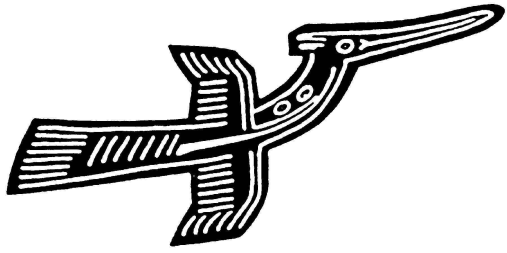
MON-KHMER STUDIES

A JOURNAL OF
SOUTHEAST ASIAN LANGUAGES
AND CULTURES
Volume 39



MAHIDOL UNIVERSITY AT SALAYA, THAILAND
&
SIL INTERNATIONAL

2010

ISBN: 978-1-55671-257-9
ISSN: 0147-5207

Printed in the United States of America

Camera copy prepared by Mahidol University
Research Institute for Languages and Cultures of Asia

Copies of this and other publications of SIL International® may be obtained from

SIL International Publications
7500 West Camp Wisdom Road
Dallas, TX 75236-5699 USA

Voice: 972-708-7404
Fax: 972-708-7363
Email: academic_books@sil.org
Internet: http://www.ethnologue.com

EDITORIAL NOTE

On behalf of the Mon-Khmer Studies Vol. 39, we would like to express our sincere thanks to all authors who have contributed their academic knowledge to this volume. In this volume, there are ten articles, six languages: Vietnamese, Kammu, Old Khmer, Lavɨəʔ, Cua and Tai. The linguistic aspects are rhythmic patterns, body part terms, copula, metaphors, origin of alphabet, phonology, loanwords and sound changes. There is one article on Thai folk healer monks; a book review and conference reports are also presented.

As usual, we welcome your articles on Austroasiatic linguistics and languages as well as studies on cultures for the coming volumes. Please find more information on our webpage at www.mksjournal.org.

Sophana Srichampa
For the MKS Editorial Board

MON-KHMER STUDIES is a journal devoted to the study of Austroasiatic languages, and the cultures of their Mon-Khmer and Munda speakers.
MON-KHMER STUDIES is produced annually, and welcomes articles or notes on linguistics, cultural description, comparison, bibliography, historical development, sociolinguistics, stylistics, orthography, paleography, and other relevant topics.
MON-KHMER STUDIES may on occasion publish articles that concern other languages, and which advance our understanding of the nature, history, and development of the Austroasiatic family.

Editorial Address:
MON-KHMER STUDIES
Research Institute for Languages and Cultures of Asia
Mahidol University at Salaya, Nakhon Pathom 73170, Thailand
TEL 66 2 800-2308-14 ext. 3213
FAX 66 2 800-2332
E-mail: monkhmerj@gmail.com
Homepage: www.mksjournal.org

TABLE OF CONTENTS

ABBREVIATIONS

AE	*American Ethnologist*
AL	*Anthropological linguistics*
ALH	*Acta Linguistica Haniensia*
BEFEO	*Bulletin de l'Ecôle Françoise d'Extrême-Orient*
BLS	*Berkeley Linguistics Society*
CLAO	*Cahiers de linguistique-Asie Oriental*
CTT	*Consonant Types & Tones*
IJSL	*International Journal of Scottish Literature*
IL	*Indian Linguistics*
IPLS	*Indo-Pacific Linguistics Studies*
ISCA	*International Speech Communication Association*
JASA	*Journal of the Acoustics Society of America*
JP	*Journal of Phonetics*
JSAS	*Southeast Asia Studies*
JSEALS	*Journal of the Southeast Asian Linguistics Society*
JSS	*Journal of Siam Society*
LS	*Language Sciences*
MKS	*Mon-Khmer Studies*
OL	*Oceanic Linguistics*
PL	*Pacific Linguistics*, Canberra
PSJ	*Phonetic Society of Japan*
SALS	*Southeast Asia Linguistics Studies*
UHL	*Universals of Human Language*

