ANY ADVANCE ON ELEVEN?
LINGUISTIC TONETIC CONTRASTS IN A BIDIALECTAL THAI SPEAKER

Phil Rose
Department of Linguistics (Arts), Australian National University

ABSTRACT The use of a Thai bidialectal with at least ten different tones is demonstrated in establishing linguistic tonetic contrasts, and testing the suitability of tonal normalisation strategies and descriptive adequacy of current feature systems. A descriptive framework for linguistic tonetics is outlined.

INTRODUCTION

How do tone languages and dialects differ with respect to their tonal phonetics? One way of investigating this lingusitic tonetic question is to use bilingual speakers (Maddieson 1979), where it is likely that observed differences between varieties are not due to individual differences in vocal tract anatomy - since output from the same individual is involved - but realise *bona-fide* lingusitic phonetic differences. Bilinguals are unlikely to be homogeneous in their competence and performance. However, it is reasonable to assume that if a bilingual speaker produces a contrast in forms between their two languages, this reflects a real perceived difference. With the bilingual approach we are in a sense seeing through the ears of the native speaker.

In this paper, the tones of two Thai varieties - Pakphanang and Standard Thai - are compared acoustically using a female bidialectal speaker. These two varieties differ in their tonal inventory, and especially in their tonetics. Pakphanang (PPhN) is a Southern Thai dialect which contrasts seven tones on unstopped syllables. The auditory and acoustic characteristics of the tones are described in detail in Rose (in print). The PPhN system comprises four distinctive pitch contours - convex, level, falling and concave - for the first three of which height is also contrastive. The pitch shapes of the tones are as follows. T(one)1 very high convex, with its falling portion a little more salient than the rising. Voice quality is often falsetto, possibly as a consequence of the high pitch target. T2 is upper-mid level. T3 has convex pitch in the lower half of the pitch range, with the rising portion longer and more salient. T4 is mid-level. T5 has a short level pitch in the mid pitch range followed by a clear fall. T6 is low rising into the mid pitch range with an initial low level component. T7 has low falling pitch. It can be seen from this that the PPhN pitch range is neither optimally nor neatly exploited: the convex pitched tones 1 and 3 lie more or less at the extremes of the pitch range, but the level pitched tones 2 and 4 are concentrated in the mid pitch range, and the falling pitched tones 5 and 7 lie in the lower half of the pitch range. Standard Thai (Sd.T), which is based on the Central Thai variety spoken in Bangkok, contrasts five tones on unstopped syllables. Their conventional names are Mid, Low, Fall, High, and Rise but their realisation on unstopped syllables is somewhat more complex, involving more contouricity than implied by the descriptions, which reflect assumed tonemic distinctiveness. Thus the /Rise/ tone has a prolonged low onset, the /Mid/ and /Low/ tones fall somewhat; and the /High/ tone is high rising in many speakers (Gandour 1979:96). The /Fall/ tone is also heard with a prominent initial level component. These were the realisations the bidialectal speaker used for her Sd.T tones. Of the twelve tones in these two varieties, the speaker claims that two pairs are the same: Sd.T /Low/ and PPhN low falling tone 7, and Sd.T /Rise/ and PPhN low level-rising tone 6. The historical development of Sd.T and PPhN has involved differential mergers and considerable phonetic divergence. Consequently, cognates are, with the possible exception of Sd.T /Mid/ : PPhN mid falling tone 5, phonetically very different. For example, the Sd.T low rising /Rise/ tone corresponds to PPhN high convex tone 1, and the Sd.T high rising /High/ tone corresponds to PPhN low falling tone 7.

PROCEDURE

Two sets of PPhN and Sd.T words with comparable segmental structure were recorded in the phonetics laboratory at A.N.U. by a 46 year-old female native speaker in 1981. The speaker, like most well-educated local professionals, belongs to the sociolinguistic class of bidialectals who are 'typically fluent in standard Central Thai' and who have the 'ability to switch without mixing the seven-tone Southern and five-tone Central tonal systems' (Diller 1987:150-155).

