 differences are intrinsically caused by the Af differences, which themselves result from the extrinsic glottal area settings necessary for the contrast in +/- [continuant] on the syllable initial consonant.

Firstly, an excellent corrrelation exists between F0 differences and corresponding Af differences at the same duration point ("difference" means the difference between a value at a given duration point after [x] and the value after [k] at the same duration point. (This cannot be shown by simply plotting F0 against Af, because Af is effectively the same for most tones, whereas F0, of course, is different). Figure 3 shows a two degree polynomial fitted to F0 differences plotted against Af differences for 36 duration points. The r^2 for this fit is 0.86. Although the frequency response of the electroaerometer is characterised (Baken 1987, 293) as only "moderate", relative to a pheumotachograph, for registering the Af changes found in all speech, its value of DC to 200 Hz or more (Smith n d,1) makes it unlikely that the temporal extent of the Af differences is an instrumental artefact.

Correlation does not necessarily of course imply causality. A causal relationship is, however, predicted by the received myoelastic-aerodynamic theory of phonation (Lieberman 1967, 14-18). The kinetic part of the second component of the theory which accounts for the Bernoulli force predicts that, other things like vocal cord tension being equal, differences in glottal airflow will be positively correlated with differences in F0. (Greater glotttal airflow will result in stronger Bernoulli force which will suck the cords together quicker; it can be assumed that the other, aerostatic, term of the aerodynamic component which has to do with subglottal pressure remains effectively constant during the

It is theoretically possible that originally intrinsic F0 differences associated with initial consonants come to be deliberately produced - this is part of the explanation for tonogenesis (Hombert et al., 1979). However, a direction of causality in the TP data whereby deliberate changes in F0 are causing changes in Af is excluded because under the normal conditions of F0 control by vocal cord tension a *negative* correlation between F0 and Af might be expected. This is because it is claimed that the reduced mean glottal area that occurrs during F0 increases results in a reduced airflow (Ohala 1978, 18). Extrinsic F0 changes would

not be expected to positively correlate with Af, unles they were produced by extrinsic pulmonic effort, and this is unlikely for Tai varieties (Rose 1984, 152, 154).

Table 1 (con't)

SP	T2 [xaat 44] n=2				T2 [kaat 44] n=3			
	F0		Af		F0		Af	
	X	sd	X	sd	X	sd	X	sd
0	242	7	233	22	224	4	60	2
5	-	-	169	10	-	-	40	11
10	231	13	129	5	219	11	34	11
20	226	11	89	6	221	10	40	15
40	227	8	64	26	227	8	41	19
60	225	6	56	22	228	8	47	26
80	226	6	79	24	226	8	68	27
90	-	-	107	23	-	-	101	32
100	223	4	120	15	225	10	118	40
D	19.4, 1.1				23.5, 1.6			

SP	T5 [xaat 21]				T5 [kat 21] n=2			
	F0		Af		F0		Af	
	X	sd	X	sd	X	sd	X	sd
0	240	7	254	31	243	11	52	10
5	-	-	176	29	221	17	-	-
10	229	7	125	29	-	-	29	11
20	218	8	78	21	207	10	26	4
40	203	9	51	17	203	3	26	7
60	185	7	44	19	192	1	28	7
70	175	4	45	18	-	-	29	6
80	170	5	59	16	183	1	35	1
90	147	10	81	18	-	-	-	-
100	-	-	98	28	173	4	54	9
D	19.8, 1.2				8.7, 0.4			

Finally , it is possible that the F0/Af relationship is epiphenomenal, and the good correlation between Af and F0 differences reflects differences in longitudinal vocal cord tension covarying with extrinsic glottal width. An intrinsic relationship between the two factors of horizontal tension and glottal width has been entertained, but dismissed, by Hombert (1987, 89,90). Also, in the feature system proposed by Halle and Stevens (1971), the two dimensions of glottal aperture ([+/- spread, +/- constricted glottis]), and cord tension ([+/- stiff, +/-slack cords]) are considered orthogonal. However, a stronger reason for rejecting the epiphenomenal argument that high F0 values are attributable to deliberate [+stiff] cords on the [x] is the incompatibility between stiff cords and phonation at large glottal widths. Since cord stiffness acts to inhibit phonation at large glottal widths (Halle and Stevens 1971) it is unlikely that the cords at phonation onset in TP [x] are in fact stiff.