Rhythmic pattern and corrective focus in Vietnamese polysyllabic words

Nguyễn, T. ANH-THƯ
Australia

Abstract

This paper reports a study which examined the rhythmic patterns in Vietnamese polysyllabic words and their acoustic correlates, particularly the scope of forward planning, reflected in rhythmic prominences and edge/boundary effects; modified by or in interaction with (corrective) focus effects. The aim of the study is pursued by examining the rhythmic patterns and their acoustic correlates in polysyllabic reduplicative words (2-,3-,4-,5-,6-syllable pseudo-words). Ten native speakers of Vietnamese (Saigon dialect) participated in the study. The results showed that there is a tendency of syllable coupling indicated mainly by syllable duration pattern and supported by the native listeners' perception results, suggesting that polysyllabic words in Vietnamese tend to be parsed into bi-syllabic iambic feet with a rightward or retrograde rhythmic pattern, suggesting that bi-syllabic right-headed foot is a prosodic unit above the syllable in Vietnamese.

1. Introduction

Vietnamese is a contour tone language which has a system of lexically distinctive tones and is strongly syllabic in its phonological organization and morphology. Most syllables are independent morphemes and every syllable in an utterance bears an independent lexical tone specification which is not neutralised (become toneless) in context. The existence of stress at the word level in Vietnamese has been a matter of unresolved controversy (Thompson 1965; Nguyễn Đăng Liêm 1970, among others). However, recent series of studies (Nguyễn & Ingram 2007a; Nguyễn & Ingram 2007b; Nguyễn & Ingram unpublished) showed clear evidence of prominence asymmetry; that bi-syllabic reduplicative and compound words are phonetically right-headed, realised mainly by syllable duration, full vowel and tonal shape. This prominence pattern is further supported by the tone sandhi which is confined to first syllables, suggesting that "tone sandhi is a reduction phenomenon occurring on prosodically weak positions" (Shih 2005) and, by implication, that Vietnamese shows phonetic evidence of prosodic constituency at the level of the bisyllabic word. However, the status of the prosodic unit – whether it constitutes a stress foot or not – is yet to be determined. The aim of this study was to further follow up the status of word stress/stress foot and the rhythmic structure of Vietnamese, particularly the scope of forward planning, reflected in rhythmic prominences and edge/boundary effects; modified by or in interaction with corrective focus effects.

Foot is a prosodic category of sound organization (Hayes 1985, 1995; Nespor and Vogel 1986; McCarthy and Prince 1990, 1996; Halle and Vergnaud 1987, and others). Feet represent the rhythmic structure of the word; words are made up of rhythmic units called feet. Each foot has a unique head (its strong, or only syllable), and optional weak syllables (Kager 1995). A foot can comprise just a single word or a group of words. There are two kinds of feet; left-dominant and right-dominant. Languages use either one or the other type. Left-dominant feet have a strong first syllable with the following syllables weak. Right-dominant feet have a strong final syllable with preceding syllables weak. English is a left-dominant language. For example, "consultation" has two feet, /kɔn.səl/ and /tæɪ.ʃən/. In each of these feet, the first or left-most syllable is strong and the second is weak, that is, left-dominant. In stress languages, prosody is often represented with the foot headed by a stressed syllable (Hayes 1995). Pitch accent languages have been represented with a high pitch associated with the head of a foot (Hayes 1995). In a tone language such as Mandarin, the foot is built on the basis of the tonal structure of syllables. There is a constraint in Mandarin, as well as in several other dialects examined by Yip (1980) that allows only one fully toned syllable per foot. In a foot with more than one syllable, only the first bears tone; the others are toneless. Under certain circumstances, a foot may consist of two originally toned syllables, but in these cases a tone-deletion rule applies to make the second syllable toneless in conformity with the general constraint on foot structure. Vietnamese has no toneless syllables like Mandarin; therefore, it is wondered how the rhythmic patterns are represented in Vietnamese polysyllabic words (as well as in longer stretches of speech). The empirical findings in our recent studies show that bi-syllabic reduplicative and compound words are phonetically right-headed, realised mainly by syllable duration, full vowel and tonal shape (Nguyễn & Ingram, 2007a, 2007b). A recent perceptual study showed that word stress or prosodic asymmetry at the level of the bisyllabic foot is merely a phonetic tendency in Vietnamese; a 'sub-phonological' threshold phonetic effect originating in rhythmic or metrical tendencies at the level of post-lexical phonology or perhaps the level of 'motor programming' or speech gesture co-ordination and control (Nguyễn & Ingram, unpublished). On the basis of previous researchers' observations about rhythmic patterns (Jones and Huỳnh 1960, Trần 1969, review below) and our recent empirical findings, it is predicted that there will be a consistent alternation of strong and weak syllables represented by syllable duration, intensity and/or F0 shape such that the weak syllables will have shorter duration, weaker intensity and/or less fully realised tones in compared with the strong ones.