The corpus is given in table 1. It comprised the five Sd.T and seven PPhN tones demonstrated on unstopped syllables with [a:] and [u:] Rhymes. The syllable-initial consonant was a voiceless aspirated stop [kh] or [ph] except in PPhN tones 3 and 4, where phonotactic constraints require unaspirated stops, and in the PPhN tone 7 with u:, where /r/ was chosen due to a lexical constraint. The words were elicited using Standard Thai orthography. Diller (1987:156) 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 PPhN forms were read first, followed by the Sd.T. The [a:] forms were read first for each variety, and the speaker paused after each item. Three repeats of the corpus were recorded. The informant did not know how many repeats were going to be requested.

PAKPHANANG

Historical Category	A H; B H	C H	A M; B M	C M	A L	B L	C L
Tone	1	2	3	4	5	6	7
[a:]	**khaa**	**khaa**	**kaa**	**kaa**	**khaa**	**khaa**	**khaa**
	leg	*kill*	*crow*	*mark*	*grass*	*value*	*trade*
[u:]	**khuu**	**phuu**	**kuu**	**kuu**	**khuu**	**phuu**	**ruu**
	threaten	*male*	*I*	*borrow money*	*ditch*	*wasp*	*know*

STANDARD THAI

Historical Category	A M,L	B H,M	C H, M; BL	C L	A H
Tone	/Mid/	/Low/	/Fall/	/High/	/Rise/
[a:]	**khaa**	**khaa**	**khaa**	**khaa**	**khaa**
	grass	*galangal*	slave	*trade*	*leg*
[u:]	**khuu**	**khuu**	**khuu**	**khuu**	**khuu**
	ditch	*threaten*	*pair*	*bend*	*much*

Table 1. Corpus illustrating the seven Pakphanang and five Standard Thai tones on [a:] and [u:] Rhymes. Historical tonal [A,B,C] and consonantal [H,M,L] categories are also shown.

The duration of the Rhyme was first determined from wide band spectrograms with their good time-domain resolution. Fundamental frequency (F0) was then sampled at percentage points of this duration from aligned expanded narrow-band spectrograms. Slightly different duration bases were used for some tones; these are described below. This method of F0 measurement has an accuracy of +/- 4Hz at the 90% confidence level. Radiated amplitude (Ar) was sampled from the oscillographic trace from an F/J intensity meter, using 20 msec. 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.

Arithmetical mean and standard deviation values for F0, Ar and D for the five Sd.T tones tones with [a:] and [u:] vowels, as well as VOT of initial obstruents are given in table 2. The corresponding values for the seven PPhN tones can be found in Rose (in print). The percentage points at which F0 and Ar were sampled are shown under "SP". So for example in /Mid/ tone words with [a:] vowels, the mean F0 at the 50% sampling point was 174 Hz, with a standard deviation of 1 Hz. The 50% sampling point occurred at (58.2 csec. x 50% =) csec. 29.1. Tones with final rising pitch, that is PPhN low rising tone 6, Sd.T /High/ and /Rise/ tones, were sampled as a function of duration to F0 peak. F0 on the PPhN low convex tone 3 was treated as a spline and sampled as a function of duration from phonation onset to F0 peak, and peak to offset.

In order to obtain acoustic values which better reflect the Sd.T and PPhN tones, two things were done. Many F0 and Ar perturbations observable before phonation offset were very similar to those which occur over the last 15 to 25 csec of phonation in Tai-Phake (Rose 1990:395-6). Concomitant airflow data indicate the Tai-Phake perturbations to be caused by abduction of the vocal cords in anticipation of the voicelessness of the following pause. These perturbations do not therefore constitute part of the individual tones' acoustical properties, and so they were discarded by ignoring F0 and Ar values after 80% of duration (PPhN tones 1,2,4,5,7; Sd.T /Mid/, /Low/ and /Fall/ tones); after F0 peak (PPhN tone 6; Sd.T /High/ and /Rise/ tones); and after 60% of duration from peak (PPhN tone 3) (these values gave the best agreement with the Tai-Phake data). Secondly, intrinsic differences associated with the [a] and [u] allotypes were also factored out by calculating the arithmetical means of all six tokens of a given tone. The resulting mean values are graphed as functions of absolute duration in figure 1. For ease of comparison, the F0 shapes of the tones are shown again in Figure 2 in two panels: the upper panel compares mean F0 for PPhN tones 1,3,4 and 5 and Sd.T /Fall/ and /Mid/ tones; the lower panel shows

