

# JOURNAL OF THE SOUTHEAST ASIAN LINGUISTICS SOCIETY



Volume 5  
2012



# JSEALS

Journal of the Southeast Asian Linguistics Society



Volume 5, 2012

## *Managing Editor*

Paul Sidwell (The Australian National University & CRCL, Bangkok)

## *Editorial Advisory Board*

Mark Alves (Montgomery College, Maryland)  
George Bedell (Payap University)  
Mark Brunelle (University of Ottawa)  
Gerard Diffloth (Ecole française d'Extrême-Orient, Siem Reap)  
San San HNIN TUN (INALCO / LACITO, Paris)  
Deth Thach Joseph (INALCO, Paris)  
Marlys Macken (University of Wisconsin-Madison)  
Brian Migliazza (SIL International)  
Peter Norquest (University of Arizona)  
Amara Prasithrathsint (Chulalongkorn University)  
Martha Ratliff (Wayne State University)  
Keralapura Shreenivasaiah Nagaraja, (CIIL Mysore)  
Sophana Srichampa (Mahidol University)  
Alice Vittrant (Université de Provence / CNRS-LACITO, Paris)  
Justin Watkins (SOAS, London)

JSEALS is the peer-reviewed journal of the Southeast Asian Linguistics Society, and is devoted to publishing research on the languages of mainland and insular Southeast Asia.

JSEALS was formally established by decision of the SEALS 17 meeting, held at the University of Maryland in September 2007. It supersedes the Conference Proceedings, previously published by Arizona State University and later by Pacific Linguistics.

JSEALS welcomes articles that are topical, focused on linguistic (as opposed to cultural or anthropological) issues, and which further the lively debate that characterizes the annual SEALS conferences. Data papers, book reviews, and letters are welcome submissions, subject only to internal review.

Publication is continuous online, and consolidated annually in December. Papers should be submitted electronically to the Managing Editor, Paul Sidwell (paulsidwell@yahoo.com).

Submission is open to all, although we expect that most JSEALS articles have been formally submitted for peer review and publication after having the opportunity to be presented and discussed at the SEALS conference. Note that papers are expected to be written in English.

As a service to the community, non-reviewed conference presentations are archived on our website (www.jseals.org). The SEALS On-Line Archives holds the texts of most of the yearly conference publications.



*Asia-Pacific Linguistics: Open Access*  
THE AUSTRALIAN NATIONAL UNIVERSITY  
ISSN 1836-6821



# Contents

Volume 5, 2012

Notes on Grammatical Vocabulary in Central Vietnamese Mark J. Alves	1
A Tonal Analysis of Contemporary Tai Khuen Varieties <i>R. Wyn Owen</i>	12
Tones of Thai Song Varieties <i>Somsonge Burusphat</i>	32
Relative Clauses in Xong (Miao-Yao) <i>Adam Sposato</i>	49
Identity Maintenance in Lanna (Northern Thai) <i>Phattharathanit Srichomthong.</i>	67
Planning and Implementing Patani Malay in Bilingual Education in Southern Thailand <i>Suwilai Premsrirat &amp; Uniansasmita Samo</i>	85
Co-phonology vs. Indexed Constraint Theory: A Case Study of Perak Dialect Partial Reduplication <i>Sharifah Raihan Syed Jaafar</i>	97
Applicative Verbs and Applicative Construction in the Bantik Language <i>Atsuko Utsumi</i>	107
On the Historical Semantic Changes of the Thai Morpheme <i>hâj</i> <i>Kiyoko Takahashi</i>	126



## NOTES ON GRAMMATICAL VOCABULARY IN CENTRAL VIETNAMESE<sup>1</sup>

Mark J. ALVES

Montgomery College, USA

<Mark.Alves@montgomerycollege.edu>

### Abstract

This study focuses on grammatical vocabulary in Central Vietnamese<sup>2</sup> (CV, hereafter) with the goal of exploring historical-linguistic developments in Vietnamese in general. One purpose of focusing on grammatical vocabulary is to provide a sample of how this type of vocabulary can innovate over time in different dialectal regions. For research on Central Vietnamese in particular, this focus also allows some identification of historical patterns of lexical preservations (e.g., forms shared with other Vietic languages), innovations or general variation, and the effects of language contact. While CV is quite obviously Vietnamese, there are a number of distinctive traits in its grammatical vocabulary which highlight some of the differentiation over time between CV and standard Vietnamese<sup>3</sup> (StV, hereafter). In a number of instances, data suggests reconstructability at the level of Proto-Việt-Mường or even Proto-Vietic forms. These historical points are dealt with throughout this paper and summarized in the conclusion.

**Keywords:** Vietnamese, Proto-Vietic, grammatical vocabulary

**ISO 639-3 language codes:** vie, mtq, scb

### 1. Central Vietnamese as a dialect

The historical-linguistic position taken in this paper follows that of Nguyễn Tài Cẩn (1979). Vietnamese belongs to the Vietic branch of Mon-Khmer, and within Vietic, there are the major sub-branches Pong-Chứt (examples of languages include Rục, Thà Vưng, and Pong) and Việt-Mường. Within Vietnamese, there are geographic regions where dialects can be distinguished. There are differences among them in pronunciation and lexical items which can be marked by isoglosses on a map of Vietnam.

Based on differences in both the lexicon and phonology, which can significantly impair the ability of Vietnamese from other regions from understanding CV speakers, CV must be considered a dialect, not merely a regional accent. This three-way regional distinction is well-known in Vietnam and has been documented in major linguistic works, such as that of Hoàng Thị Châu (1989: 90-103), as well as numerous shorter works (e.g., Friberg 1973, Vu 1982, etc.). Hoàng (Ibid.) even provided a list of lexical differences among several key grammatical types in the three regions of CV, Northern Vietnamese (NV, hereafter), and Southern Vietnamese (SoV, hereafter). Three areas—Bắc kỳ, Trung kỳ, and Nam Kỳ—and their lexical and phonological distinctions were recognized even in the cultural work of Phan K.B. at the beginning of the twentieth century (2011:422-435). Rather than discussing a distinction between NV, CV, and SoV, Thompson (1965:80-85) describes differences among Vietnamese dialects with major cities as the approximate dialect centers. In this view, Vĩnh and Huế, which are typically both considered to be CV dialects, are treated separately due to phonological distinctions.

<sup>1</sup> This paper is a modified version of the article published in 2011 in *Hội Thảo Quốc Tế Đào Tạo và Nghiên Cứu Ngôn Ngữ Học ở Việt Nam: Những Vấn Đề Lý Luận và Thực Thiện* (International Conference on Linguistics Training and Research in Vietnam: Theoretical and Practical Issues); *Trường Đại Học Khoa Học Xã Hội và Nhân Văn* (Hanoi. The University of Social Sciences and Humanities). ed. by Nguyễn Hồng Côn, et. al.

<sup>2</sup> The following is a complete list of the abbreviations: CV=Central Vietnamese, NV=Northern Vietnamese, NCV=North-Central Vietnamese, SV=Southern Vietnamese, StV=Standard Vietnamese.

<sup>3</sup> The term ‘standard Vietnamese’ is a translation of the Vietnamese term *tiếng chuẩn*, literally ‘standard speech,’ which refers to the generally accepted standard throughout Vietnam. It is generally similar to Northern Vietnamese. However, in some cases, even Northern Vietnamese has phonological and lexical traits which differ from the accepted standard.

While Thompson's approach circumvents deciding on major dialect regions, it also recognizes the fact that the matter is more complex, and that considering a three-way division may not be sufficient. It has also been noted that North-Central Vietnamese (NCV, hereafter), in the region around the city of Vinh, can possibly be considered a sub-region of CV due to some additional lexical and phonological distinctions. Hoàng (Ibid. 261) notes the special grammatical vocabulary in the NCV region and provides a list of key lexical items. Vietnamese spoken in the region of Nghệ Tĩnh and Bình Trị Thiên have been posited by Nguyễn Tài Cẩn in his 1995 historical-linguistic textbook on Vietnamese as being part of a special region of Vietnamese, one with characteristics of both NV and Mường (1995:275-278). The substantial dialectal variety even within small distances in that region, such those described in Alves 2007, suggests that this area is perhaps among the oldest of Vietnamese. It is also worth noting that NCV is a region where Mường and Việt-Chứt languages were spoken. However, more investigation would be needed to strengthen such a position.

## 2. Method of data collection and analysis

For this study, Đặng Thanh Hòa's entire 2005 Vietnamese dialect dictionary Từ Điển Phương Ngữ Tiếng Việt was studied, with grammatical vocabulary identified. The words and phrases were then entered into a spreadsheet with various categories noted, such as region, grammatical function, historical origins when identifiable, and other details. After that, other references were consulted, including Vương Lộc's 2002 dictionary of archaic Vietnamese, Nguyễn Tài Cẩn's 1995 comparative lists of Việt-Chứt language, Nguyễn Văn Lợi's 1995 description of the Rục language, Khang et.al.'s 2002 dictionary of the Mường language, and an unpublished manuscript by Michel Ferlus of reconstructions of Proto-Vietic, including complete lists of available vocabulary.

It is important to understand the method of selecting the vocabulary for this study. Determining which words and expressions are grammatical is not an absolute matter, and one may posit that there are degrees of grammaticality. In general, grammatical vocabulary tends to have more abstract meanings which function more to indicate syntactic relationships and even pragmatic functions, whereas content words, which are not included in this study, are the elements for which grammatical terms provide this semantico-syntactic context. For instance, the possessive marker *của* is clearly grammatical; *hôm nay* 'today' is slightly more semantically concrete but still expresses the grammatical function of time; while *sách* is a clearly content word.

Thus, words which express abstract relationships between nouns, such as locatives, conjunctions, and comparatives, are considered grammatical. Those expressing modality, negation, time, and degree to modify verbs and adjectivals are also treated as grammatical in this study. Quantity expressions and measure words, while not always as abstract as numbers and classifiers, are included in this study due to their generalized, open function and somewhat similar syntactic function to other clearly grammatical words. Pronouns and deictics are non-controversially considered grammatical, but also in this study, so are terms of address, which in Vietnamese have pronominal functions. The following non-exhaustive list is representative of the criteria used to select the vocabulary for this study.

- Pronominal reference (pronouns, deictics, and terms of address)
- Location (prepositions and locative nouns)
- Time and frequency
- Interrogative function
- Negation and modality
- Quantity and units of measurement
- Clause-connecting elements
- Comparison, degree, and intensification
- Pragmatic functions (sentence final particles, exclamations)
- Causation/causative verbs

In the database, over 140 items in CV were identified, of which over 100 are unique to CV and not generally shared with either SoV or NV. Of the 140, 35 are also seen in SoV, while only two are shared with NV. Overall, it can be said that there is a noticeably large number of CV grammatical vocabulary distinct from both NV or SoV, including at least a few crucial items. The significant overlap with SoV grammatical vocabulary, in addition to a number of phonological similarities, suggests a closer historical-linguistic relationship between those two regions than with NV despite the reduced mutual intelligibility between CV and either SoV and NV, a matter anecdotally attested among Vietnamese. More importantly, it must be assumed that the original approximate areas of where modern NV and CV are spoken had diverged a good deal before the migration of Vietnamese to Southern Vietnam. That migration occurred largely over the past

several centuries since the fall of the Champa Empire in the 1400s and resulted in expansion of Vietnamese settlements to modern day Southern Vietnam.

Based on the data collected, the types of CV grammatical vocabulary fall into the categories shown in Table 1. The first two categories are briefly explained in this section and then referred to in subsequent sections, while the third category, unique etyma, constitutes the bulk of rest of the paper. A brief note on archaic Vietnamese preserved in CV is provided in a subsection below.

**Table 1:** Categories of CV Grammatical Vocabulary Relative to StV

1.	Compounds with one cognate in common with StV and one non-cognate morph
2.	Cognates in common with StV but with alternate pronunciations
3.	Completely unique etyma (both archaic Vietnamese words and strictly CV forms)

Instances of mixed cognate morphemes show recurring patterns primarily among adverbial expressions. Among intensifying expressions with *hết* plus another morpheme, the second morpheme may differ in CV and StV, such as CV *hết cỡ* ‘extremely’ versus the synonymous StV item *hết mức*. In several compounds expressing time, CV uses *khi* in contrast with StV *lúc*, such as CV *khi đêm* ‘at night’ versus the StV equivalent *lúc đêm*. Among terms expressing time or quantity, several compounds in both CV and StV contain *một*, while the second morpheme alternates, such as CV *một tẹo* versus StV *một tí*. It is impossible to determine whether these are calques or simple alternates, but regardless, these word formation patterns are clearly part of the linguistic region.

#### A. Patterns of phonological correspondence

As historical phonological patterns between CV and StV are reasonably well described (Nguyễn Đình Hoa 1965, Friberg 1973, Vũ 1982, Hoàng 1989, Nguyễn Tài Cẩn 1995, etc.), it is generally easy to identify cognates among the dialects. Some of these alternations appear among high-frequency or significant CV grammatical words, as listed in Table 2, containing comparative samples of the phonemes shown with Vietnamese orthography. Cognates in the Vietic language Rục are included and highlight the conservative phonological aspects in CV. Some of the changes represent synchronic dialectal differences as well as diachronic changes, though that has been described elsewhere (c.f., Nguyễn Tài Cẩn 1995; Ferlus 1976 and 1981) and is beyond the scope of this paper.

