

Mr. White goes to Market - Running Speech and Citation Tones in a Southern Thai Bidialectal

Phil Rose^{1,2}

¹ College of Asia and the Pacific, Australian National University, Australia

² Division of Humanities, Hong Kong University of Science & Technology, SAR China

philjohn.rose@gmail.com.au

Abstract

The nature of running speech tones and their relationship to citation tones are examined for two varieties of Thai - Standard and Southern - spoken by a bidialectal female. It is shown that representative tones from these varieties are readily extracted from running speech and that their mean values generally relate quite straightforwardly to contemporaneously elicited citation tones when control for general utterance position is taken into account.

Index Terms: running speech tones, citation tones, bilingual tones, Southern Thai, Standard Thai

1. Introduction

In the description and analysis of linguistic tone, for practical as well as theoretical reasons, citation forms play a central role. Defined as occurring in isolation and in opposition to connected speech [1], citation forms are easy to elicit - often because they relate clearly to an existing orthography - and this in turn makes them easy to transcribe and quantify. Above all, they are well-controlled.

The representative status of citation forms is, however, problematic [2, 3]. Whereas many emphasize the scarcity of citation forms in normal speech, at the same time there is a belief that they are closer in some unspecified way to *Langue*, or *Competence*:

"The citation form of a [Thai] monosyllabic word may be viewed as bearing the ideal manifestation of a tone. Of course, except for the occasional one-word sentence, such ideal forms do not often occur in running speech." [4].

Labovian socio-phonetics showed a long while ago that citation forms constitute just one category in the range from word-list style through casual speech. To the Labovian paradigm we owe also the insight that citation forms may display phonological contrasts that are not found in less formal speech, and vice-versa. So it should not be taken as axiomatic that one can derive, or otherwise relate, running speech from citation forms; rather, it should be demonstrated. Kratochvil [5] articulates this, almost Jonesian, position:

... a phonological class, such as a particular tone ... is understood as an abstraction which must be based on the properties of the phonetic signal shared by all or most concrete instances of that class. ... The norm of a tone ... is then taken to be the sum of the phonetic properties on which that tone as a phonological class is based. ... a concrete instance of that tone is seen as a compound of the norm and the effects of some other factors, rather than, for example, an approximation of an ideal independently existing standard. ... The norm must be extractable analytically ... from all the available instances of the given tone, it should not be sought outside the instances. ... This excludes, in particular, the use of the properties of isolated citation form tones as norms in an investigation of tones in normal speech".

An additional consideration is the representativeness of citation forms when based, as they usually are, on a restricted

set within an established tradition, e.g. *ma* syllable tones for Mandarin, *khaa* for Thai etc.. The assumption is that the acoustic forms in the citation phonological environment will be representative of the variety, but this is highly questionable: the acoustics of the Standard Thai /Fall/ tone on *khaa* will be rather different from those on *khaang khang* or *paa*, for example.

This paper is motivated primarily by Kratochvil's pioneering work, dating from the mid-sixties, on running speech tones - measured on each and every syllable - in Peking dialect e.g. [6, 5]. Apart from this, and other early work on Thai e.g. [4, 7], it is a pity that, given how easy it is to quantify them now, true running speech tones (as opposed to tones in, say, isolated polysyllabic words) still remain linguistically largely unexplored. The aim of this paper, then, is to look at running speech and citation tones in Standard and Southern Thai to see to what extent (1) coherent tones can be extracted from uncontrolled running speech; and (2) the citation forms give a good idea of what is going on in real speech.