/MID/ TONE

	[kha:]		[khu:]	
SP	F0	Ar	F0	Ar
0	- -	18.0,2.2	- -	16.8,0.2
5	198,13	18.7,2.3	214,2	23.6,0.3
10	193,4	21.8,2.9	209,3	24.5,0.4
20	183,1	21.4,2.0	203,4	26.1,0.3
30	180,2	22.5,2.1	202,1	26.8,0.2
40	178,3	21.9,1.7	199,2	26.9,0.3
50	174,1	21.5,1.3	196,2	27.3,0.3
60	172,2	20.5,1.1	191,3	25.7,1.1
70	163,4	18.6,1.2	180,10	22.4,1.6
80	154,4	14.9,1.5	166,19	17.8,3.7
90	152,1	11.8,0.9	161,29	13.0,1.3
95	144,6	- -	157,29	- -
100	139,5	3.8,1.3	154,33	4.5,1.0
D	58.2,0.9		54.8,0.5	
VOT	7.0,1.1		9.0,0.1	

/LOW/ TONE

	[kha:]		[khu:]	
SP	F0	Ar	F0	Ar
0	186,2	17.1,1.9	- -	16.7,2.0
5	181,2	21.8,3.3	182,2	24.1,2.9
10	174,1	22.6,2.8	180,2	25.2,2.6
20	161,2	23.5,2.0	172,4	26.1,2.0
30	155,1	22.9,2.1	166,7	26.0,0.6
40	149,2	21.4,1.6	161,5	24.7,0.9
50	142,0	19.9,1.8	154,4	22.7,1.2
60	138,1	17.9,1.5	148,5	20.0,0.4
70	135,1	14.8,1.8	144,4	18.1,1.4
80	136,4	11.4,1.3	138,5	13.8,3.1
90	132,4	6.8,1.4	147,5	13.0,1.0
95	137,8	- -	151,9	- -
100	135,7	2.8,1.0	145,8	3.7,0.6
D	61.6,0.3		55.0,2.7	
VOT	6.7,0.4		8.2,0.4	

/HIGH/ TONE

	[kha:]		[khu:]	
SP	F0	Ar	F0	Ar
0	- -	11.3,3.2	- -	14.9,0.8
5	187,2	15.1,1.1	198,5	20.9,1.1
10	187,2	16.7,0.8	203,2	23.0,2.0
20	180,2	17.5,1.3	204,1	22.9,1.4
30	180,1	18.2,1.2	206,3	23.2,1.4
40	183,1	18.3,1.2	211,5	23.0,1.0
50	184,1	17.8,1.0	214,2	23.2,1.2
60	187,2	17.6,1.1	217,3	23.4,1.9
70	190,3	15.8,1.7	221,6	23.2,1.9
80	195,4	13.4,1.9	228,3	18.8,2.8
90	204,3	11.0,1.2	235,5	14.1,1.7
100	220,7	3.2,1.0	245,6	3.5,0.9
F0off	209,10	- -	230,10	- -
D	48.2,0.7		42.4,2.1	
Doff	51.2,0.4		46.5,0.4	
VOT	5.3,0.3		6.0,0.8	

/RISE/ TONE

	[kha:]		[khu:]	
SP	F0	Ar	F0	Ar
0	- -	13.4,0.4	- -	13.2,2.3
5	152,2	17.8,0.7	159,5	19.9,1.4
10	145,5	18.3,1.0	158,6	20.8,1.3
20	138,3	17.8,0.8	152,7	20.3,0.9
30	136,1	17.7,0.7	151,4	20.8,0.9
40	137,2	16.9,0.5	151,4	21.0,1.1
50	138,4	15.1,0.8	153,4	20.8,1.1
60	140,7	14.8,1.4	154,3	19.2,1.0
70	143,7	12.7,1.2	157,5	17.4,0.3
80	153,9	10.4,1.7	171,6	14.3,1.3
90	171,12	7.8,2.1	198,8	13.8,1.2
100	198,18	5.0,1.0	226,13	4.4,1.2
F0off	192,20	- -	218,13	- -
D	51.9,2.2		45.1,1.7	
Doff	56.1,2.4		47.9,1.9	
VOT	7.0,1.1		8.3,1.8	