The aim of the study is pursued by examining the rhythmic patterns and their acoustic correlates in polysyllabic reduplicative words (2-,3-,4-,5-,6-syllable pseudo-words). The aim of the study is two fold:

(1) To examine the rhythmic patterns in Vietnamese polysyllabic words and their acoustic correlates. We predict to find a consistent alternate right-headed prominence pattern in these polysyllabic words which will be realized mainly by syllable duration, intensity and/or F0 shape.

(2) To investigate the acoustic correlates of corrective focus on subcomponents of four-syllable words (S1, S2, S3, S4, S1S2, S3S4 and the whole word) to examine how the rhythmic pattern is modified by corrective focus (i.e. focus elicited by subjects' correction of the error prompt produced by the experimenter). It is predicted that corrective focus will induce syllable lengthening, F0 range expansion, F0 height enhancement, more fully realised tonal shape/contour and increased intensity. In terms of durational patterns, there will also be a spill-over effect such that there is lengthening not only on the pragmatically focused syllable but also on the neighboring, particularly immediately following syllable such as in other languages (Mandarin: Chen 2006; English: Turk and White 1999; Cambier-Langeveld and Turk 1999).

1.1 Stress pattern in Vietnamese

Vietnamese has no system of culminative word stress; nevertheless, it is widely accepted that there is stress in the sense of accentual prominence at the phrasal level (Thompson 1965; Nguyễn Đăng Liêm 1970). Duration, intensity, full tonal realisation of accented syllables have been observed to be important parameters for describing stress in Vietnamese (Đỗ 1986; Chaudhary 1983; Hoàng & Hoàng 1975; Gsell 1980). Generally, these studies examined stress in the sense of phrasal-level accentual prominence but do not address the question of stress at the word level. Regarding the stress patterning in utterances, it is generally agreed by some researchers that there is an alternating pattern of strong and weak syllables. Thompson (1965) stated that the majority of the syllables have medium stress. In a sequence of syllables, alternating ones are slightly louder (but not in a distinctive manner): "each pause group has at least one heavy stress and weak stresses are fairly frequent in rapid passages, rarer in carefully speech" (p. 50). Jones and Huỳnh (1960) stated that "normally the stresses in a Vietnamese utterance are conditioned by the junctures," and regarded the fundamental stress pattern of Vietnamese as consisting of the alternating occurrence of a strong and weak stress, with the last word of the phrase receiving a strong stress. According to Trần (1969), "the stress rhythm of Vietnamese phrases showed a retrograde pattern, i.e., it is the last syllable before the pause which determines the stress pattern of the entire group. In a basic or regular stress pattern, the strong and weak stresses alternate in the pause group" (p. 109). However, in contrast with Thompson (1965) and Jones & Huỳnh (1960), who postulated that there is only one fundamental stress pattern which starts regressively with a strong stress, Trần's analysis shows two fundamental patterns of stress in the Vietnamese pause group, i.e. one which ends with a weak stress and one which ends with a strong stress. Whether the pattern begins (in terms of the retrograde pattern) with a strong or weak stress, depends on the grammatical structure of the utterance, particularly the syntactic nature of the syllable which occurs immediately before the pause. Nevertheless, all of these studies were based on impressionistic investigation rather than on experimental data and instrumental analysis, which this study relies on.