2. Citation tones

Standard Thai, which is based on the Central Thai variety spoken in Bangkok, contrasts five tones on sonorant-final syllables. Their (citation!) realisation involves more pitch contouricity than implied by their conventional names of **Mid**, **Low**, **Fall**, **High** and **Rise**, which reflect assumed tonemic distinctive features. Thus the /Rise/ tone has a low dipping pitch; /Mid/ and /Low/ tones fall somewhat; and the /High/ tone is high rising in many speakers [8]. The /Fall/ tone is also heard with a prominent initial level component. The top row of fig. 1 shows the mean acoustics (F0 as a function of absolute duration) of a female speaker's Standard Thai tones from non-contemporaneous recordings separated by five years. Like most well-educated local professionals, she was typically fluent in both Standard Thai and her local dialect and could switch effortlessly between the seven tone Southern and five tone Standard without mixing [9]. The first recording was in 1982; the speaker was then 41. The F0 data in fig. 1 represent citation tones on low vowel /aa/ Rhymes only, using the conventional and well-known Standard Thai aspirated velar onset minimal quintuplet of *khaa grass* (**Mid** tone), *khàa galanglal* (**Low**), *khâa slave* (**Fall**), *khâa trade* (**High**), *khâa leg* (**Rise**). The second recording, made in 1987, used the same corpus as the first, supplemented with *khâaw rice*, *khâaw white*, *khâaw news*, the extra forms being included for comparability with a running speech text. The top right panel of fig. 1 also includes the mean F0 for a /Fall/ tone with a coincident VOT /p/ onset, to demonstrate the abovementioned point about the potential un-representativeness of citation forms with fixed segmental structure: it can be seen to differ considerably from the F0 on the tokens with *kh* onset.

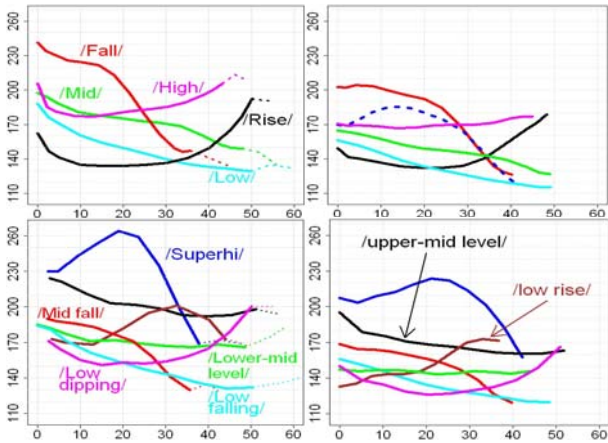


Figure 1: Mean F0 of one female's Standard (top row) and Southern Thai sonorant-final citation tones on two non-contemporaneous occasions (left column = 1982, right column = 1987). X-axis = duration (csec.), y-axis = F0 (Hz.) Blue dotted line = Standard Thai /Fall/ tone with [p] onset.

The bottom row of fig. 1 shows mean acoustics, on /aa/ Rhymes, of citation forms from the Southern Thai dialect of Pakphanang, which contrasts seven tones on syllables ending in a sonorant [10]. Named after their pitch, these are as follows.

Superhi, with convex pitch very high in the pitch range: *khaa leg*, *yaa type of fish*, *naa thick*. **Upper-mid level**, with level pitch in the upper part of the mid pitch range: *khaa kill*, *yaa grass*, *naa face*. **Low-rising**, with rising pitch in the lower half of the pitch range. An optional falling component results in a convex pitch: *kaa crow*, *paa jungle*, *yaa medicine*. **Lower-mid level**, with level pitch a little lower than the middle of the pitch range: *kaa mark*, *paa aunt*, *yaang roast*. **Mid-fall**, with pitch falling from the middle of the range to low: *khaa thatch grass*, *yaa don't*, *naa rice field*. **Low dipping**, with low level pitch rising into the mid pitch range: *khaa value*, *naa worth doing*, *yaa father's mother*. **Low fall**, with falling pitch in the lower third of the pitch range: *khaa trade*, *ruu know*, *naa mother's younger sibling*.

Fig. 1 immediately shows one problem with citation tones: they can vary. The speaker's citation tones, both Standard and Southern, differ over time in both F0 range and location. The earlier sets also appear to have been said more slowly, but this is because they included some obvious offset perturbations. There are also some differences in individual tones: the pitch/F0 of the Southern Thai /low rising/ tone is convex in the earlier recording but rising in the later. The F0 of the Southern Thai /low dipping/ tone differs from the Standard Thai /Rise/ tone at offset in the first recording and at trough in the second. Otherwise, the configurations are largely stable, with both recordings showing one shared tone only - Standard /Low/ and Southern /low fall/ - and the Southern Thai having a considerably greater F0/pitch range by virtue of its /superhi/.