SUMMARY

This paper has argued that the extensive F0 perturbations observed in Thai Phake tones on syllables with initial [k] and [x] are due to differences in airflow caused by the extrinsic
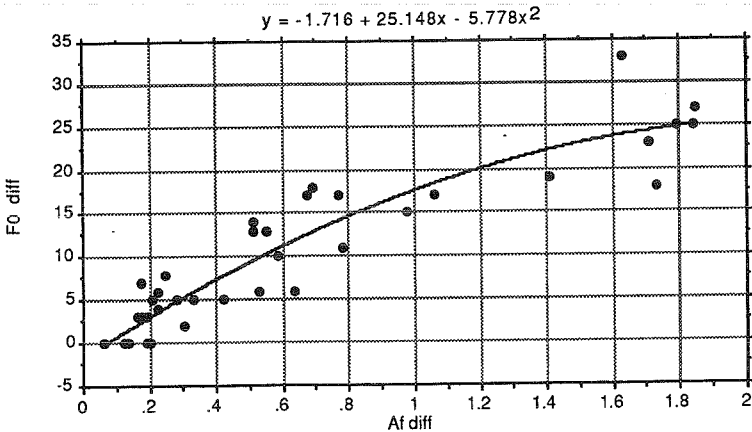
$$y = -1.716 + 25.148x - 5.778x^2$$

Figure 3. Scatter plot of differences in F0 against differences in Af in Thai Phake tones.

glottal width adjustments necessary for these consonants. This provides evidence against the received view that aerodynamic factors are not responsible for F0 perturbations. It is well known that the

phonological history of Tai involves cases of tone split conditioned by the two natural classes of voiceless aspirated stops and voiceless fricatives on the one hand, and voicless unaspirated stops on the other (Haas,1958). The former class -- so called [+spread glottis] segments in the Halle-Stevens feature system -- gives rise to higher tones, the latter -- [- spread glottis] segments -- to lower. The Thai Phake data provide a nice demonstration of how the initial stage of such splits could have occurred. It is also worth noting that the aerodynamic connection thus revealed provides the explanatory link between the appropriate articulatory feature [+/- spread glottis] and pitch (i.e. F0) height.

REFERENCES

Baken, R.J. (1982) *Clinical Measurement of Speech and Voice*, Taylor and Francis.

Halle, M. and Stevens, K.N. (1971) "A note on laryngeal features" *Quarterly Progress Report*, Research Lab. of Electronics MIT 101, 198-213.

Haas, M. (1958) "The tones of four Tai dialects' *Bulletin of the Institute of History and Philology*, Academia Sinica, 817-826.

Hombert, J-M. (1978) "Consonant Types, Vowel Quality and Tone" in V.Fromkin (ed.) *Tone A Linguistic Survey*, Academic Press, 77-111.

Hombert, J-M. Ohala, J. and Ewan, C. (1979) "Phonetic Explanations for the Development of tones'"*Language* 55/1, 37-58.

Klatt, D. Stevens, K.N. and Mead, J. (1968) "Studies of Articulatory Activity and Airflow During Speech" Annals of the New York Academy of Sciences Vol. 155/1, 42-55.

Lieberman, Philip (1967) *Intonation, Perception, and Language*, MIT press.

Ohala, J. (1978) "Production of Tone" in V.Fromkin (ed.) *Tone A Linguistic Survey*, Academic Press, 5-39.

Rose, Phil (1984) "The Role of Subglottal Pressure and Vocal Cord Tension in the Production of Tones in a Chinese Dialect" In B. Hong (ed.) *New Papers in Chinese Language Use*, A.N.U., 133-168.

Rose, Phil (1990) "Thai Phake Tones: Acoustic Aerodynamic and Perceptual Data on a Tai Dialect with Contrastive Creak." In R.Seidl (ed.) *Proc. 3rd Australian Intl. Conf on Speech Science and Technology:*, ASSTA, 394-399.

Smith, Svend (nd) *Instruction Manual for the Electro Aerometer Type AM 510/4.*