**Table 2:** Common sound correspondences between CV and StV in grammatical vocabulary

CORRESPONDENCES	CV	StV	Vietic	GLOSS
ươ vs. a	đương	đang	NA	progressive marker
ây vs. ai	cây	cái	Rục <i>ke</i> <sup>4</sup>	general classifier
-ng vs. -nh	eng	anh	NA	term of address (male, older, respectful)
i vs. ay	mì ni	mày này	Rục <i>mi</i> <sup>1</sup>	you (informal) this
ch vs. gi	chi	gì	NA	what
tr vs. gi	trưa	giữa	Rục <i>troah</i> <sup>2</sup>	between
ui vs. ôi	tui	tôi	NA	I
r vs. nh	rau	nhau	NA	reciprocal marker

Of the 140 plus items in the database, only about 20 entries fall into this category, with about ten additional forms which cannot be considered with high certainty to be phonologically related cognates. Still, considering the high grammatical functionality of some of these words, those phonological differences, combined with the many other distinct etyma, certainly contribute to decreased mutual comprehensibility between CV and speakers in other regions. CV *mi* ‘you (informal)’ for StV *mày* and CV *ni* ‘this’ for StV *này* are quite similar to Mường and Việt-Chứt languages, such as Rục or Thà Vượng (Suwilai 1998). A few other highly functional words which are cognates with different pronunciations include SoV *chi* versus StV *gì* ‘what,’ CV *trưa* versus StV *giữa* ‘between/the middle,’ CV *tui* versus StV *tôi* ‘I,’ and the reciprocal in CV *rau* versus StV *nhau*. Many of these represent competing sound changes.

### B. Archaic words

Throughout this paper, CV words which are related to Mường or Pong-Chứt languages are noted, but in addition to these kinds of lexical preservations, there are various Vietnamese words which are not part of modern StV but which are listed in Vương Lộc's 2002 dictionary of archaic Vietnamese words. These words and phrases are shown in Table 3. How frequently these are still used in parts of Central Vietnam and where they are used in that region may vary and would require additional research to clarify. Nevertheless, this overall list highlights some of the more archaic parts of the CV lexicon (archaic in that they are no longer used, though few have cognates with other Vietic languages), even in the restricted domain of grammatical vocabulary. This is not to say that NV and SoV do not have vocabulary which is preserved from early stages of Vietnamese, but in addition to the other older Vietic vocabulary, this data is worth considering in exploring the history of Vietnamese.

**Table 3:** Archaic Grammatical Vocabulary in CV No Longer Used in StV

CV	StV	GLOSS
bao lăm	bao nhiêu	how much (implied negative)
bậu	(từ nam giới dùng để gọi vợ)	term used by husband to refer to wife lovingly
in	y	exactly as
lung	nhiều, quá mức	a lot
mụ	bà	grandmother
năng	hay, thường	often
ngghi	nó, hấn	he/she
nhủ	bảo	make/command
phải chi	giả như, giả mà	supposing that
phương chi	huống gì	much less/not to say
rày/rây	nay	at this time
rồi	xong	finished

### 3. Types of Central Vietnamese grammatical vocabulary

As noted in the previous section, the words were selected based on their grammatical functions. This casts a wide net but ultimately provides a more complete picture of the range of grammaticalized vocabulary in CV. The rest of this section covers nine types of grammatical vocabulary.

#### A. Interrogative words

Table 2 above shows phonological distinctions of CV, with cognates in the Vietic language Rục, in grammatical vocabulary. Beyond phonological distinctions, CV is well known for having several distinctive interrogative words, separating it from both StV and SoV. These include CV *mô* 'where' versus StV *đâu*, CV *răng*<sup>4</sup> 'how' versus StV *sao*, CV *chi* 'what' versus StV *gì*. The latter form *chi* is pronounced essentially the same as in Mường, an apparent older Vietic pronunciation, at least to proto-Việt-Mường. CV *mô* is notable as having cognates in several other Pong-Chứt languages. While not an interrogative word (though it is a causative verb having a relatively grammatical function), CV *mần* 'do' versus StV *làm* is combined with CV *răng* to express 'how'. *Mần* is also a Việt-Chứt cognate seen in Rục, suggesting a possible Proto-Vietic form.

**Table 4:** Interrogative words in various Vietic languages

StV	CV	Rục	Mường	Gloss
đâu	mô	tumô	no	where (or which)
gì	chi	chamo <sup>3</sup>	chi	what
làm	mần	mun <sup>2</sup>	là	make/do (used in interrogative compounds)

This sort of distinction between CV and StV is among the most persuasive data, showing that CV and StV, or NV to be precise, diverged as dialects. Moreover, in some cases, such differences also illustrate lexical and phonological preservations that highlight the closer relationship between CV and other Vietic

<sup>4</sup> It is also worth noting that, in this study, several CV grammatical words with deictic reference were found to have initial /r/. These include *răng* 'how', *ri/rúa* 'thusly', and *rày/rây* 'at this time'. This is not unlike patterns in basic demonstratives noted in StV by Thompson in which such words systematically have initial /d/, /n/, and /b/ (1965:142).

languages. One might even hypothesize that NV diverged from earlier Vietic, as more closely represented by CV.

### B. Negation and Modals

CV has a number of distinct negation and words expressing modality. In some cases, the difference is merely in pronunciation, as in items 3 and 4 in Table 5. Of the negation words, CV *nỏ* ‘not,’ which is more of a NCV form than CV in general, is an etymon worthy of attention. Whether it is a possible CV innovation is not clear. Cognates of the item have not been found in other Vietic languages. Other items on the list are interesting, though it is not clear based on the dictionary how common the other items are in CV.

**Table 5:** Negation words and modals

	CV	StV	GLOSS
1.	nỏ	không, chẳng	no, not
2.	dòn	khỏi	no need
3.	làm chi	làm gì	nothing, not
4.	đương	đang	(progressive marker)
5.	rồi	xong	finished
6.	thầu	nổi, hám ý phủ định	able
7.	lựa chi	thà rằng	had better

### C. Pronouns and Terms of Address

The number of differences, both phonological and etymological, among pronouns in CV clearly highlights the distinct dialect status of CV. Items 1 through 4 are cognates with StV forms, though they differ phonologically (c.f., section 3A). The pronoun pluralizer *bầy* may be related to StV *mấy*. A likely cognate is seen in Ruc, the pluralizer *bợl*<sup>5</sup>. Items 9 through 13 are all entirely distinct from StV. At this point, they are of uncertain etymological origin, without immediately recognizable cognates among Mường or Việt-Chứt languages.

**Table 6:** Pronouns

	CV	StV	Gloss
1.	mi	mày	you (informal)
2.	ni	này	this
3.	tau	tao	first-person singular, informal
4.	tui	tôi	I
5.	bay/bây	chúng mày	you, plural
6.	bầy	bọn/tụi	plural marker
7.	bầy choa	bọn/tụi tôi/tao	we
8.	bầy tui	bọn/tụi tôi	we
9.	nghi	nó, hần	he/she
10.	choa	chúng ta	we, I
11.	ri	thế này	thusly
12.	rúa	thế	thusly
13.	nó	ấy	that

As for terms of address, often derived from familial terms, again, CV shows a number of significant differences. Worthy of note are the various forms which are not of Chinese origin, unlike in StV. CV *ả* ‘elder sister’ shows a contrast with the likely Old-Sino-Vietnamese *chị* (standard Sino-Vietnamese *tỷ* 姐 (Mandarin *jǐe*)), *mệ* and *mự*<sup>5</sup> ‘grandmother’ for the Chinese loanword *bà* 婆 (Mandarin *pó*), and CV *o* ‘sister of father’ in contrast with StV *cô*, also of Chinese origin 姑 (Mandarin *gū*). Are these preservations of older Vietic, innovations in CV, or something else? The latter item CV *o* in particular has a cognate in Ruc, suggesting

<sup>5</sup> Admittedly, Chinese has phonologically similar 母 ‘mother,’ standard Sino-Vietnamese *mẫu*. However, having other forms meaning ‘mother’ (rather than ‘grandmother’), including StV *mẹ* and SV *má*, altogether make clear identification of words with phonological differences, despite having initial /m/, problematic.



that this is possibly older in Vietic. Vương 2002 claims that CV *bậu* has a cognate in Mường *pầu*, which means ‘they/those people,’ though the semantic difference is significant and cannot be easily explained. It is a possible older Việt-Mường word, but the evidence must be considered tentative at this point. Vương also notes a possible Pong-Chứt cognate for CV *tam*, namely, Thà vưng *saam*, which matches a tendency seen in Sino-Vietnamese /t/ for Middle Chinese /\*s/. Clearly, these items are worth investigating further as they could shed light on what might not be Chinese elements in Vietnamese and reflect a pre-Chinese influenced era.

**Table 7:** CV Terms of Address

CV	StV	GLOSS
ả	chị	sister, elder
bậu	(từ nam giới dùng để gọi vợ)	(term used by husband to refer to wife lovingly)
bọ	cha	father
eng	anh	(term of address for somewhat older man than the speaker)
mệ	bà	grandmother
mụ	bà	grandmother
o	cô	sister of father
tam	em	(term to address younger person)

#### D. Locative and Deictic Terms

CV has a few distinct etyma for locative words and deictics, in addition to those with phonologically related cognates, as shown in Table 8. These forms are fairly conservative and appear to be early Vietic words or have more archaic pronunciations. The initial /tr/ in CV *trữa* ‘between’ versus the palatalized StV *giữa* form shows a conservative tendency in sound (cf. the cognate in the Vietic language Rục *troah*<sup>2</sup>). Similarly, CV *nì* ‘this’ versus StV *này* are essentially the pronunciation in Mường and Pong-Chứt languages. Finally, the CV locative *côi* ‘top’ versus StV *trên*, the latter of which is spread throughout Vietic languages, is noteworthy for having a likely cognate in Mường, synonymous *cỏi* (though the tone is readily explainable), showing at least a potential proto-Việt-Mường form. There is also an interesting parallel in the tone change from *ngang* to *huyền* in the apparent non-cognates for the two distals indicating ‘over there’ and ‘over there, farther.’ This is an interesting tonal, morphological calque, though which form is the earlier cannot be determined at this point. Altogether, these grammatical words again highlight the distinctiveness CV has as a dialect.

**Table 8:** Locative and Deictic Terms

CV	StV	Gloss
trữa	giữa	between
côi	trên	top
tê	kia	over there
tề	kìa	over there, farther
đàng tê	đàng kia	over there
đàng tê tề	đàng kia kìa	over there, farther
đó tề	đó kia	over there
dựa	kề, cạnh	next to
nì	này	this

#### E. Measure Words

Measure words are not as grammaticalized as what are considered true classifiers, yet they have the function of serving as complements of numerals and taking other nouns as complements, a relatively grammatical function. There are several distinctive CV measure words, as listed in Table 9. CV *âu* is also seen in Mường. CV has *rê*, while StV *bánh* is originally of Chinese origin (standard Sino-Vietnamese *bính*, Chinese 饼, Mandarin *bǐng*). Similarly, the CV measure word for bread *ố* parallels the function of the StV Old Chinese loanword *chiếc* (Sino-Vietnamese *chích*, 只, Mandarin *zhī*). These forms have clearly undergone semantic extension and grammaticalization. More research needs to be done to identify the etymological origins of these words and trace their historical relationships. Nevertheless, these examples highlight some differential patterns in this category.

**Table 9:** CV Measure Words

CV	StV	Gloss
âu	liễn	pot (terracotta) of liquid
vây	(lượng cỏ/rác...)	a quantity of grass, trass, etc.
ngoai	vòng, bận	a round of an activity
ô	chiếc	a loaf of bread, etc.
rê	bánh	a cake of tobacco

**F. Quantity and Intensifier Words and Expressions**

Basic numerals are the same in CV and StV, with the phonological exception of initial /l/ in NV *năm* ‘five’ in numeral compounds (e.g., ‘25’ *hai mươi lăm*). However, in expressing degree and generalized quantity, CV shows a number of differences from StV, as listed in Table 10. The etymological sources of these are not immediately identifiable. One interesting word is the CV word *lố* for ‘dozen,’ in contrast with the Sino-Vietnamese cognate in StV *tá*, Chinese 打 (Mandarin *dǎ*). It may be a cognate with the Hokkien word for dozen, 罗 *lô*, which has also been borrowed by Cambodian, Thai, and Laotian (Pou and Jenner, 77).

**Table 10:** Distinct CV intensifying words and expressions

CV	StV	GLOSS
băm	nhiều, lắm	very, much
bộn	nhiều	much
ét	đơn, chỉ có một	only one
lên	đầy, nhiều	fully, a lot
lố	tá	a dozen
loạn	lắm, nhiều	a lot
mê thiên	rất nhiều	very many
rền	đầy, nhiều	fully, a lot
ba hột	chút ít, không đáng kể	very little
sáp	bọn, tụi	a group of X

Some intensifying expressions are partial cognates, including those with *một*, *hết*, and *thấy* (with the additional element referring to one’s grandmother). These pattern in an almost prefix-like, calque-like way. Other words in the list are completely distinct etyma, though their specific origin is not yet clear.

**Table 11:** CV intensifying words and expressions with mixed etyma

CV	StV	GLOSS
hết cỡ	hết mức	excessively
một tẹo	một tí	a little
một thỉ	một tí	a little
thấy mụ nội	thấy bà	terribly
nhứt	nhất	most
dễ sợ	cực kì	extremely
đui	rất, khi	very
gắt	ở mức độ rất cao	extremely
lung	nhiều, quá mức	a lot
òm	quá, hết mức	extremely
bất	rất, quá	very

Some additional adverbs include those expressing approximation, as shown in Table 12. It is notable that the CV approximative *cỡ* is homophonous with the word meaning ‘size,’ which is parallel with StV approximative words *khoảng* and *độ*, both of which have homophonous meanings of a unit of time or place and height or period of time, respectively. This is suggestive of a pattern of grammaticalization in which a

unit of measurement becomes a generalizing expression (c.f., English approximative expression ‘in the *area* of X’).

**Table 12:** CV adverbs expressing approximation

CV	StV	GLOSS
cỡ	khoảng	about
cùng	khắp cả, hết tất cả	everywhere completely
lối	chừng, khoảng chừng	about, approximately
lối chừng	chừng, khoảng chừng	about, approximately

### G. Time Words and Expressions

In this category, many of the words are compounds, and of many of those compounds, there are many instances of shared cognates, as shown in Table 13. For example, CV *khi* parallels StV *lúc* and indicates time in many compounds. The other morphs in the compounds have the same meaning in many instances, such as CV *khi sớm* ‘while early’ versus StV *lúc sớm*.