3. Running Speech Tones

The running speech used to furnish tones for comparison with the citation forms was a short story recorded at the same time as the second set of citation forms. The story was ingeniously composed by the informant around some of the minimally contrasting words used in the citation forms: *khàa galangal*; *khàaw news*; *khàa value*, *slave*; *khàaw rice*; *khàa trade*; *khàa leg*; *khàaw white*. The story, told using a written prompt, recounts how Mr. White the merchant spent a successful day at market. It is transcribed below in Standard Thai

romanisation with periods indicating (pause-delimited) utterance boundaries.

naay khàaw. tâng bânruan yùu rim mǎnâam. mii aachiiip khàakhàay.
Mr. White built a house on the riverside. He was merchant by trade.

khàw dâai khàaw wâa. mii ngaan wât.

He got news of a temple festival.

cing aw sǐnkháa. sày rǔa pay khàay.

Then he took his produce, put it in a boat and went to sell.

khàapay. sǐnkháa tem rǔa.

On the way there the boat was full of produce.

khàakláp. lǔa sǐnkháa. phiang nǐtǔy.

On the way back there was only a little left.

sǐnkháa khǒng khàw khàay dii. phrǔ' raakhaa mây phaang.

His produce sold well, because the price wasn't expensive.

khàw dâay kamray khúm khàa nuây.

The profit he made was worth all the effort.

khàw khít wâa. khráng tǔo pay. yàak khàay khàaw kaang dǔay.

He thought that in future, he'd sell both rice and curry,

yàak phât phêt plaalâay. sày khaa mǎakmǎak.

and fried spicy eel, and put in lots of galangal.

phrǔom thàng tham kàytǔmkháa.

At the same time he'd make chicken soup with galangal.

khàw mây mii khàathâat chuây lǔa. tǔng tham khon diaw.

He didn't have a servant to help him. He had to do it on his own.

To extract the running speech tonal acoustics, a wideband spectrogram of a small portion consisting of two to three syllables was made in Praat with a superimposed pitch (sic: F0) trace using a fixed *time step* setting of 0.005 seconds. Then, following [5], the tonally relevant F0 for each sonorant-final syllable was extracted along with its duration by reference to the segmental structure of the syllable. This was done because it has been shown for several Asian tone languages that citation tonal F0 is distributed over the sonorant portion of a syllable Rhyme, but not over the syllable Onset, and F0 on voiced onsets (which can be sonorant or obstruent in Thai) had therefore to be excluded. (Tones on stopped syllables, e.g. *wât temple*, *mâak very*, were not frequent enough to warrant quantification.) The F0 and duration information was then copied to a text file for further processing in R. One of the auditorily salient aspects of Thai prosody is the extended length and well-defined tonal pitch contour of utterance-final syllables, and in order to see to what extent running speech tones are modified by utterance position, the data were coded for tone and immediate tone environment allowing partitioning into utterance-initial - medial and -final tones. Tones were identified and checked with the help of an expert on Southern Thai. To parametrise the tonal F0, code was written to fit a seventh order orthogonal polynomial, and tonal F0 then sampled from it at 10% sampling points of the sonorant Rhyme.

Fig. 2 shows the distribution of the F0 time course in the first two utterances of the story in both Standard and Southern Thai. Different colors are used to pick out the F0 on the constituent tones, which are named at the bottom of each panel, and the thicker grey line shows the polynomially-modeled tonal F0. A rough segmentation is also provided. Non-tonally relevant F0 on voiced onsets, i.e. [n] in *naay*, [b] in *bân*, [ɹ] in *rǔan*, [j] in *yùu*, [m] in *mǎa*, is shown with a dotted line. The smoothing effect of the polynomial modeling on the speaker's fairly jittery raw F0 is clear on e.g. both Standard and Southern *khaaw*.

Fig. 2 shows firstly that there is a longer pause between the first and second Standard Thai utterances than the Southern, but otherwise the tempo is very similar for both (the Standard is about 0.1 second longer). Of central importance is the obvious preservation of citation-like tonal shapes in the running speech. All five Standard tones are exemplified, and

all but the /upper-mid level/ Southern tones, and it would be possible to identify the running speech tones rather well by comparison to their citation form shapes in fig. 1. Several sandhi effects were clearly audible in the Southern Thai story. Most related to morphemes with the /low rising/ tone which had a pitch indistinguishable from the /low falling/ tone (yuu at in fig. 2 is an example; plaa fish in plaaalay eel another). The /superhi/ was only realised by convex pitch utterance-finally, otherwise it was high falling.