A more systematic durational pattern has been found in other languages such as Mandarin. Xu (1999) found that a disyllabic word in Mandarin shows a clear short-long duration pattern regardless of whether it is

focused or not, or whether it is utterance-initial or utterance-final. Chen (2006) shows that the syllables in a monomorphemic quadrisyllabic word exhibit a 3 1 2 4 duration pattern (larger number indicating longer duration) in the utterance final position, but a 4 1 2 3 pattern in an utterance medial position. Xu and Wang (2009) found similar results: in a multi-syllabic unit the final syllable is the longest, while the first syllable is the second longest. The first medial syllable is always the shortest, while the second medial syllable is the second shortest. These patterns pose problems for all of the previous theories that assume a strong-weak pattern for Mandarin (Duanmu 2000; Feng 1998) if duration is assumed to be correlated with stress. That is, if a four-syllable group is assumed to consist of two disyllabic feet, then the first and second feet would have opposite stress/strength patterns no matter what kind of stress-duration relation is assumed. Xu and Wang (2009) also found that syllable grouping in Mandarin is encoded directly by syllable duration, while the corresponding variations in F0 displacement, vp/d ratio and velocity profile are the consequences of duration control.

1.2 Acoustic correlates of focus

Studies on focus in stress-accent languages such as English and Dutch found that while an increase in both duration and intensity are observed to accompany focus, F0 is the most reliable acoustic realization of focus (see Sluijter (1995) and references therein). In Mandarin, Chen (2001) found that all three acoustic cues Fo, overall intensity, and duration are employed, with different degrees of consistency, to convey contrastive focus. In regards to phonetic cues of contrastive or corrective accentual focus in Vietnamese, some authors, such as Hoàng and Hoàng (1975), or Gsell (1980) consider that full tonal realization of accented syllables is one of the positive marks of prominence (accent) at phrasal level in Vietnamese. In a recent study on the effect of emphasis on glotalised and nonglotalised Vietnamese tones (the Hanoi creaky falling tone (i.e. the nặng tone) in obstruent vs. sonorant final consonant environment respectively), Michaud and Vũ (2004) found that in Vietnamese emphasis, syllable lengthening appears as a speaker-dependent variable, whereas a stable correlate of emphasis is curve amplification, manifested as increased slope of Fo curve or as Fo register raising.

1.3 Domain of accentual lengthening under focus

In terms of the domain of the lengthening effect under focus, Chen (2006) examined patterns of durational adjustment of mono-morphemic four-syllable words when different constituents of the word are focused for correction in Mandarin. Her results showed that the domain of lengthening is the constituent under focus. When a focused domain is multi-syllabic, the distribution of lengthening is non-uniform: there is a strong tendency toward an edge effect with the last syllable lengthened the most. There is also a spill-over lengthening effect such that there is lengthening not only on the pragmatically focused syllable but also on its neighbors, particularly the immediately following syllable. The magnitude of such lengthening is conditioned by prosodic boundaries in that word boundaries attenuate lengthening more than syllable boundaries.

In general, the duration patterns reported for Mandarin by Xu (1999), Chen (2006) and Xu and Wang (2009) bear certain resemblance to those reported for English, including constituent-initial lengthening and constituent-final lengthening effect (Hirst & Bouzon 2005; Turk and Shattuck-Hafnagel 2000), and a spill-over lengthening effect (in English: Turk and White 1999; in English and Dutch: Cambier-Langeveld and Turk 1999). It raises the question of how the duration pattern is manifested under focus in Vietnamese and how the rhythmic pattern is modified by corrective focus.