/FALL/ TONE

	[kha:]		[khu:]	
SP	F0	Ar	F0	Ar
0	- -	14.3,1.4	- -	14.9,1.9
5	240,1	15.8,1.9	244,6	19.6,2.4
10	237,5	18.1,1.5	245,4	22.1,1.8
20	229,8	19.6,1.2	248,7	23.8,1.8
30	228,4	20.4,1.2	246,5	23.8,2.6
40	220,5	20.4,1.1	242,5	23.4,1.9
50	202,8	19.0,1.1	221,4	23.8,1.2
60	176,9	17.8,1.3	196,4	22.2,1.8
70	157,5	15.7,1.8	167,9	18.3,2.1
80	150,5	12.3,1.7	148,4	13.8,0.8
90	139,4	8.5,4.0	145,6	9.5,2.8
95	141,3	- -	- -	- -
100	136,10	3.3,0.5	139,13	3.0,0.9
D	47.2,1.7		43.8,2.6	
VOT	7.2,0.6		7.3,0.5	

Table 2. Means and standard deviations (x,sd) for fundamental frequency (F0), radiated amplitude (Ar), duration (D), and voice onset time (VOT) in one Pakphanang speaker's unstopped Standard Thai tones on syllables with [a:] and [u:] rhymes. n = 3. Units are Hz (F0), dB (Ar), and csec. (D,VOT).
SP = sampling point.
D off = duration to phonation offset.
F0 off = F0 at phonation offset.

mean F0 for PPhN tones 2,6 and 7, and Sd.T /High/, /Rise/ and /Low/. Note the slightly different vertical frequency scales in both figures.

DISCUSSION

As the PPhN and Sd.T tones were produced by the same speaker in the same session under almost identical elicitation circumstances, the F0 data are directly comparable. This means they can be taken to give a correct picture of how the tones of the two varieties relate. The lower panel in figure 2 shows that the F0 shapes of the PPhN low falling tone 7 and the Sd.T /Low/ tone are almost exactly congruent. Their Ar values do not differ significantly either, once allowance is made for the effect of the different initial consonants (in /ruu/ vs. /khuu/). Sd.T /Low/ and PPhN tone 7 are therefore productionally the same, at least for this speaker. It may be thought that this should come as no surprise, given the speaker's prior claim as to their identity.

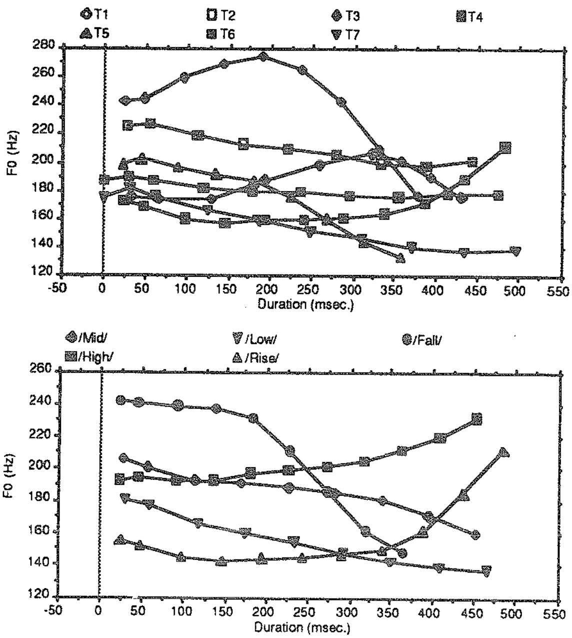

Figure 1. Mean tonal F0 and duration values for (above) Pakphanang and (below) Standard Thai.

However such inferences are apparently not valid, given the difference between the other pair - Sd.T /Rise/ and PPhN low rising tone 6 - for which identity was also claimed. The bottom panel of figure 2 shows both these tones have the same F0 peak, and contour, but the PPhN tone 6 lies about 15 Hz statistically significantly higher over the first 70% of its duration. This difference is also audible. It is difficult to explain this in terms of assimilatory effects due to different position in the elicitation sequence. Perhaps the PPhN tone is free to occupy a higher position in the speaker's tonal range, because there is no other higher concave or rising tone in the paradigm with which it might be confused. Such a tone *is* present in the Sd.T configuration, however, since the speaker's /High/ tone is phonetically clearly rising. If this is correct, it would add structural knowledge, of paradigmatic tonal contrasts, to the list of factors, like intrinsic vowel F0, which mediate speech perception. In any case it appears that the speaker reports two tones to be the same which she produces differently. Whether this counts as a lingusitic tonetic contrast is difficult to say, but it is clear that the speaker is controlling at least ten different tonal F0 shapes.