**Table 13:** CV *Khi* versus StV *Lúc*

CV	StV	GLOSS
khi đầu	lúc đầu	at the beginning
khi đêm	lúc đêm	at night
khi hồi	lúc này	just now
khi ni	lúc này	at this time
khi nớ	lúc ấy	at that time
khi sớm	lúc sớm	while early
khi tê	lúc trước	previously

One morph of historical note is StV *giờ*, which has been posited (Pulleyblank, 283) as an Old Sino-Vietnamese loanword (standard Vietnamese *thì*, Chinese 时, Mandarin *shí*) originally meaning ‘time’ and which has come to mean in Vietnamese ‘hour’ or more generally ‘at this time.’ CV has the clearly related cognate *giữ*, but also the likely cognate *chừ*. The two readings *giữ* and *chừ* appear to be phonological alternants in CV, but are still related to the StV form of the same etymon. The likely older form is *chừ* considering its non-palatalized initial, which, if accurate, would not be an insignificant fact in understanding Sino-Vietnamese historical linguistics.

**Table 14:** CV words meaning ‘now’

CV	StV	GLOSS
bây chừ	bây giờ	now
bây giữ	bây giờ	now
giữ	giờ	now

The large number of time words and expressions in CV which differ from those in StV clearly draws attention. Table 15 shows this substantial list, some of which involves a contrast between CV *bữa* and StV *hôm* in several compounds.

**Table 15:** Various CV time words and expressions

CV	StV	GLOSS
bữa điếp	hôm kia	day before yesterday
bữa ni	ngày hôm nay	today
bữa rày	bữa hôm ấy đến nay	from that day to today
bữa tê	hôm kia	day before yesterday
giác	độ, khoảng, lúc	around this time
hồi tê	hồi sau, lúc sau	at a previous time
mại lồi	ngày trước nói khái quát	a previous day, non-specific
mấy hồi	chẳng mấy lúc	not long
mấy nă	chẳng được mấy chốc	not even a moment
mốt	ngày kia	two days later
xém	súyt	nearly, about to
rày	nay	at this time

CV also shows a number of distinct time adverbs of frequency. Notable among these are some reduplicants, such as CV *hoài hủi* and *lần lần*, both of which parallel reduplication in StV. As noted in section 3B above, CV *năng* is an archaic form of Vietnamese.

**Table 16:** Time adverbs of frequency

CV	StV	GLOSS
lút	tuốt, thẳng một mạch	continuously
thét	mãi	continuously
hoài hủi	mãi, luôn luôn	continuously, forever
lần	dần	gradually
lần lần	dần dần	gradually
năng	hay, thường	often

### H. Conjunctions

Among connective words in CV are some phonological differences and some compounds with CV etyma, such as *răng* and *rứa*, as noted above. As seen in other cases, these compounds appear almost like calques, though it is impossible to determine whether CV or NV first developed the terms and which variety calqued which originally. Regardless, the semantic patterns in these words are suggestive of a kind of linguistic area.

**Table 17:** CV Connective Words

CV	StV	GLOSS
dù răng	dù sao	although
dưng	nhưng mà	but
bởi rứa	bởi thế	thus
giả thí như	ví dụ như	for example
hèn chi	hèn gì, hèn nào	that is why
nên chi	vì vậy, cho nên	thus
phải chi	giả như, giả mà	supposing that
phương chi	huống gì	much less

### I. Sentence Particles

The dialect dictionary contains at least ten CV sentence particles, though it is difficult to determine whether some of these are cognates with phonological variants or distinct etyma. It is tempting, for instance, to consider CV *nà*, an emphatic particle, to be related to StV *nào*, and the CV question particle *hỉ* appears phonologically similar to StV *hả*. However, neither of these forms fall under a common pattern of phonological correspondences between CV and StV. Sentence particles are semantically highly abstract and less readily perceived phonetically. It is thus, perhaps, not surprising that many of the dialectal forms, while



having similar functions, vary phonologically but only slightly. Consider the semantic and phonological similarities between CV particles *nghe* and *nha* with StV *nhé*, all of which express urging, though again without phonological patterns to clearly related them.

**Table 18:** CV sentence particles

CV	StV	GLOSS
hấy	hả	emphatic question
hầy	nhỉ	emphatic question
hỉ <sub>1</sub>	nhỉ	emphatic question
hỉ <sub>2</sub>	hả	question
hè	nào	urging
nà	nào	urging
nghe	nhé	urging
nha	nhé	urging
nì	này	urging
nờ	này, nào	urging

It is perhaps best to assume that there are various cognates among these words, but that due to their more peripheral nature in the lexicon, being nearly devoid of concrete meaning and most often occurring in sentence-final position, it is often the case that cognates cannot be claimed with certainty. It may be more realistic to posit a loose or even phonaestheme-like element for this category of words.

#### 4. Concluding Thoughts

Despite the focus on grammatical vocabulary, the data repeatedly demonstrate how CV is distinct from StV as a dialect, thereby suggesting how it is a distinct dialect region historically. It also shows in many instances—both phonologically and lexically—how close CV is to the older stages of Việt-Mường and Vietic in Vietnamese linguistic history. At the same time, the data are illustrative in various ways of a kind of linguistic area. In numerous instances, compounds with mixed cognates are suggestive of how CV and StV parallel each other in word-formation strategies. Indeed, in some cases, one may posit that these word-formation strategies are suggestive of derivational morphology, a counter-argument to the position that Vietnamese has strictly isolating morphology. This small study illustrates how grammatical vocabulary can serve to clarify the relationships among dialects.

#### References

- Alves, Mark J. 2007. A Look at North-Central Vietnamese. Electronic Publication E-4. SEALS XII Papers from the 12<sup>th</sup> Annual Meeting of the Southeast Asian Linguistics Society (2002), 1-8. ed. Ratree Wayland, John Hartmann, and Paul Sidwell. Pacific Linguistics Research School of Pacific and Asian Studies. The Australian National University.
- Alves, Mark and Nguyễn Duy Hương. 2007. Notes on Thanh-Chương Vietnamese in Nghệ-an Province. Electronic Publication E-2. SEALS VIII Papers from the 8<sup>th</sup> Annual Meeting of the Southeast Asian Linguistics Society (1998), 1-10. ed. Mark Alves, Paul Sidwell, and David Gil. Pacific Linguistics Research School of Pacific and Asian Studies. The Australian National University.
- Đặng, Thanh Hòa. 2005. Từ Điển Phương Ngữ Tiếng Việt. Nhà Xuất Bản Đà Nẵng. Trung Tâm Từ Điển Học.
- Ferlus, Michel. 1976. Du nouveau sur la spirantisation ancienne en Vietnamien. Bulletin de la Société de Linguistique de Paris 71:305-312.
- . 1981. Sự biến hóa của các âm tắc giữa (obstruents mediales) trong tiếng Việt (Changes of medial obstruents in Vietnamese). Ngôn Ngữ Học 1981.2:1-21.
- Friberg, Barbara. 1973. Generative phonology as applied to Vietnamese dialects: a study based on middle Vietnamese, comparing the three major dialects of modern Vietnamese. Saigon University MA thesis.
- Hoàng, Thị Châu. 1989. Tiếng Việt trên các Miền Đất Nước (Phương Ngữ Học). Nhà Xuất Bản Khoa Học Xã Hội – Hà Nội.
- Nguyễn, Đình Hoà. 1966. Vietnamese-English Dictionary. Rutland, Vermont: Charles E. Tuttle Co.

- Nguyễn, Tài Cẩn. 1995. Giáo trình lịch sử ngữ âm tiếng Việt (Textbook of Vietnamese historical phonology). Hà Nội: Nhà Xuất Bản Giáo Dục.
- Nguyễn, Văn Khang, Bùi Chỉ, and Hoàng Văn Hành. 2002. Từ Điển Tiếng Mường-Việt. Nhà Xuất Bản Văn Hóa Dân Tộc – Hà Nội.
- Nguyễn, Văn Lợi. 1993. Tiếng Rục (The Ruc Language). Hà Nội: Nhà Xuất Bản Khoa Học Xã Hội.
- Phan, Kế Bính. 2011 (reprint of the 1915 publication). Việt Nam Phong Tục. Nhà Xuất Bản Văn Học – Hà Nội.
- Pou, Saveros and Philip Jenner. 1973. Journal of Oriental Studies 6.1: 1-90. Hong Kong University Press.
- Pulleyblank, Edwin G. 1981. Some notes on Chinese historical phonology. Bulletin de l'école Française d'Extreme-Orient 277-288.
- Suwilai, Premsrirat. 1998. So (Thavung)-English-Thai glossary Part I. Mon-Khmer Studies 28:189-218.
- Thompson, Laurence. 1965. A Vietnamese grammar. Seattle: University of Washington Press.
- Vũ, Thanh Phương. 1982. Phonetic properties of Vietnamese tones across dialects. Papers in Southeast Asian Linguistics, No. 8:55-76, Tonation. Pacific Linguistics A-62. Ed. David Bradley. Canberra: Australian National University.
- Vương, Lộc. 2002. Từ Điển Từ Cổ (A Dictionary of Ancient Words, 2<sup>nd</sup>. Ed.). Hà Nội: Nhà Xuất Bản Đà Nẵng, Trung Tâm Từ Điển Học.



# A TONAL ANALYSIS OF CONTEMPORARY TAI KHUEN VARIETIES<sup>1</sup>

R. Wyn OWEN

*Payap University, Chiang Mai*

<[wyn\\_owen@sil.org](mailto:wyn_owen@sil.org)>

## Abstract

This paper presents an analysis of the tones of five contemporary Tai Khuen varieties in order to investigate the differences in the number of distinctive tones reported in the literature. The present study shows that while some contemporary speakers have a tone system with six tones, most speakers have only five tones. Comparison of the distribution and phonetic characteristics of the tones in the contemporary varieties with previous studies shows that the five tone system was derived by the coalescence of two tones in the six-tone system. Investigating the factors that determine which tone system a particular contemporary speaker uses leads to the conclusion that language contact with five-tone Shan was the cause of the change.

**Keywords:** Phonetics; Tones; Language contact

**ISO 639-3 Language codes:** shn, kkh

## 1 Introduction

Understanding the tone system of any particular language is important not only in its own right but for cross-linguistic purposes as well. For example, tonal distributional patterns have been used as key criteria in the classification of Southwestern Tai (SWT) languages (Chamberlain 1975:50; Robinson 1994:141; Edmondson 2008:194-195). In the literature on Khuen phonology, different authors writing at different times have reported different numbers of distinctive tones. This means that in order to fully understand Khuen tones, there is a need for a synchronic description of the tone systems of Khuen varieties from various locations. Such a description can form a basis for investigating the origin of the differences reported in the literature. Furthermore such a description is an important element in evaluating the adequacy of the 700-year old Khuen writing system for representing contemporary spoken Khuen. This study is therefore based on the analysis of contemporary data from Khuen speakers from various Khuen communities. The distributional patterns of the tones will be established and the phonetic characteristics of each tone will be described in detail as a basis for comparison between different Khuen varieties.

The remainder of this section provides a brief introduction to Khuen culture and language and in Section 2 previous research on Khuen tones is summarised and compared and calibrated with Pittayaporn's (2009a:267-282) reconstruction of the tonal history of Tai languages. Sections 3 and 4 respectively describe the methodology and present the findings of the present study. Section 5 presents some implications of the work and makes suggestions for further research.

### 1.1 Khuen culture

Khuen (k<sup>h</sup>u̯n) is a Tai language spoken primarily in the Keng Tung basin of Eastern Shan State, Myanmar where there are estimated to be 115,000 Khuen (Lewis 2009). Lewis further estimates a population of 6,280 in Khuen communities in Northern Thailand. Shliessinger (2003:90) reported one Khuen village of around 600 inhabitants in Laos.

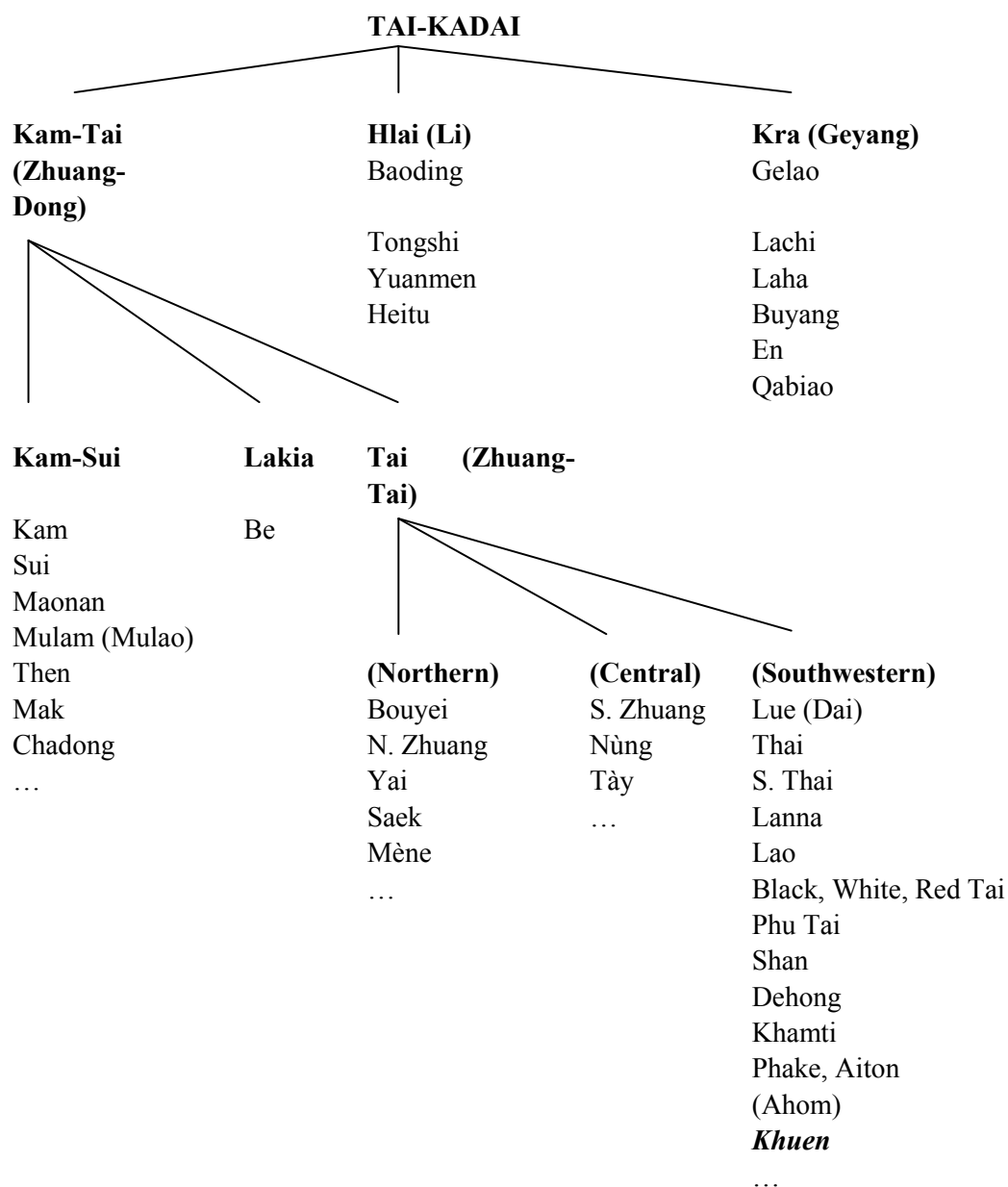
The Khuen are valley-dwelling wet rice growers — one of many common characteristics of Tai groups (Lebar et al. 1964:187). Khuen culture has been strongly influenced by that of the Lan Na kingdom founded by Mang Rai in the late 13<sup>th</sup> Century A.D. and based in Chiang Mai (Wyatt 2003:70). The adoption of the Lan Na script coupled with the prominent role of the Buddhist clergy in the intellectual life of the community were strong factors in support of Khuen language use and literacy (Penth 1994:13; Peltier

<sup>1</sup> I am grateful to Dr. Sigrid Lew and the anonymous referees for comments on an earlier version of this paper. Any remaining errors are entirely my own.