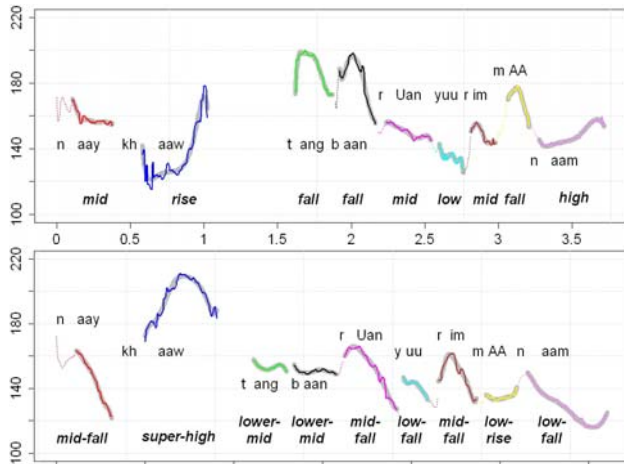


Figure 2: F0 of the first two utterances of story in Standard (top) and Southern (bottom), showing segmentals and tone. Thin lines = raw F0, thicker lines = polynomially smoothed F0, dotted lines = F0 on onset segments. X-axis = duration (sec.), y-axis = F0 (Hz.)

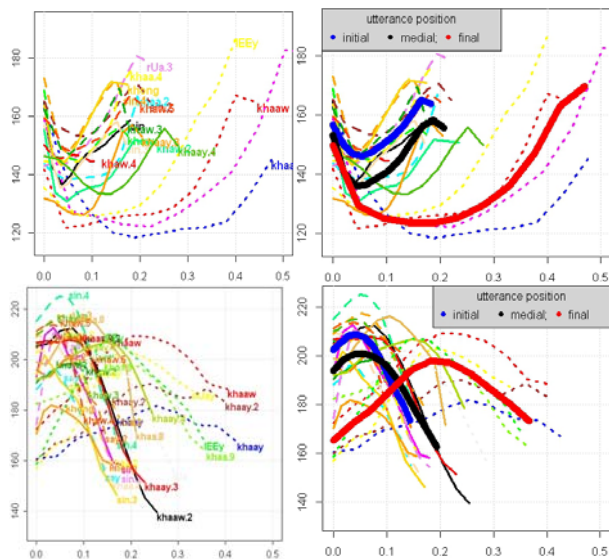


Figure 3: Acoustics of running speech tones. Top row = Standard Thai /Rise/, bottom row = Southern Thai /super-hi/. Thick lines show means correlated with utterance position. X-axis = duration (sec.), y-axis = F0 (Hz.)

The F0 range difference between Standard and Southern noted above appears to be maintained: the peak of the utterance-final Southern /superhi/ in *khaaw* is still ca. 10 Hz higher than the peak of the utterance-initial Standard /Fall/ in *tang*. Fig. 2 illustrates well the rather different tonal F0 shapes for the same cognates in both varieties, for example the diametrically opposed low dipping pitch of the Standard /Rise/ on *khaaw* compared to the high convex pitch of the Southern /superhi/. This apparent display of tonal virtuosity involved in

manipulating such different systems is commonplace.

As expected from the different lexical frequency of tones in running speech, the number of Standard Thai tokens available for each tone was not uniform. Broken down by position, their incidence was as follows (/tone/, initial, medial, final = total): /Mid/, 4, 18, 6 = 28; /Low, 2, 4, 2 = 8; /Fall/, 2, 10, 3 = 15; /High/ 2, 5, 2, = 9; /Rise/, 9, 6, 4, = 19. Given its seven-tone system, the incidence of some of the Southern Thai tokens was sparser still: /super-hi/, 11, 8, 7 = 26; /upper-mid level/, 1, 2, 0 = 3; mid fall/, 3, 11, 3 = 17; /lower mid/, 2, 6, 1 = 9; /dipping/, 0, 3, 1 = 4; /low fall, 2, 9, 3 = 14; /low rise/, 0, 3, 4 = 7. The small number of tokens for some tones in some positions means of course less confidence for their mean values.