With regard to a lingusitic phonetic representation, Anderson (1978:133,134) states "Any phonetic property ...in terms of which utterances in Language X are systematically different from utterances in Language Y...must find some reflection in the system of phonetic transcription provided by lingusitic theory." Even a cursory glance at the linguistic-tonetically different F0 shapes in figures 1 and 2 shows that they will provide serious representational problems for current feature systems. There are, for example, eight significantly different F0 levels - three more than are allowed for by feature systems constructed on the basis of maximal contrast within languages. The contrast between PPhN tones 5 and 4 and CT /Mid/ is especially difficult to capture, involving different durations of effectively level F0 before a fall. In Yip's (1980) system, for example, only two falling tones are possible: [+U,HL]; [-U,HL], yet the combined PPhN/Sd.T data contain four: Sd.T /Fall/, Sd.T /Low/ and PPhN tone 7, PPhN tone 5, and CT /Mid/. Note that CT /Mid/ cannot be considered as level, because of the contour contrast with the true level PPhN tone 4 [-U,H]. The representation of the high convex PPhN [+U, LHL] is also a problem in relationship to Sd.T /Fall/ [+U,HL], since it implies the identity of falling parts (HL) which are clearly different.

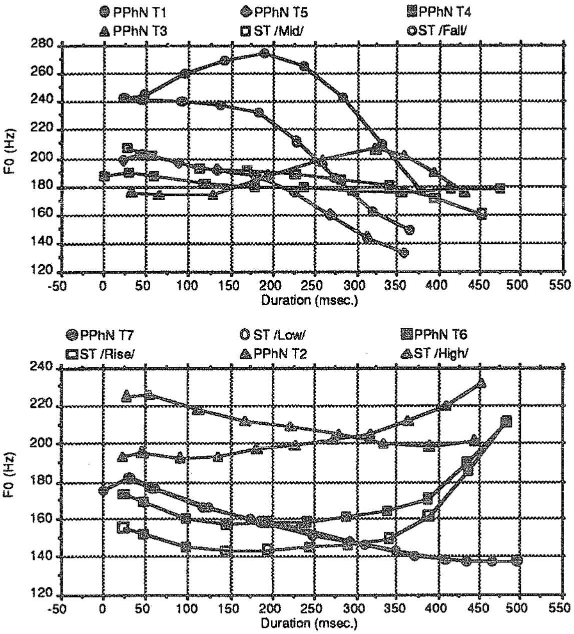

Figure 2. Pakphanang and Standard Thai (ST) tonal F0 compared.

An alternative approach, suggested by the shapes in figures 1 and 2, is to describe the F0 in terms of two components: *Position* and *Contour*. The Position component locates the contour in a particular region of the F0 range, and is continuous. Evidence for a continuous Position component comes from pairs like PPhN tone 5 and CT /Fall/ (or PPhN tone 6 and CT /Rise/, if it is decided to call it a linguistic tonetic contrast). The F0 shape of both tones in these pairs has the same contour in the well-defined sense of Rose (1982:35), but differs in Position. The contour component is internally structured, consisting of a basic contour plus optional modifications to its onset and offset. The PPhN/Sd.T F0 data show three basic contours: R(ise) F(all) and L(evel), and the same modifications (symbolised in lower case): r(ise) f(all) and l(evel). Thus, apart from differing in Position, the PPhN convex tones 1 and 3 appear to instantiate a contrast in basic contour of Fall vs. Rise, with the Fall modified by an initial rise in tone 1: [r.F], and the Rise by a final fall in tone 3: [R.f]. The CT /Fall/ is another example of a basic Fall, like PPhN tone 1, but it has a level onset: "l.F". The abovementioned problematic threeway contrast between PPhN tones 4 and 5, and Sd.T /Mid/ can be represented as [L] (PPhN 4) vs. [L.f] (Sd.T /Mid/) vs. [l.F] (PPhN 5).