1993:20). Khuen literacy continues to be strongly promoted by Buddhist temples and the Khuen Cultural and Literature Committee based in Keng Tung (Peltier 1993:21; Owen 2008:5).

## 1.2 Khuen language

Khuen belongs to the Southwestern branch of the Tai language family (Li 1964:7; Pittayaporn 2009a:301). The linguistic classification of Khuen as Southwestern Tai (SWT) is not disputed in the literature although various models have been proposed for subdividing SWT (for example Chamberlain 1975:50; Robinson 1994:141; Pittayaporn 2007a:10-11). The classification tree in Figure 1 is adapted from Diller (2008:7) and places<sup>2</sup> Khuen in the broader context of the Tai-Kadai languages.



**Figure 1:** Classification tree of the Tai-Kadai languages (adapted from Diller 2008:7)

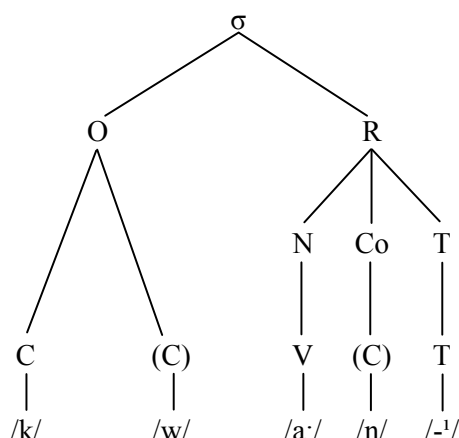
Tai languages in general display many common characteristics. Basic constituent order is Subject-Verb-Object (see Morey 2005:288-303 for a good discussion of this). In common with many other languages in Southeast Asia the Tai languages are isolating in structure and a high proportion of morphemes are monosyllabic (Morey 2005:208). Nevertheless Pittayaporn (2009a:40) argues that a sesquisyllabic reconstruction for Proto-Tai best accounts for the variation in comparative data. Many polysyllabic words in modern varieties are loan words. For Khuen in particular, Petsuk (1978:26) asserts that most of the words with more than one syllable are loan words from Burmese, Pali, Sanskrit and English. Notwithstanding, there

<sup>2</sup> Khuen was not explicitly included in the original diagram, being one of those languages denoted by '...'.



are many polysyllabic words in Khuen which are not loan words. Just as Hudak (1990:38-40) described for Standard Thai, Petsuk (1978:56-67) showed that affixation, compounding and reduplication are productive derivational processes in Khuen.

The diagram in Figure 2 represents the Khuen syllable types described in Petsuk (1978:26-29). As can be seen from the parentheses, the second consonant in the onset is optional, as is the consonant coda. The onset and rime on the other hand are compulsory. The minimal syllable consists of a consonant onset followed by a vowel nucleus with a distinctive tone. The maximal syllable consists of a consonant cluster onset followed by a vowel nucleus and with distinctive tone and a consonant coda. Initial consonant clusters are limited to /kw/ and /k<sup>hw</sup>/ in natural Khuen words although various other clusters occur in loan words from Burmese and Shan.



**Figure 2:** Khuen syllable structure (based on Petsuk 1978:26-29)

As shown for Khuen in Figure 2, Tai languages in general exhibit lexical tone, that is, the pitch of the word can change the meaning of a word (Yip 2002:1). In fact some additional laryngeal feature such as glottalisation or creakiness is often found to play a crucial role in making a particular tone distinctive (Morey 2005:113). Patterns of tonal distribution are one of the most important features in comparing Tai varieties (Gedney 1972:423). The standard framework for analysing and describing tone systems in Tai languages is the Gedney tone box framework as set out in Figure 3 (Gedney 1972). The framework is based on the historical reconstruction of Proto-Tai (PT) and consists of the cross tabulation of four categories of PT initial consonants with five categories of PT tones resulting in 20 separate tone boxes representing mutually exclusive environments.

		PT tones				
		A	B	C	D-Short	D-Long
Initials at time of tonal splits	Voiceless friction sounds <i>*s, *hm, *ph, etc.</i>	A1	B1	C1	DS1	DL1
	Voiceless unaspirated stops <i>*p, etc.</i>	A2	B2	C2	DS2	DL2
	Glottal <i>*ʔ, *ʔb, etc.</i>	A3	B3	C3	DS3	DL3
	Voiced <i>*b, *m, *l, *z, etc.</i>	A4	B4	C4	DS4	DL4
		Smooth <sup>3</sup> syllables			Checked <sup>4</sup> syllables	

**Figure 3:** Gedney's tone box framework (adapted from Gedney 1972:434)





Although the Gedney box framework provides 20 separate tone boxes, in general a particular Tai language will have between 4 and 7 distinctive tones (Morey 2005:115). In determining the number of tones for a particular variety the convention is to number each distinctive tone on smooth syllables but to associate the tones on checked syllables with the smooth syllable tone to which it is most similar phonetically.

<sup>3</sup> Also known as 'live syllable', this is a syllable ending in a voiced sound, i.e. a vowel, semi-vowel or nasal.

<sup>4</sup> Also known as 'dead syllable' this is a syllable ending in one of the voiceless stops *p, t, k*, or *ʔ*.

Pittayaporn (2009a: 267-282) reconstructed the PT tones as register complexes comprising pitch height, pitch contour, vowel duration and voice quality. The phonetic characteristics of the tones are described in Table 1.

**Table 1:** Phonetic Characteristics of PT Tones (adapted from Pittayaporn 2009a:271)

	*A	*B	*C	*D
Pitch Height	Mid	Low	High	Low
Contour	Level	Rising	Falling	Rising
Vowel Duration		Long	Short	
Voice Quality	Modal	Creaky	Glottal Constriction	Creaky <sup>5</sup>
				

It is generally accepted that some time after PT diversified into various daughter languages tone splits occurred, conditioned primarily by the laryngeal features of the initial consonants (see Pittayaporn 2009a:24). All modern varieties display evidence of a split between onsets that at the time of the splits were voiced and onsets that were voiceless, but aspiration, glottalisation and frication have caused additional splits in some languages (Pittayaporn 2009:241). Pittayaporn (2007b:1421-1422) proposed three mechanisms for tone change. These are briefly summarised in Table 2.

**Table 2:** Mechanisms for Tone Change (Pittayaporn 2007b:1421-1422)

No.	Change Mechanism	Description of Change Mechanism
1.	Segment-tone interaction	initial pitch is normally lower after a voiced onset, but some languages have higher initial pitch after a voiced onset; the pitch at the end of a syllable is not affected by the interaction between the initial consonant and the initial pitch.
2.	Contextual variation	the phonetic realisation of a tone varies with the phonetic and phonological context. There are two kinds of variation: contour reduction and peak sliding.
a.	Contour reduction	Dynamic tones are realised with reduced contours. The tonal onset is not affected by contour reduction
b.	Peak Sliding	It is more likely for the peak of a tonal contour to move rightward than leftward. The tonal onset remains the same, but the peak is delayed.
3.	Perceptual maximisation	Tonal contours may change to become more easily distinguished from other tones in the same system.

These mechanisms can be used to explain how modern tonal reflexes develop from the PT tones. In the present context they will be used in Section 2.2 to sketch a possible pattern of development of modern Khuen tones from the PT tones.

## 2. Synthesis of previous research

In this section previous research on Khuen tones is presented to provide both the context and motivation for the present study. The map in Figure 4 shows selected village tracts<sup>6</sup> and specific village locations around Keng Tung which is the heartland of Khuen language use and therefore the main area of focus for this study.

<sup>5</sup> In Pittayaporn's original table (Table 6-20, p271) the voice quality of \*D is not stated, but it is explicitly described as creaky on p277.

<sup>6</sup> In Myanmar a village tract (or 'village cluster') is an administrative unit immediately below the township and above the village. The number of villages in a particular village tract varies depending on the population: the mean number of villages per village tract is 4.7.



**Figure 4:** Map of selected village tracts around Keng Tung

The shading on the map shows two areas that have been associated with different varieties of Khuen in the studies described in Section 2.1. The Kang Murng<sup>7</sup> area to the north west of Keng Tung city is comprised of four village tracts namely Doi Long, Wat Sao, Kat Htai and Kat Tao. By contrast the Murng Lang village tract lies to the east of Keng Tung city. The village tracts of Kat Fah to the northeast and Yang Keng to the north are also shown because the home villages of some of the informants in this study lie in these village tracts. The map also shows that the Kang Murng area as well as the village tracts of Yang Keng and Kat Fah are better served by the main roads from Keng Tung than is Murng Lang. This implies that Murng Lang inhabitants are less likely to come into contact with non Khuen speakers than are people living in the other areas.

The following section presents an overview of previous research on Khuen tones where the focus will be on the number of tones and the patterns of distribution. Details of the phonological characteristics of the tones will be given in the discussion in Section 2.2.

### **2.1 Overview of Khuen tone research**

Egerød described Khuen phonology and script but did not identify his phonological description with any particular location or variety other than stating that his research ‘commenced with a stay in Kengtung in 1957’ (1959:123). He identified six distinctive tones and described their distribution which is laid out in Figure 5.

<sup>7</sup> The Romanised spelling of place names in this paper follows local conventions where these exist and otherwise attempts to represent Khuen phonetic pronunciation rather than historical derivation. Gedney ([1964] 1994:977) refers to ‘Klang Muong Khuen’ and Petsuk (1978:2) refers to ‘Khün Kang Muang’.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	5	2	3
2. Voiceless unaspirated stops					
3. Glottal	2	4	6	6	4
4. Voiced					

**Figure 5:** Tone distribution in Keng Tung Khuen from Egerød (1959:127)

In 1964 Gedney collected extensive field notes on two varieties of Khuen ('Klang Muong Khuen' and 'Baan Veng Khuen') although these were not published for another 30 years (Gedney [1964] 1994a; Gedney [1964] 1994b). The informant for Klang Muong Khuen was born in Keng Tung whereas the informant for Baan Veng Khuen was from the village of Baan Veng<sup>8</sup>, which lies about 5 miles east of Keng Tung. The tonal systems and realisations of individual tones were the same for the two varieties. Gedney identified six distinctive tones<sup>9</sup> whose distribution is laid out in Figure 6.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	5	2	3
2. Voiceless unaspirated stops					
3. Glottal	2	4	6	6	4
4. Voiced					

**Figure 6:** Tone distribution in Khuen varieties from Gedney ([1964] 1994a; [1964] 1994b)

Petsuk (1978:2) asserted that there were three varieties of Khuen, namely, 'Khün Kang Muang', 'Khün Muang Lang' and 'Weng Khün'. She described in detail the phonology of 'Khün Kang Muang'. The principal informant for her study was from Ban Hai<sup>10</sup>, a village about 5 kilometres from Keng Tung in the Kat Htai village tract (Petsuk 1978:4). Petsuk identified five distinctive tones whose distribution is laid out in Figure 7. The most notable feature of this distribution is the lack of tone split in the B and DL columns. Also notable is that Tone 2 exhibits tonal sandhi: the selection of one of its two phonetic variants is primarily determined by the pitch height of the preceding syllable.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	4	2	3
2. Voiceless unaspirated stops					
3. Glottal	2		5	5	
4. Voiced					

**Figure 7:** Tone distribution in Kang Murng Khuen from Petsuk (1978:109)

In addition to the foregoing studies of Khuen around Keng Tung, there are two studies of Khuen communities in Thailand. Akharawatthanakun (2012) studied Khuen along with four other Tai languages spoken in Nan Province. She identified six distinctive tones whose distribution is laid out in Figure 8.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	5	2	3
2. Voiceless unaspirated stops					
3. Glottal	2	4	6	5	4
4. Voiced					

**Figure 8:** Tone distribution in Khuen of Nan Province from Akharawatthanakun (2012)

Kullavanijaya et al. ([1998] 2000) studied 88 Tai speech varieties in China, Vietnam, Laos and Thailand. Particular attention was paid to the tone systems which were used to determine subgroups of SWT. They identified six distinctive tones for Khuen in Chiang Mai Province Thailand and noted 'a rather peculiar phenomenon in Tai Khuen, i.e. C123, has merged with DL4' (Kullavanijaya et al. [1998] 2000:282).