4. Results

4.1. Variation in running speech tones

The results for both Standard and Southern Thai show a clear conditioning of tonal acoustics by utterance position. Fig. 3 plots the acoustics of Standard Thai /Rise/ and Southern Thai /superhi/ in running speech. Although the raw individual data in the left panels look messy, it can be seen from the right panels, where tonal means are plotted separately for utterance-initial, -medial and -final position, that they resolve coherently by taking their utterance position into consideration. It is clear that the initial and medial tones are shorter than the final forms, and also differ from them in F0 shape. With the Southern Thai /superhi/ tone, for example, the abovementioned difference between the utterance-final convex pitch and non utterance-final high falling pitch is very clear. This allotony has also been demonstrated for the /superhi/ tone in disyllabic words in the neighboring variety of Ron Phibun [11]. The non utterance-final tones have similar duration and F0 shape, with the utterance-initial forms lying slightly higher. Although the other tones in both Southern and Standard varieties have fewer tokens in some positions, the same patterning was also seen with them.

4.2. Running speech and citation tones

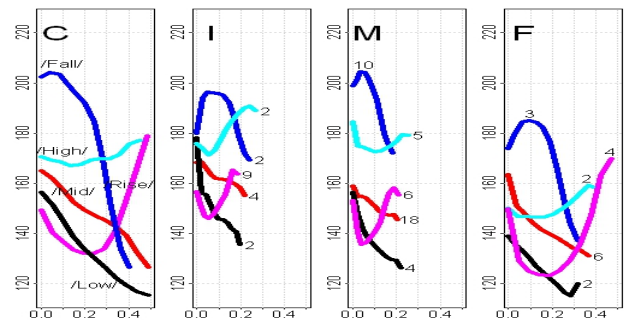


Figure 4: Standard Thai citation and running speech tones compared. C,I,M,F = Citation, Initial, Medial, Final. Number of tokens measured is indicated on plot. X-axis = duration (sec.), y-axis = F0 (Hz.)

Fig. 4 shows the relationship between mean citation and running speech tones for Standard Thai. It can be seen that the running speech has five tones in all positions, and also that, despite the small numbers of tokens in some positions, their mean acoustic shapes relate rather clearly to their citation forms. In configuration, the running speech tones resemble the citation tones most closely in final position. In medial and initial position the main configurational difference is in the so-called dynamic /Fall/ and /Rise/ tones, which neither fall nor

rise as much as in citation form. Importantly, this replicates a unexplained finding from an earlier study on Standard Thai running speech tones [7]. A possible reason for the apparent truncation may lie in the way running speech tonal acoustics align with their segmentals. Tonal acoustics were sampled here, as in [7], over the so-called tone-carrying part of the syllable, which excludes the Onset. However, in running speech it may be that the vocal cord tension gesture generating the tonal F0 spills over from the Rhyme onto the Onset of a following syllable (a nice example of this can be seen in the F0 on /n/ in last two syllables mAA naam in fig. 2). This can only be part of the story because truncation also occurs a little with word-final tones, but it is obviously a hypothesis to be tested by resampling tonal F0. All these findings apply *mutatis mutandis* to Southern Thai.

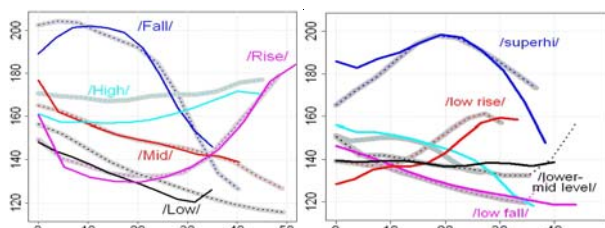


Figure 5: Left: Standard Thai citation tones predicted from utterance-final running speech tones. Right: Southern Thai utterance-final running speech tones predicted from citation tones. X-axis = duration (sec.), y-axis = F0 (Hz.)

Given the degree of similarity between the citation and running speech tones in fig. 4, it is relatively straightforward to demonstrate their relationship quantitatively. The left panel of fig. 5 shows, for example, how well the mean Standard Thai citation tone F0 in the top right panel of fig. 1 can be predicted from the running speech utterance-final tones (actual values are grey with colored dots; corresponding predicted values are solid colored). To do this, the configuration of the running speech utterance-final tones was first modeled with a z-score normalisation, and the normalised values then transformed with reference to the long term mean and standard deviation of the running speech. Relative duration was then adjusted slightly to give the best visual fit (increasing the duration of the utterance-final tones by a factor of 1.1%). In essence this is saying that the citation tones can be understood as the *configuration* of the utterance-final tones when realised with the *range* of the running speech tones. It can be seen that, apart from the /High/ tone, the fit is rather good.