The next point to note is the difference in F0 range between the two varieties. Both have very similar lower limits, but the PPhN upper limit, defined by the peak F0 of the high convex tone 1, is higher by some 30 Hz than that of the Sd.T tones, at the onset of the /Fall/ tone. This difference of 30 Hz is not small: it is about 30% of the Sd.T range. This is a clear indication that F0 range is a linguistic tonetic parameter: Pakphanang has a greater F0 range than Standard Thai. Differences in F0 range have been claimed between Vietnamese dialects (Phuong 1981), on the basis of comparison of mean F0 values of different speakers, but it is not clear how such an accentual difference could be conclusively demonstrated without recourse to bilinguals. (It is not clear if it could be perceived by a transcriber, either, but that is another matter.)

Because the bidialectal data reveal the correct relationship in tonal F0 between the two varieties, they can be used to evaluate normalisation strategies. Any strategy which resolves the F0 of the two varieties in a fashion similar to that in figure 2 can be considered successful. It is apparent from the data that a percent-of-range normalisation using maximum and minimum values would considerably distort the relationship between the two F0 configurations, because of the significant difference in F0 range.

However, the minimum value of a low falling tone, and peak of a low rising tone might be useful range-defining points, since they are clearly comparable between the two sets. More interesting is the fact that both sets of tones have very similar standard deviations: 32.2 Hz (PPhN); 30.8 Hz (Sd.T), and their means only differ by 5 Hz: 188.4 Hz (PPhN), 182.8 Hz (Sd.T). (These were calculated on F0 values at sampling points from 5%-80% inclusive.) A z-score normalisation using these parameters would therefore yield very nearly the correct relationship. It is perhaps surprising that such different F0 configurations would yield such similar means and standard deviations, especially in the light of Maddieson's (1979) findings of significant between-language differences in these parameters. Further research on bilinguals, with, for example, the copious data in Anivan (1985) will show whether this is a fortuitous result, or whether it is perhaps related to the complexity of the F0 configurations involved. Maddieson (1979) tested African tone languages with only two or three level tones. Perhaps in tone languages with a sufficiently complex set of tones, for example five or more, including contour tones, means and standard deviations are comparable.

ACKNOWLEDGEMENTS

I should like to thank Dr. Tony Diller for all his help and encouragement in this study, and Ms. Sutira Wacharaboworn for acting as informant during her stay as visiting fellow at the A.N.U., and effortlessly controlling ten (eleven?) tones.

REFERENCES

Anderson, S.R. (1978) "Tone features", in V.A. Fromkin ed. *Tone: a Linguistic Survey* , pp. 133-175, (Academic Press:New York).

Anivan, Sarinee. (1985) An Acoustical Study of Thai Tones. Unpublished Ph.D. thesis, Macquarie University.

Diller, A.V.N. (1987) "Reflections on Tai Diglossic Mixing", Orbis 32 1-2, 147-166.

Gandour, J.T. (1979) "Tonal Rules for English Loanwords in Thai", in T.L. Tongkhum et al. eds. *Studies in Tai and Mon-Khmer Phonetics and Phonology in Honour of Eugenie Henderson*, pp. 94-105, (Chulalongkorn University Press: Bangkok).

Maddieson, I. (1979) "Tone spacing: evidence from bilingual speakers", UCLA Working Papers in Phonetics 45, 84-88.

Phuong, V.T. (1981) The acoustic and perceptual nature of tone in Vietnamese. Unpublished Ph.D. thesis, A.N.U.

Rose, Phil. (1982) "Acoustic Characteristics of the Shanghai-Zhenhai syllable-types", in D.Bradley ed. *Tonation*, pp. 1-53, Pacific Linguistics Papers in South-East Asian Linguistics 8.

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*, pp. 394-399, (Australian Speech Science and Technology Assn.:Canberra).

Rose, Phil. (in print) "A Seven Tone Dialect in Southern Thai with Super-high: Pakphanang Tonal Acoustics and Physiological Inferences", to appear in a Festschrift.

Yip, M.J.W. (1980) *The Tonal Phonology of Chinese* (Indiana University Linguistics Club:Bloomington).