<sup>8</sup> The present convention is to write 'Wan Weng' as in the map in Figure 4.

<sup>9</sup> Note that the tone labels are changed to make them consistent with other authors presented in this chapter.

<sup>10</sup> The present convention is to write 'Wan Hai' as in the map in Figure 4.



Although the authors do not make the point, the implication of their observation is that for this Khuen variety  $B \neq DL$ . The distribution is laid out in Figure 9, where ‘-’ denotes cells for which no information was reported.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	5	-	-
2. Voiceless unaspirated stops				-	-
3. Glottal	2			-	-
4. Voiced		4	6	-	5

**Figure 9:** Tone distribution in Khuen of San Patong from Kullavanijaya et al. ([1998] 2000:281)

## 2.2 Discussion

The six tonal systems described in the previous section are compared in Table . The studies are grouped according to geographical location and listed chronologically within each group. The numbers in boldface in the table are the tone labels. Where the authors provided sufficient detail of the phonological characterisation of the tones, this information is included. The phonological characterisation is based on the system proposed by Chao (1930) in which a five-point scale is used with ‘5’ representing the highest pitch and ‘1’ the lowest pitch. Although most authors specified which tones the tones on checked syllables were phonetically similar to, only Akharawatthanakun provided sufficient detail to include the phonological characterisation for checked syllables.

**Table 3:** Comparison of tone systems from previous Khuen research

	Keng Tung Township				Thailand	
	Keng Tung	Murng Lang	Kang Murng		Chiang Mai	Nan
Tone Box Reference	Egerød (1959)	Gedney (1964)	Gedney (1964)	Petsuk (1978)	Kullavanijaya et al. (1998)	Akharawatthanakun (2012)
<b>A1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>	<b>1</b>
<b>A2</b>	/35/	/35/	/35/	/15/		
<b>A3</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b> <sup>11</sup>	<b>2</b>	<b>2</b>
<b>A4</b>	/44/	/44/	/44/	/33/		
<b>B1</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>3</b> /21/	<b>3</b>	<b>3</b>
<b>B2</b>						
<b>B3</b>						
<b>B4</b>	4 /21/	4 /22/	4 /22/		<b>4</b>	4 /31/
<b>C1</b>	<b>5</b>	<b>5</b>	<b>5</b>	<b>4</b> /33ʔ/	<b>5</b>	<b>5</b>
<b>C2</b>						
<b>C3</b>						
<b>C4</b>	6 /52/	6 /51ʔ/	6 /51ʔ/	5 /51/	<b>6</b>	6 /41ʔ/
<b>DS1</b>	<b>2</b>	<b>2</b>	<b>2</b>	<b>2</b>	-	<b>2</b>
<b>DS2</b>					-	
<b>DS3</b>					-	
<b>DS4</b>	<b>6</b>	<b>6</b>	<b>6</b>	<b>5</b>	-	5 /44/
<b>DL1</b>	<b>3</b>	<b>3</b>	<b>3</b>	<b>3</b>	-	<b>3</b>
<b>DL2</b>					-	
<b>DL3</b>					-	
<b>DL4</b>	<b>4</b>	<b>4</b>	<b>4</b>		<b>5</b>	4 /31/

<sup>11</sup> The phonological characterisation given here is /33/ because in isolation the phonetic realisation of the tone is [23]. In certain other phonetic environments the tone is realised as [45].

All six varieties reveal the same pattern of splits and mergers for the A column, namely A12-34. The C and DS columns also pattern in the same way for all reported<sup>12</sup> varieties, namely C123-4 and DS123-4. A further feature that holds true for all varieties apart from the variety described by Kullavanijaya et al is the identical patterns of the B and DL columns, i.e., B=DL. The main difference between the varieties is whether or not the tone in B4 is distinct from the tone in B123. All varieties apart from Petsuk (1978) show a distinction. Other differences exist in the tones on the checked syllables in the varieties spoken in Thailand, viz. DL4 is associated with C123 in the Chiang Mai variety and DS4 is associated with C123 in Nan. While of interest, it is beyond the scope of this study to investigate these differences further. In addition to the high degree of similarity between varieties in terms of the patterns of tonal splits and mergers, there is also a high degree of similarity in terms of the phonological characterisations of the various tones. Aside from differences in the B column which are caused by the difference in tone split in Petsuk's study, the main difference is the nature of Tone 2. The first four studies report a level tone whereas Akharawatthanakun reports a mid rising tone.

In seeking an explanation for the differences in the B column, time seems to be the obvious factor. If these data can be taken as representative, it would appear that half a century ago all<sup>13</sup> Khuen varieties had six distinctive tones. Twenty years later however, in the Kang Murng area at least, there had been a coalescence of the two tones in the B column resulting in a five-tone system. The development of the six-tone system is consistent with the general pattern of tonal development in Tai languages in which each of the original tonal categories underwent a binary split yielding six separate tones. Various modern Tai varieties have tonal systems with fewer than six tones, caused by the merger of two of the six categories (see Pittayaporn 2009a:245-246 for examples). Taking this to be the case, we now sketch a possible process of development of the six-tone system described by Gedney (1996a&b [1964]), depicted in Figure 10.

	A	B	C	DS	DL
1. Voiceless friction sounds	1 = /35/	3 = /12/	5 = /33?/	2	3
2. Voiceless unaspirated stops					
3. Glottal	2 = /44/	4 = /22/	6 = /51?/	6	4
4. Voiced					

**Figure 10:** Gedney's Khuen six-tone system with phonemic characterisations

First note that voicing of initial consonants is associated with higher tonal onset in each column, in other words, Tones 2, 4 and 6 have higher initial pitch than their counterparts 1, 2 and 3 which have voiceless initial consonants. We conclude that segment-tone interaction was the main mechanism that caused the initial tone splits (cf Pittayaporn 2009a:260,262). For \*C (the C tone immediately prior to the tone split) we posit a high falling tone with glottalisation, akin to Tone 6. Tone 5 can be derived from Tone 6 firstly by voiceless onset lowering the initial pitch and then the process of contour reduction creating a level tone from the falling tone. For \*B a low rising tone can be posited akin to Tone 3. Tone 4 could develop from the low rising tone by segment-tone interaction and/or contour reduction. For \*A a high level tone can be posited, akin to Tone 2. Tone 1 could develop from the high level tone by segment-tone interaction lowering the initial pitch followed by perceptual maximisation raising the final pitch to make the contour more distinctive. The three tones posited for \*ABC closely resemble the PT tones reconstructed by Pittayaporn (2009a:271) and depicted in Table 1.

In order to provide a possible explanation of what brought about the merger of the tones in the B column yielding the five-tone system described by Petsuk (1978:109), it is necessary to consider both internal factors, such as the phonological change mechanisms proposed by Pittayaporn (2007b:1421-1422) and external factors such as language contact. During the period between Gedney's and Petsuk's research the number of Shan speakers in the Keng Tung area was increasing<sup>14</sup>. In fact even in 1957 Shan was 'very widely spoken in Kengtung' (Egerød 1959:123). As a trading centre and regional capital Keng Tung attracted Shan people from the areas further west in Shan State where they lived in greater numbers. So in the city Khuen came into close contact with Shan. Wardhaugh (2004:192-3) describes how innovations in language often spread from cities to surrounding rural areas. It is conjectured therefore that the change in Khuen tones began in Keng Tung because of contact with Shan and spread out from the city, most readily to those areas that

<sup>12</sup> Kullavanijaya et al (1998) do not give any information about the DS column.

<sup>13</sup> There is no data from that era for the Thailand varieties, but given that their contemporary patterns are the same as the older descriptions of varieties from Keng Tung Township, the simplest explanation is that they all derive from a common six-tone system.

<sup>14</sup> Personal communication with the elderly residents of Kengtung who have witnessed the changes during their lifetime.

have the best access to Keng Tung. It is not possible to directly verify this hypothesis since the only published data from Keng Tung are Gedney ([1964] 1994a) and Egerød (1959). Egerød was sure however that Khuen in Keng Tung had been influenced by Shan:

Khün has been heavily influenced by Shan, especially as far as borrowing of vocabulary is concerned. The language of many Khün speakers is now corrupted to the point of being almost Shan. (Egerød 1959:124).

Egerød (1959:125) further attributed to the influence of Shan the Khuen 'tendency to confuse' short high vowels with their long counterparts. In addition to the general influence mentioned by Egerød, there is anecdotal evidence to the effect that elderly Khuen residents of Keng Tung cannot remember anything other than a five-tone system, indicating that the five-tone system has been in use for most if not all of their lives. This study includes data from Keng Tung city as well as the areas of Kang Murng and Murng Lang for comparison with previous studies as well as the areas of Yang Keng and Kat Fah.

### **3 Methodology**

This study followed the general two-phase approach for determining tonal systems in Tai languages described by Tingsabadh (2001:212-214). In the first phase a wordlist based on Gedney's tone checklist is elicited (Gedney 1972:434-436). Analysing the responses to the first phase establishes the patterns of tonal splits and mergers and in particular the number of tones – normally taken to be the number of distinct tones on smooth syllables. Having established the number of tones, the researcher develops a second list of words that contrast tones in identical or analogous environments. Akharawatthanakun (2002, 2004) used the methodology to investigate tonal patterns in dialect contact areas. Gedney's checklist works well where the Tai variety under investigation uses words cognate with the words on the checklist. Where this is not the case however, the chances are that the elicited word will not lie in the desired tonal category. The methodology for the present study is based on that of Akharawatthanakun (2012) who increased the number of items in the wordlist for the first phase from 60 to 140 to provide a much more robust coverage of the whole range of tonal categories. The following sections describe the informants and data collection procedures used in the study.

#### **3.1 Informants**

As described in Section 2.2, age and location are factors that might affect pronunciation. Furthermore, since the social roles of males and females are often different, gender was also included as a factor to be investigated. Data was primarily collected over a period of 7 days from male and female Khuen informants across a broad spectrum of ages from Keng Tung city and Kang Murng and Murng Lang areas as well as from other village tracts that do not lie in either of these areas. A total of 12 informants (7 male; 5 female) were included in the study: 2 from Keng Tung (both male); 2 from Kang Murng (1 male; 1 female); 2 from Murng Lang (1 male; 1 female); 2 from Yang Keng (1 male; 1 female); and 4 from Kat Fah (2 male; 2 female). Females were aged between 15 and 78; males were aged between 27 and 83. Apart from the tonal differences described below, there were no noticeable differences in pronunciation across age, gender and location categories.

#### **3.2 Data collection**

There were two phases of data collection. The purpose of Phase 1 was to determine the pattern of tonal splits and mergers for a particular informant. The 140-item wordlist devised by Akharawatthanakun (2007) was used for this purpose, but due to time constraints with informants usually only a subset of items were used. The items were elicited by means of pictures, miming or pointing. The responses were transcribed by hand but not recorded. Over the 12 informants, the average number of words used to determine the pattern of tonal distribution for an individual was 54.

The purpose of Phase 2 was to provide acoustic data from a set of words which as far as possible had tone as the sole differentiating feature. The wordlist used in Phase 2 is thus referred to as the 'Analogous Wordlist' and is shown in Figure 11 displayed on the Gedney box framework.

	A	B	C	DS	DL
1	ခဝ /k <sup>h</sup> a:35/ 'leg'	ခဝ /k <sup>h</sup> a:12/ 'galangal'	ခဝ /k <sup>h</sup> a:332/ 'slave'	ခဝ /k <sup>h</sup> at44/ 'to polish'	ခဝ /k <sup>h</sup> a:t12/ 'to be torn'
2	ကဝ /ka:35/ 'crow'	ကဝ /ka:12/ 'charm'	ကဝ /ka:332/ 'to dance'	ကျ /kap44/ 'with'	ကပ /ka:p12/ 'coconut husk'
3	ဝါ /ba:44/ 'thin'	ဝါ /ba:12/ 'shoulder'	ဝါ /ba:n332/ 'village'	ဝါ /bat44/ 'ticket'	ဝါ /ba:t12/ 'wound'
4	ဝါ /ka:44/ 'car'	ဝါ /ka:22/ 'fee; cost'	ဝါ /ka:512/ 'to trade'	ဝါ /kap53/ 'tight'	ဝါ /ka:p22/ 'to grip in teeth'

**Figure 11:** Analogous wordlist displayed on Gedney box framework

In each box in Figure 11 the word to be elicited is given in Khuen script with a phonetic description based on Gedney ([1964] 1994a:978) and finally an English gloss. The six distinctive tones of the Khuen tonal system as described by Gedney are covered by the set of words found in cells A2;B2;C2;A4;B4;C4 which are minimally differentiated by tone. The phonetic transcription of tone is based on Chao (1930). Items from every box in the Gedney box framework were elicited and could be compared with the conclusions from Phase 1.

Items on the analogous wordlist were elicited by means of pictures which were presented to the informant in random order. Over the 12 informants the average number of replications recorded of each item in the analogous wordlist is 4, that is on average approximately 90 tokens were recorded for each informant. Tokens were recorded using a Sony MZ-RH1 minidisc recorder before being transferred to computer where Speech Analyzer<sup>15</sup> was used to produce and view graphs of pitch contours and raw waveforms.

The procedure for data collection for Shan was slightly different to that for Khuen. Having previously elicited and recorded a general 406-item wordlist from a 32-year old female Shan informant from Keng Tung, the tonal distribution for this informant had already been established. For the purposes of this study therefore only minimal contrasts on open syllables were elicited and recorded. Two sets of contrasts as shown in Table 4 were elicited and recorded in random order. Set I was devised by Poo-Israkij (1985:95) and Set II by the author.

**Table 4:** Sets of words for Shan tone contrasts

Tone Label	Set I		Set II	
<b>Tone 1</b>	တၢ် /ta:35/	'eye'	တၢ် /ka:35/	'crow'
<b>Tone 2</b>	တၢ်: /ta:45/	'to smear'	တၢ်: /ka:45/	'car'
<b>Tone 3</b>	တၢ်. /ta:212/	'CLF for a meal'	တၢ်. /ka:212/15	'to go'
<b>Tone 4</b>	တၢ်: /ta:332/	'landing'	တၢ်: /ka:332/	'to dance'
<b>Tone 5</b>	တၢ်. /ta:532/	'to challenge'	တၢ်. /ka:532/	'to trade'

For Set I two tokens of each word were recorded, and because of its affinity with the wordlist used for Khuen, eleven tokens of each word in Set II were recorded. Each word in the table has a transcription in Shan script, a phonemic transcription based on Poo-Israkij (1985:91) and an English gloss.