The transformation is of course reversible, so that one can predict the running speech tones from the citation tones. This is illustrated in the right panel of fig. 5, which shows how well Southern Thai utterance-final tones can be predicted from citation tones (this time the normalised citation tones were transformed with the mean and standard deviation of the raw running speech utterance-final tones, with a 1.1% duration adjustment). There were insufficient numbers of dipping and upper-level tones for the comparison, but otherwise, again, it can be seen that the fit is good.

This bidirectional relationship means that, at least for this Standard and Southern Thai bidialectal, there is no obvious reason why any unidirectional derivational relationship should be preferred over the default phonological position of considering both citation and running speech tones as conditioned realisations of a separate, *emic*, form. To compare tones usefully between varieties, one needs a fully-specified *emic* representation, perhaps along the lines of [5], where it was shown that the F0 of running speech tones in Beijing

dialect could be predicted from their duration (the longer the duration, for example, the more pronounced the tonal contour). It would seem from figs. 4 and 5 that the Thai data may also exemplify this behaviour: it is worth finding out.

5. Summary

This paper has shown how representative tones can be recovered rather easily from running speech, by the simple process of getting mean values from a large sample of each tone category. It has also demonstrated that, perhaps not surprisingly, citation tones and running speech tones are closely related. However, given the ease of running speech tone extraction it may be sensible in future to work directly with them, rather than rely indirectly on citation forms. Certainly, when collecting data for cross-dialectal tonal comparison, examples of running speech should be included long enough to furnish many occurrences of all tones of differing segmental structure in all utterance positions.

6. Acknowledgements and Dedication

I thank Prof. Tony Diller for recording in Thailand the citation tones and the story of Mr. White, and for help in identifying the southern Thai tones. This paper is dedicated to Pavel Kratochvil, who, like our lovely language informant Ms. Sutira Wacharaboworn, is no longer around to receive the acknowledgement he deserves. Paul never stopped arguing that in order to properly study tones in real speech, you should actually study real speech, not citation tones. It must have seemed to him during our Friday supervisions down *The Three Horseshoes* some 40 years ago that his Popean message was falling on the enthusiastic, but nevertheless unreceptive ears of the next generation. But, it did actually register.

7. References

- [1] Crystal, David, A Dictionary of Linguistics and Phonetics [3rd edition], Blackwell, 1991.
- [2] Noss, Richard B., "How useful are citation forms in synchronic Thai phonology?", in J.G. Harris and J.R. Chamberlain [Eds.] *Studies in Thai Linguistics*, Central Institute of English Language, 274-284, 1975.
- [3] Tingsabadh, M.R. Kalya, "Loss of Preeminence of Citation Forms in the Study of Tones in Thai Dialects", in Bamroongraks et al. [Eds.] *The International Symposium on Language and Linguistics*, 224-227, 1988.
- [4] Abramson, Arthur.S., "Lexical Tone and Sentence Prosody in Thai", *Proc. 9th Intl. Cong. Phonetic Sciences*, 380-387, 1979.
- [5] Kratochvil, Paul, "Variable Norms of Tones in Beijing Prosody", *Cahiers de linguistique Asie orientale* 14(2): 153-174, 1985.
- [6] Kratochvil, Paul, *The Chinese Language Today: Features of an Emerging Standard*, Hutchinson University Library, 1968.
- [7] Tingsabadh, M.R. Kalya and Deeprasert, Daranee, "Tones in Standard Thai Connected Speech", in A.S. Abramson [Ed.] *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*, 297-307, 1997.
- [8] Gandour, J.T., "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*, 94-105, Chulalongkorn University Press, 1979.
- [9] Diller, A.V.N. "Reflections on Tai Diglossic Mixing", *Orbis* 32 (1-2): 147-166, 1984.
- [10] Rose, P. "A Seven-tone dialect in Southern Thai with Super High: Pakphanang Tonal Acoustics and Physiological Inferences", in A. Abramson [Ed.] *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*, 191-208, 1997.
- [11] Thompson, Napier Ian Guy, "Tone Sandhi Between Complex Tones in a Seven Tone Southern Thai Dialect", *Proc. 5th ICSLP* (2): 53-56, 1998.