<sup>15</sup> Speech Analyzer was developed by JAARS (an affiliate with SIL International.) The tool has pitch and spectro analysis capabilities.



#### 4 Analysis of tone systems in contemporary Khuen varieties

The tonal distribution was determined for each individual informant and two patterns emerged: a six-tone system following the pattern of Gedney ([1964] 1994a) as displayed in Figure 6 and a five-tone system following the pattern of Petsuk (1978) as displayed in Figure 7. Before presenting the results, it is worth noting some comments made by one of the informants. The 38-year old male informant from Yang Keng commented that in his area people used to have six tones but now younger generations only have five. This begs the question of how much older someone would need to be to be still using the six-tone system, in his area or other areas. In this respect, note that the informant for Petsuk (1978) would be around 55 today. Thus it might be expected that informants would need to be well over 55 to be using the six-tone system. This is the case for the other informant from Yang Keng in this study (78-year old female) who still uses the six-tone system. The male informant from Yang Keng further commented that the loss of one tone came about in order to pronounce the tone for Gedney boxes B123 more like Shan. This latter point will be discussed in detail in Section 4.3.

A total of nine out of the twelve informants have a five-tone system and the remaining three have six tones. Both informants from Murng Lang (a 68-year old male and a 30-year old female) have six tones as does the 78-year old female informant from Yang Keng. The two informants from Kang Murng both have five tones as might be expected given their ages (27; 56) in the light of Petsuk's study. If the change is indeed attributable to language contact with Shan, then it is not surprising that the 49-year old male informant from Keng Tung should also have five tones. As far as the 83-year old male informant from Keng Tung is concerned, the fact that he has five tones suggests either that he once had six tones and eventually lost one of the tones or that he always had five tones in which case the change to a five-tone system had begun some time before Petsuk's informant was born. The lack of change in Murng Lang is consistent with the fact that Murng Lang has poorer road connections with Keng Tung and hence there is likely to be less contact between Murng Lang inhabitants and outsiders. While the 67-year old male informant from Kat Fah claimed there was a difference between B123 and B4, in practice his pronunciation of tone was the same for these categories. This perception that these tones were different despite the lack of a phonetic difference can be interpreted as evidence that in the past there was a phonetic difference which has been lost over time. Alternatively the informant's perception could be due to his knowledge of the Khuen script which provides for separate tones, but even this points to a six-tone system in the past, whether in the experience of this particular informant or not.

In order to compare the two systems one informant of each type was chosen for further analysis. For each individual informant, an average was taken over five tokens of each tone. This was done by first normalising the duration of each tone. The effect of this was to create more representative contours for each tone. The duration of each tone was then de-normalised to produce a graph that better reflects the actual tonal contours in relation to one another. The next two sections describe respectively the six- and five-tone systems.

##### 4.1 Six-tone system

The pattern of tonal distribution for those informants having a six-tone system is shown in Figure 12. The pattern is the same as that of Egerød (1959:124) and Gedney (1994a; 1994b). The tones are labelled 1,...,6 for ease of reference.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	5	2	3
2. Voiceless unaspirated stops					
3. Glottal	2	4	6	6	4
4. Voiced					

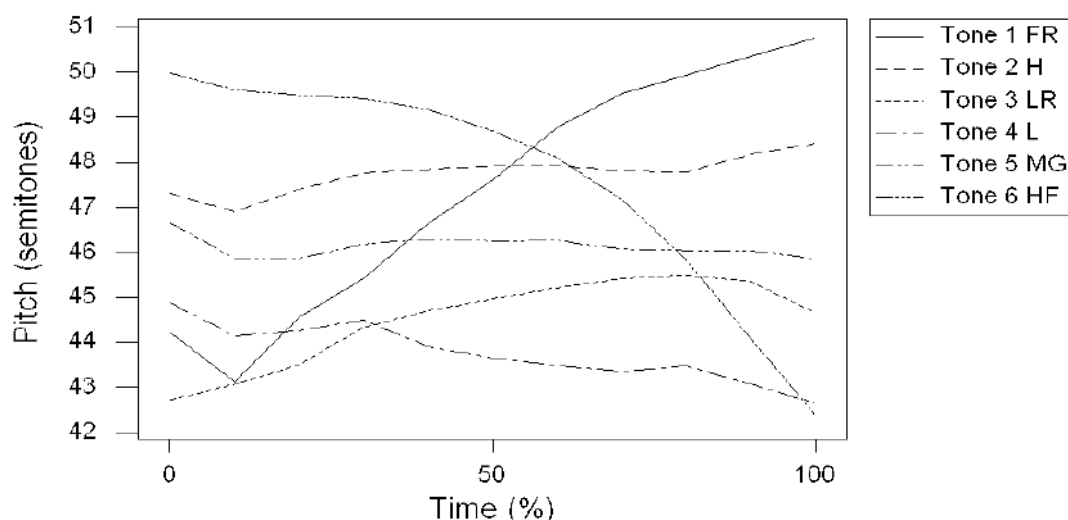
**Figure 12:** Tone distribution for contemporary six-tone system

The six tones are characterised both in words and numbers in Table 5. For the sake of comparison with previous studies, phonological characterisations of the tones from Gedney ([1964] 1994a) are also given.

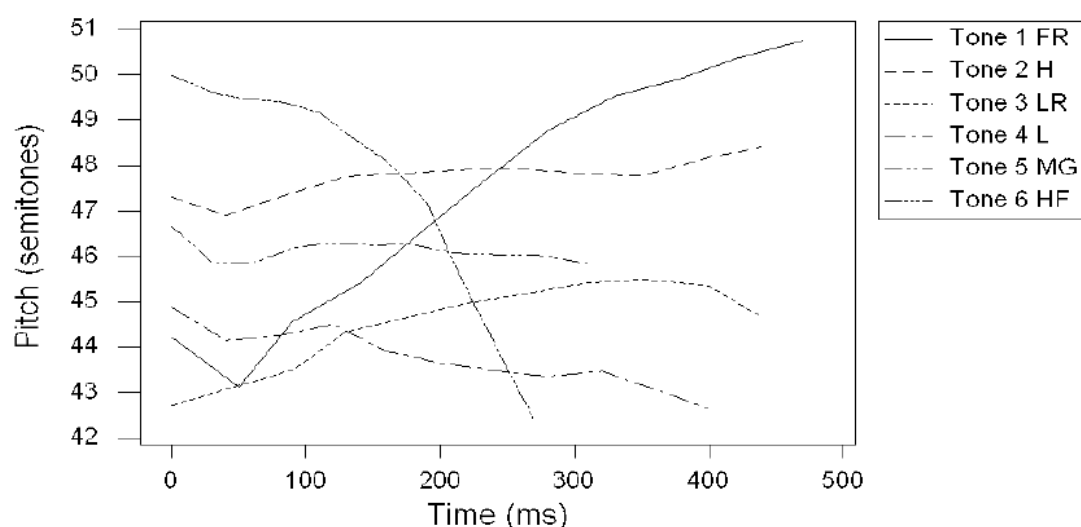
**Table 5:** Description of tones in six-tone system

Tone Label	Contemporary Six-Tone System			Gedney [1964]
	Word description	Phonetic characterisation	Phonological characterisation	Phonological characterisation
Tone 1	falling rising (FR)	[215]	/215/	/35/
Tone 2	high (H)	[44]	/44/	/44/
Tone 3	low rising (LR)	[13]	/13/	/12/
Tone 4	low (L)	[21]	/22/	/22/
Tone 5	mid glottalised (MG)	[33ʔ]	/33ʔ/	/33ʔ/
Tone 6	high falling (HF)	[51]	/51/	/51/

In Table 5 the phonological characterisation follows Yip (2002:23) in interpreting the fall in low Tone 4 as being too slight to make it a contour tone. Gedney (1972:423) observed that most tonal systems in Tai languages include tones with some kind of glottal constriction. This is a suprasegmental feature which differs from the segmental glottal in checked syllables (Court 1998:4). It is no surprise therefore to find that mid glottalised Tone 5 ends in glottal closure. Moreover since consonant clusters do not occur in the coda (c.f. Figure 2) interpreting the glottalisation as a segmental glottal final would violate the established syllable structure. The glottalisation is therefore associated with the tone and is naturally represented on the tone label because both pitch and glottalisation are laryngeal features. One further point to note is how closely the contemporary tones match the historical ones phonologically. The normalised pitch contours of the six distinctive tones are plotted in Figure 13.

**Figure 13:** Normalised pitch contours in contemporary Khuen six-tone system

As can be seen from the pitch contours in Figure 13 the tones are clearly differentiated by pitch and the nature of the contour, either rising, falling or level. The differentiation becomes even clearer however when duration is taken into consideration. The real-time pitch contours of the six-tone system are plotted in Figure 14.



**Figure 14:** Normalised pitch contours in contemporary Khuen six-tone system

The contours in Figure 14 show a considerable range of durations: the shortest duration (high falling Tone 6, 270ms) is only 57% of the longest duration (low rising Tone 1, 470ms). Similarly mid glottalised Tone 5 (310ms) is 66% of the duration of low rising Tone 1. It is significant that the shorter duration of the mid glottalised Tone 5 serves to distinguish it from the two other level tones, namely high Tone 2 and low Tone 4. Having both duration and glottalisation to distinguish the mid glottalised Tone 5 from the other two level tones means that level tones in the Khuen six-tone system are less likely to be subject to the kind of confusion in difficult listening conditions observed by Abramson (1976) for Standard Thai. The nature of the contour of high falling Tone 6 is highly distinctive. The high starting point and subsequent rapid fall in pitch (which inevitably leads to a shorter duration) yields a tonal contour unlike any of the other tones.

The magnitude of the difference in durations raises the question of whether the shorter durations of Tones 5 and 6 are attributable to the tones or simply due to vowel length. It is beyond the scope of this study to address this issue save for making the following comments. Firstly Abramson's classic work on Standard Thai tones showed a comparable range of durations for syllables ending in long vowels as well as syllables with short vowels ending in a sonorant (Abramson 1962:120-124). Secondly, the phonemic status of vowel length in Tai languages is an issue that has long been pondered by Tai linguists - see for example Hartmann (1976). Thirdly, vowel duration was a key feature of the PT tones reconstructed by Pittayaporn (2009a:271-274). Finally, in the present study all of the realisations of mid glottalised Tone 5 and high falling Tone 6 whether on open syllables such as those used for Figure 14 or on syllables ending in a sonorant, had durations that were noticeably shorter than for Tones 1-4 on the same range of syllable types. This final point is consistent with Pittayaporn's reconstruction of \*C as having shorter duration.

#### 4.2 Five-tone system

The pattern of tonal splits and mergers for those informants having a five-tone system is shown in Figure 15. The pattern is the same as that of Petsuk (1978:109). For ease of comparison between the five- and six-tone systems the tones in A and C columns are labelled the same. The single tone in the B column of the five-tone system is labelled '3-4' not just for the sake of comparison with the two tones labelled '3' and '4' in the six-tone system but also in order to make the correspondences of Tones 5 and 6 with their counterparts in the six-tone system more obvious.

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3-4	5	2	3-4
2. Voiceless unaspirated stops					
3. Glottal	2		6	6	
4. Voiced					

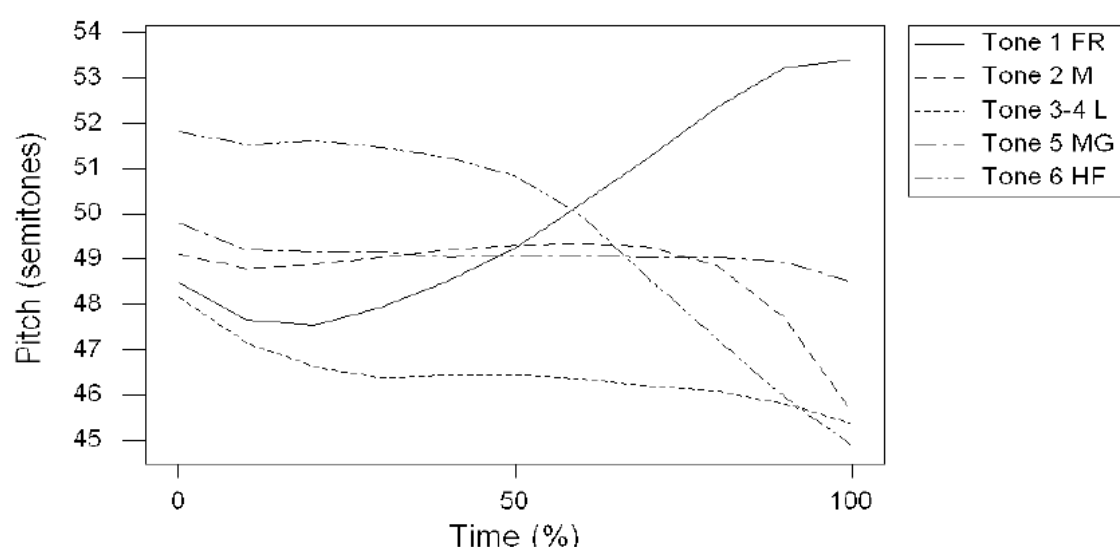
**Figure 15:** Tone distribution for contemporary Khuen five-tone system

The five tones are characterised both in words and numbers in Table 6. For the sake of comparison with previous studies, phonological characterisations of the tones from Petsuk (1978:24-25) are also given.

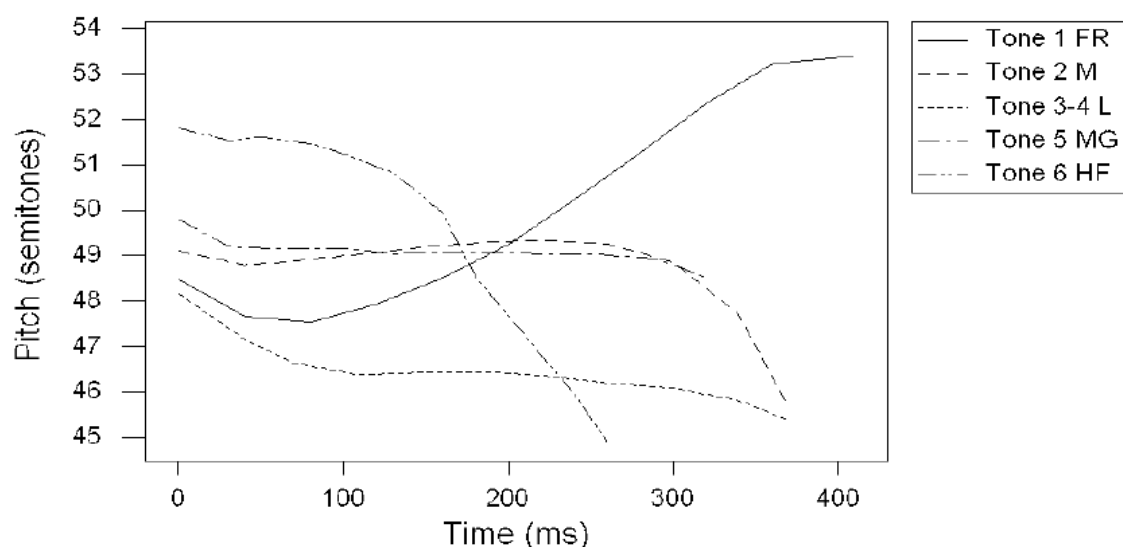
**Table 6:** Description of tones in contemporary Khuen five-tone system and Petsuk (1978:24-25)

Tone Label	Contemporary Five-Tone system			Petsuk (1978)
	Word description	Phonetic characterisation	Phonological characterisation	Phonological characterisation
Tone 1	falling rising (FR)	[325]	/325/	/15/
Tone 2	mid (M)	[32]	/33/	/33/
Tone 3-4	low (L)	[21]	/22/	/22/
Tone 5	mid glottalised (MG)	[33ʔ]	/33ʔ/	/33ʔ/
Tone 6	high falling (HF)	[41]	/41/	/51/

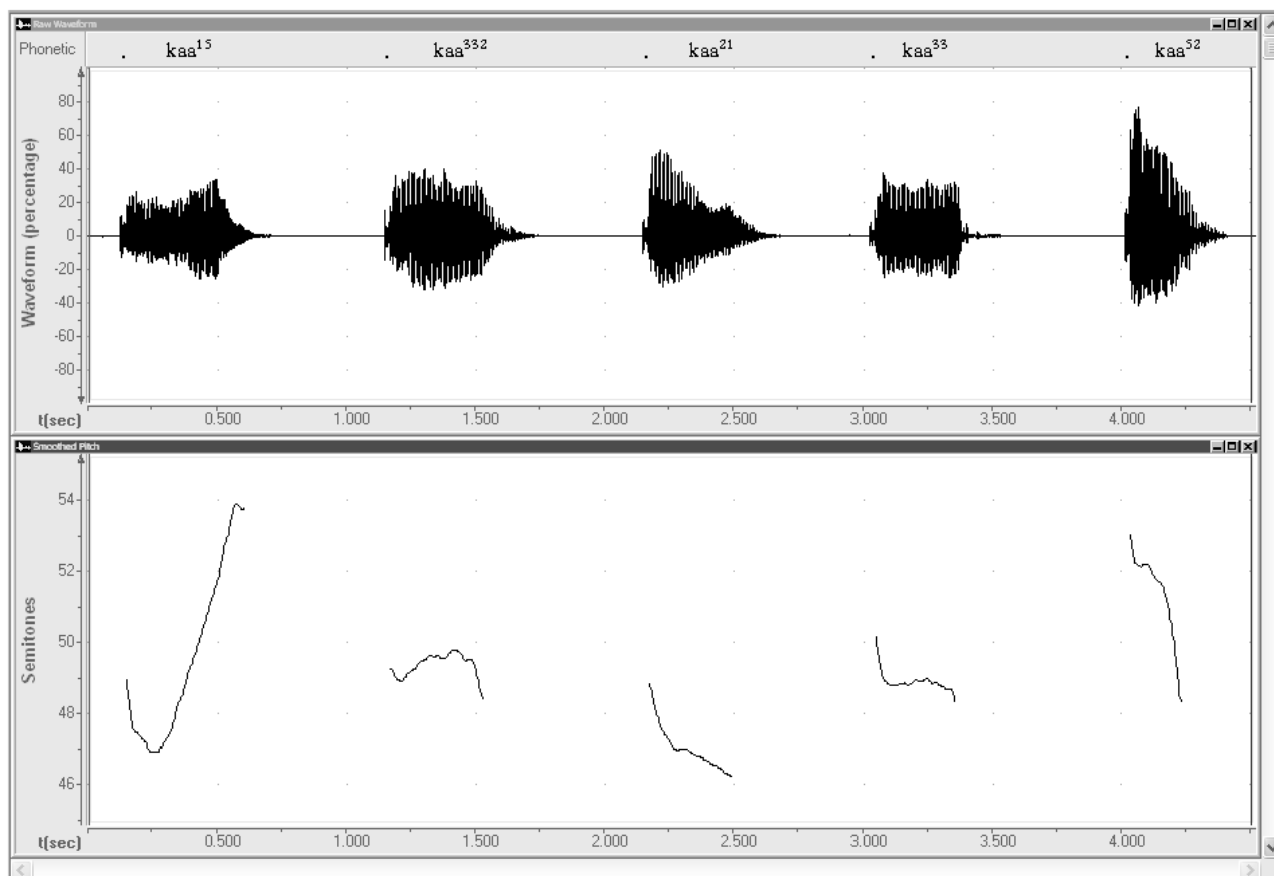
The characterisations show that the tones in the contemporary five-tone system are essentially the same as those reported by Petsuk. The normalised pitch contours of the five distinctive tones are plotted in Figure 16.

**Figure 136:** Normalised pitch contours in contemporary Khuen five-tone system

As can be seen from the pitch contours in Figure 16 mid Tone 2 and mid glottalised Tone 5 are very close in pitch for the majority of their duration. This is seen even more clearly in Figure 17 where the pitch contours are plotted on a real time scale.

**Figure 17:** Pitch contours in contemporary Khuen five-tone system

As can be seen from Figure 17 the pitch of mid Tone 2 drops at the end of its duration whereas the pitch of mid glottalised Tone 5 remains relatively constant. The main auditory cue that distinguishes the two tones however, is the glottal constriction in mid glottalised Tone 5. This may be seen in the plot of the waveform in Figure 18. The upper panel of the display in Figure 18 shows the raw waveform of the five tones in rising numerical order, Tone 1 followed by Tone 2 etc. The lower panel shows the corresponding pitch curves. The glottal constriction of mid glottalised Tone 5 can be seen by the relatively abrupt ending of the waveform corresponding to Tone 5<sup>16</sup>. The magnitude of the oscillations about the mean suddenly decreases due to the sharp glottal closure.



**Figure 148:** Waveform and pitch contour plots for contemporary Khuen five-tone system

One further point to note is that as described by Petsuk (1978:24) Tone 2 has two phonetic variants which are in complementary distribution. The variant depicted in Figures 16-18 is for single syllable words in isolation. When preceded by a syllable with Tone 1 or a checked syllable with Tone 2, the realisation of Tone 2 is higher in pitch. While it is beyond the scope of this study to identify all conditioning factors, it is worth observing that the higher-pitched variant of Tone 2 is more readily distinguished from Tone 5 than the citation form.

#### 4.3 Comparison of five- and six-tone systems

In order to explore the relationship between the two tone systems, each tone is described in prose and Chao (1930) numbers in Table 7.

<sup>16</sup> Gardner (1997:11) gives a waveform which shows very similar glottalisation on the C123 tone in the Nan dialect of Northern Thai.

**Table 7:** Description of tones for contemporary Khuen five- and six-tone systems

	Six-Tone System		Five-Tone System	
	Description	Phonological characterisation	Description	Phonological characterisation
Tone 1	falling rising	/215/	falling rising	/325/
Tone 2	high	/44/	mid	/33/
Tone 3	low rising	/13/	low	/22/
Tone 4	low	/22/		
Tone 5	mid glottalised	/33ʔ/	mid glottalised	/33ʔ/
Tone 6	high falling	/51/	high falling	/41/

Tones 1, 5 and 6 are phonetically identical or very similar in the two systems. Tone 2 is quite similar, with an additional fall in the five-tone system (see discussion in Section 4.2). The main difference is that whereas in the six-tone system Tones 3 and 4 are distinct, in the five-tone system there is but a single tone. Since the contemporary six-tone system was shown to be essentially the same as the historical one, and the difference between the two systems is caused by the coalescence of Tones 3 and 4 in the six-tone system, it is of interest to investigate the cause of this change. Khuen has been strongly influenced by Shan (see Section 2.2) and one informant in the present study commented that the pronunciation of the Khuen low rising tone (Gedney boxes B123) was changed to sound more like Shan. In the following section therefore Khuen tones will be compared with those in contemporary Keng Tung Shan.

#### 4.4 Comparison with Shan

The pattern of tonal distribution is shown in Figure 19. This is consistent with seven out of eight studies in Poo-Israkij's summary of the tone systems of Shan from various locations (Poo-Israkij 1985:101-107). In particular it agrees with Brown's data for Keng Tung<sup>17</sup> Shan (Brown [1965] 1985:153).

	A	B	C	DS	DL
1. Voiceless friction sounds	1	3	4	2	3
2. Voiceless unaspirated stops					
3. Glottal					
4. Voiced	2	4	5	5	4

**Figure 19:** Tone distribution for contemporary Keng Tung Shan

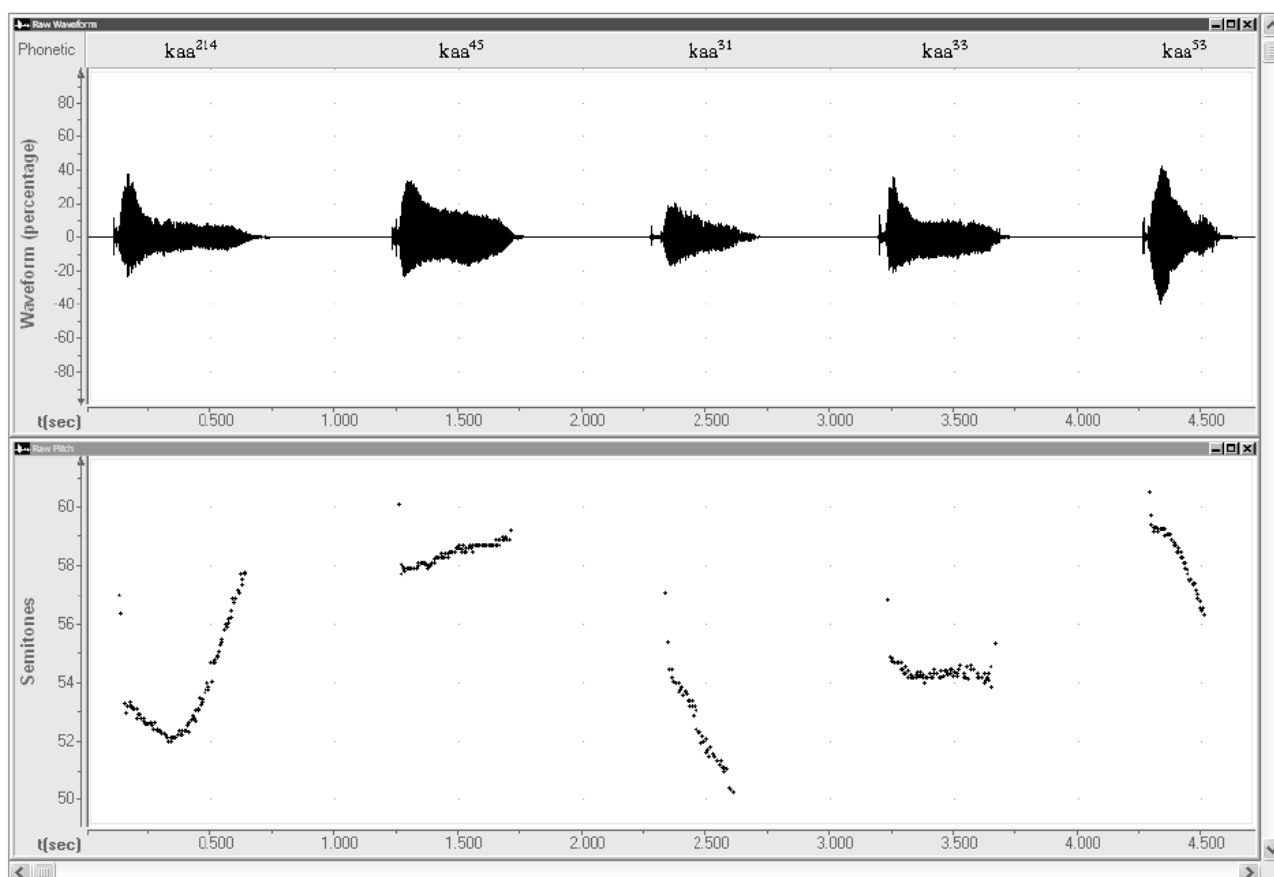
The five Shan tones are characterised both in prose and Chao (1930) numbers in Table 8. For the sake of comparison with previous studies, phonological characterisations of the tones from Brown ([1965] 1985:89) are also given.

**Table 8:** Description of Shan tones

Tone Label	Contemporary Keng Tung Shan			Brown (1985)
	Word description	Phonetic characterisation	Phonological characterisation	Phonological characterisation
Tone 1	falling rising (FR)	[214]	/214/	/24/
Tone 2	high (H)	[45]	/55/	/45/
Tone 3	mid falling (MF)	[31]	/31/	/11/
Tone 4	mid glottalised (MG)	[33ʔ]	/33ʔ/	/33/
Tone 5	high falling (HF)	[53]	/53/	/53ʔ/

<sup>17</sup> Actually, Brown's informant was from a Shan village in Thailand's Chiang Rai province that had migrated from Keng Tung about 30 years previous to the fieldwork (Brown [1965] 1985:89).

The shapes of the tones are indicated by the pitch curve plot from Speech Analyzer shown in Figure 20. The tones are plotted in rising numerical order, i.e., Tone 1 first then Tone 2, etc.



**Figure 20:** Waveform and pitch contour plots for contemporary Keng Tung Shan

To facilitate the comparison of the Khuen five-tone system with the Shan system, the two systems are represented side-by-side in Figure 21. The Gedney box framework is used to show distribution and the tones are characterised by abbreviations of the prose descriptions (FR - falling rising, etc.) and phonetic characterisations using Chao (1930) numbers.

Khuen Five-Tone System						Keng Tung Shan Tone System					
	A	B	C	DS	DL		A	B	C	DS	DL
1	FR	L [21]	MG [33ʔ]	M [44]	L [21]	1	FR [214]	MF [31]	MG [33ʔ]	H [55]	MF [31]
2	[325]					2					
3	M [32]		HF [41]	HF [42]		3					
4						4	H [45]	MG [33ʔ]	HF [53]	HF [53]	MG [33ʔ]

**Figure 21:** Comparison of tones in Khuen five-tone and Shan systems

Firstly note that both systems have three tones that are identical or near identical: Falling Rising Tone 1; Mid Glottalised Tone 3 and High Falling Tone 5. Thus 7 out of 20 Gedney boxes in the two systems have identical or near identical tones. A further 10 boxes have tones that are similar. The pitch height is the main feature that distinguishes Khuen Mid tone 2 from Shan High Tone 2 and the extent of the fall in Khuen Low Tone 3 and Shan Low falling Tone 3. The remaining 3 boxes whose tones are not similar are due to the different tonal distribution patterns. When one includes the high variant of Mid Tone 2 in the Khuen system which is very similar to the High Shan Tone 2, the proportion of cognate words that would be pronounced with near identical tones in the two languages will be even higher in connected speech.

Having seen the how closely the tones in the Khuen five-tone system resemble those of the Shan system it is natural to go one step further and consider the Khuen six-tone system. The tone shape that was 'lost' from the Khuen six-tone system (described as Tone 3 'low rising' /13/ in Table 7) was the one that had

no phonetically similar counterpart in the Shan system. This represents compelling evidence in favour of the hypothesis that Khuen tones have changed to sound more like Shan, exactly as perceived by the young male informant from Yang Keng. These results bear a strong resemblance to those of L-Thongkum (1997) who found that whereas older Mien speakers in northern Thailand had a six-tone system, younger speakers had a five-tone system with some tonal contours modified to more closely resemble those of the five-tone system of Standard Thai.

## **5 Summary and Conclusion**

The main aim of this research was to compare the tonal systems in evidence in contemporary Khuen with each other and with those described in previous studies. The locations of Keng Tung city and the Kang Murng and Murng Lang dialect areas were of particular interest. In the following sections some conclusions are drawn and suggestions made for further research.

### **5.1 Summary**

Several pieces of evidence point to a change over time in most geographical locations. Studies by Egerød (Keng Tung, 1957) and Gedney (Kang Murng and Murng Lang, 1964) showed six-tone systems. By 1978 Kang Murng had changed to a 5-tone system by the coalescence of the tone for B123 with that for B4 (Petsuk 1978:109). The present study confirmed this change for Kang Murng and showed a similar change for Keng Tung. Although no previous studies of Yang Keng or Kat Fah have been published, anecdotal evidence suggests that there used to be six tones but now there are five. One elderly informant from Yang Keng still has six tones. The one place where change was not observed is Murng Lang: both informants included in the present study use six tones, as did Gedney's informant back in 1964.

One of the informants in the study commented that change in the B123 tone from the former six-tone system was motivated by the desire to sound more like Shan, presumably for ease of communication with the growing number of Shan speakers. This is substantiated in the present study by demonstrating the phonetic similarity of the new B123 tone in the Khuen 5-tone system to the B123 tone in Keng Tung Shan. Furthermore, the geographic locations where the change has taken place, i.e., Keng Tung city, and the areas of Kang Murng, Kat Fah, and Yang Keng are those that have better road connections to Keng Tung city and hence are more likely to have contact with Shan speakers. The relative difficulty of travelling from Keng Tung to Murng Lang is hypothesised as a major factor in the preservation of the six-tone system in Murng Lang.

### **5.2 Implications**

The following implications emerge from this research:

- The present variation in tone systems is explained by language contact and the five-tone system is a recent modification of the six-tone system so for the purposes of using tonal criteria in cross-linguistic classification the six-tone system should be chosen to represent Khuen.
- The implication for the Khuen writing system is that as long as there are systems with six tones in use in Khuen communities, the writing system must be able to represent six tones.
- The evidence here is consistent with change spreading out from a city due to language contact. Khuen varieties in Keng Tung and the most accessible outlying regions have changed their tones whereas the more remote areas have not changed.

### **5.3 Further research**

In order to gain a fuller picture of the relative proportions of the population who use five- and six-tone systems more of the same kind of data can be collected to bolster the body of evidence already gathered. Other village tracts in particular to the east and south of Keng Tung should also be included in a further study.

It might be that accessibility from Keng Tung is too simplistic to serve as a predictive factor of whether communities change to the five-tone system. An extra factor could be introduced, namely the accessibility of a particular village, measured for example by how close it is to a paved road. In this case a social network approach along the line of Goetz's study of language maintenance in Tai Nuea would provide a much firmer basis on which to draw conclusions (Goetz 2001).

A fuller discussion of how phonological variation in contemporary spoken Khuen is accommodated by the Khuen writing system is in preparation by the author.



## References

- Abramson, A. S. 1962. The vowels and tones of Standard Thai: Acoustical measurements and experiments. Bloomington, IN: Indiana University Center in Anthropology, Folklore and Linguistics Publication 20.
- Abramson, A. S. 1976. Thai tones as a reference system, in Thomas W. Gething et al. (eds), 1-12.
- Akharawatthanakun, Phinnarat. 2002. Tonal variations and changes in a language mixture area: a case study of Northeastern Thailand (Isan). *Manusya* 5.2:30-51. Bangkok: Chulalongkorn University.
- Akharawatthanakun, Phinnarat. 2004. Tonal variation and change in dialects in contact: a case study of Lao. *Manusya* 7.1:56-95. Bangkok: Faculty of Arts, Chulalongkorn University.
- Akharawatthanakun, Phinnarat. 2012. Lexical and phonological variation and change in Tai dialects spoken in Nan province. Research Report submitted to The Commission on Higher Education and The Thailand Research Fund.
- Brown, J. Marvin. 1985. From Ancient Thai to modern dialects and other writings on historical Thai linguistics. Bangkok: White Lotus.
- Chamberlain, James R. 1972. Tone borrowing in five northeastern dialects. *Tai phonetics and phonology* ed. by Jimmy G. Harris and Richard B. Noss. 43-46. Bangkok: C.I.E.L.
- Chaymberlain, James R. 1975. A new look at the history and classification of the Tai languages. *Studies in Tai Linguistics*, ed. by Jimmy G. Harris and James R. Chamberlain, 49-66. Bangkok: C.I.E.L.
- Chao, Y. R. 1930. A system of tone letters. *Le Maitre Phonétique*.
- Court, Christopher. 1998. The Gedney Boxes for SW Tai: the need for a D? (Dead, ?-final) column. [Handout] The International Conference on Tai Studies, July 29-31, 1998. Bangkok: Institute for Language and Culture for Rural Development, Mahidol University.
- Diller, Anthony V. N. 2008. Introduction, in Diller, Anthony V. N. et al (eds), 3-8.
- Diller, Anthony V. N., Jerold A. Edmondson and Yongxian Luo (eds.) 2008. *The Tai-Kadai languages*. Abingdon, Oxon: Routledge.
- Dodd, William Clifton. [1923] 1996. *The Tai Race: Elder brother of the Chinese. Results of Experiences, Exploration and Research of William Clifton Dodd, D.D. Thirty-three years a missionary to the Tai people in Siam, Burma and China*. Bangkok: White Lotus Press.
- Edmondson, Jerold A. 2008. Shan and other northern tier southeast Tai languages of Myanmar and China: themes and variations, in Diller, Anthony V. N. et al (eds), 184-206.
- Egyrød, Søren. 1959. Essentials of Khün phonology and script. *Acta Orientalia* 24:123-146.
- Gardner, Jonathan. 1997. A tonal study of Kammuang in Nan and Phrae. Research Report #138 Payap Research and Development Institute and The Summer Institute of Linguistics. Chiang Mai: Payap University.
- Gedney, William J. [1964], 1994a. Introduction to Klang Muong Khuen, in Hudak, Thomas John (ed.) 977-1020.
- Gedney, William J. [1964], 1994b. Introduction to Baan Veng Khuen, in Hudak, Thomas John (ed.), 1021-1055.
- Gedney, William J. [1964], 1996. *The Lue Language: Glossaries, texts and translations*, ed. by Thomas John Hudak. *Michigan Papers on South and Southeast Asia* 44. Ann Arbor: University of Michigan Press.
- Gedney, William J. 1972. A checklist for determining tones in Tai dialects. *Studies in Linguistics in Honour of George L. Trager*, 423-437. The Hague: Mouton.
- Gething, Thomas W., Harris, Jimmy G. and Kullavanijaya, Pranee (eds). 1976. *Tai linguistics in honor of Fang-Kuei Li*. Bangkok: Chulalongkorn University Press.
- Goetz, Ruth Kelly. 2001. Investigating language maintenance: social correlates of language choice among the Dehong Dai. Ann Arbor: The University of Michigan Doctoral Dissertation.
- Hartmann, John F. 1976. The waxing and waning of vowel length in Tai dialects in Thomas W. Gething et al. (eds), 142-159.
- Hudak, Thomas John, ed. 1987. Thai. In Bernard Comrie ed., *The major languages of East and South-east Asia*, 21-28. London: Routledge.
- Hudak, Thomas John, ed. 1994. William J. Gedney's Southwestern Tai dialects: Glossaries, texts, and translations. *Michigan Papers on South and Southeast Asia* 42. Ann Arbor: Center for South and Southeast Asian Studies, The University of Michigan.

- Kullavanijaya, Pranee and Theraphan L-Thongkum. [1998] 2000. Linguistic criteria for determining Tai ethnic groups: Case studies on central and south-western Tais. Proceedings of The International Conference on Tai Studies July 29-31 1998, ed. by Somsonge Burusphat. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.
- L-Thongkum, Theraphan. 1997. Tone change and language contact: a case study of Mien Yao and Thai. In *Southeast Asian Linguistic Studies in Honour of Vichin Panupong*, ed. by Arthur S. Abramson, pp.153-160. Bangkok: Chulalongkorn University Press.
- Lebar, Frank M., Gerald C. Hickey and John K. Musgrave. 1964. *Ethnic groups of Mainland Southeast Asia*. New Haven: Human Relations Area Files Press.
- Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Tex.: SIL International. Online version: <http://www.ethnologue.com/>.
- Li Fang-Kuei. 1964. The phonemic system of the Tai Lü language. *The Bulletin of the Institute of the History and Philosophy*. *Academica Sinica* 35:7-14.
- Li Fang-Kuei. 1977. *A Handbook of Comparative Tai*. Honolulu: University of Hawaii Press.
- Morey, Stephen. 2005. *The Tai languages of Assam - a grammar and texts*. Canberra: Pacific Linguistics.
- Owen, Robert Wyn. 2008. *Language use, literacy and phonological variation in Khuen*. Chiang Mai: Payap University M.A. thesis.
- Peltier, Anatole-Roger. 1993. *A Khün reader*. Chiang Mai: Wat Tha Kradas.
- Pent, Hans. 1994. *A brief history of Lān Nā: Civilisations of North Thailand*. Chiang Mai: Silkworm Books.
- Petsuk, Rasi. 1978. *General characteristics of the Khün language*. Bangkok: Mahidol University M.A. Thesis.
- Pittayaporn, Pittayawat. 2007a. A chronology-sensitive approach to subgrouping: the case of Southwestern Tai. Paper presented at the 81st Annual Meeting of Linguistics Society of America. Anaheim, California. January 4-7.
- Pittayaporn, Pittayawat. 2007b. Directionality of tone change. In Jürgen Trouvain and William J. Barry (eds.), *Proceedings of the 16th International Congress of Phonetic Sciences (ICPhS XVI)*, 1421-24. Saarbrücken, Germany: Saarland University.
- Pittayaporn, Pittayawat. 2009a. *Phonology of Proto-Tai*. Ph.D. dissertation. Department of Linguistics, Cornell University.
- Pittayaporn, Pittayawat. 2009b. Proto-Southwestern Tai revised: a new reconstruction. *Journal of the Southeast Asian Linguistics Society*, 2: 121-144.
- Poo-Israkij, Orawan. 1985. *The phonology of Tai Yai at Amphoe Mae La Noi, Mae Hong Son Province*. Bangkok: Mahidol University M.A. Thesis.
- Robinson, Edward Raymond. 1994. *Further classification of Southwestern Tai "P" group languages*. Bangkok: Chulalongkorn University M.A. thesis.
- Schliessinger, Joachim. 2003. *Ethnic groups of Laos. Volume 3. Profile of Austro-Thai-speaking peoples*. Bangkok: White Lotus.
- Tingsabadh, Kalaya M.R. 2001. *Thai tone geography. Essays in Tai linguistics* ed. by M. R. Kalaya Tingsabadh and Arthur S. Abramson. Bangkok: Chulalongkorn University Press.
- Wardhaugh, Ronald. 2002. *An Introduction to Sociolinguistics*, 4th edition. Oxford: Blackwell Publishing.
- Wyatt, David K. 2003. *Thailand: A short history*. 2<sup>nd</sup> edition. New Haven: Yale University Press.
- Yip, Moira. 2002. *Tone*. Cambridge, UK: Cambridge University Press.