2. Method

2.1 Linguistic materials

The use of nonsense syllables in the study of prosodic phenomena, so-called "reiterant speech", has been developed to circumvent difficulties of segmental variations throughout the course of an utterance (Liberman 1978). Reiterant speech can be a powerful and effective tool for prosodic research (Liberman 1978; Liberman and Streeter 1978; Nakatani and Shaffer 1978; Nakatani, O'Connor and Aston 1981; Larkey 1983; Gandour, Potisuk and Perkins 1997). By using the same syllable everywhere in a word or sentence, prosodic regularities are only subject to the influence of factors such as stress and constituent structure (Gandour, Potisuk and Perkins 1997). For this reason, nonsense words with segmentally similar syllables are used in this study. A list of 2-,3-,4-,5-,6- syllable pseudo foreign and ethnic place names was constructed. All syllables in the words have the same rhyme and tone but different onsets; thus they can be considered to be reduplication with onset change. The onsets and codas are all voiced consonants. There were 5 words carrying 5 different tones[1] (level, rising, falling, drop, curve) in each syllable number set (2-syllable, 3-syllables, 4-syllable, 5-syllable sets), while in the 6-syllable set, there were only two words carrying level and falling tone. Within each word the same tone was repeated. The words were embedded in a constant carrier sentence (e.g., Tôi đi *La-na-ma* về [I came back from La-na-ma]).

2 syllables	3 syllables	4 syllables	5 syllables	6 syllables
La-na Màng-nhàng Lí-wí Nậm-rậm Nhổ-nổ	La-na-ma Màng-nhàng-đàng Lí-wí-rí Nậm-rậm-nhậm Nhổ-nổ-mổ	La-na-ma-ra Màng-nhàng-đàng-làng Lí-wí-rí-mí Nậm-rậm-nhậm-mậm Nhổ-nổ-mổ-lổ	La-na-ma-ra-ga Màng-nhàng-đàng-làng-nàng Lí-wí-rí-mí-ní Nậm-rậm-nhậm-mậm-lậm Nhổ-nổ-mổ-lổ-rổ	La-na-ma-ra-ga-nha Màng-nhàng-đàng-làng-nàng-gàng

[1]Note: Northerners have 6 tones, but the Southern dialect has 5 due to the merging of two contour tones (hỏi and ngã)

Corrective focus was elicited on 4-syllable words. Focused syllables were capitalized. The focus condition was elicited by subjects' correction of the error prompt in the experimenter's question. Focus was induced on one syllable, two syllables or the whole word (S1, S2, S3, S4, S1S2, S3S4 and the whole word). For example,

Subject: Tôi đi *la-na-ma-ra* về (I came back from la-na-ma-ra) **Unfocus condition**

Experimenter: Anh đi *TU-na-ma-ra* về a`? (Did you come back from TU-na-ma-ra?).

Subject: Không, Tôi đi ***LA****-na-ma-ra* về. (No, I came back from LA-na-ma-ra) **One syllable focus**

Experimenter: Anh đi *TU-LE-ma-ra* vê` a`?

Subject: Không, Tôi đi ***LA-NA****-ma-ra* về. **Two syllable focus**

Experimenter: Anh đi *TU-LE-KI-DU* vê` a`?

Subject: Không, Tôi đi ***LA-NA-MA-RA*** vê`. **Whole word focus**

2.2 Subjects

Ten speakers of the Saigon dialect (5 males, 5 females) who came from HoChiMinh city participated in the study. They were students at the University of Queensland and had been in Australia from 1-5 years.

2.3 Procedures

Subjects were given the list of pseudo foreign and ethnic place names to practice before the recording. They were asked to speak [to the experimenter] the sentence *Tôi đi "place name" về* (I came back from ……) three times. The focus condition was elicited by subjects' correction of the error prompt in the experimenter's question

2.4 Measurement

Three likely acoustic correlates of prominence (syllable duration, intensity, f0 range and contour) were measured.

The test items were segmented via the Emu Speech Tools, (Cassidy 1999). First, the Emu Labeller was used to mark the edges of the target syllables and vowels, relying primarily on the spectrographic display in the Labeller. Then the Emu-R statistical software was used to extract segment duration (ms), and fundamental frequency (Hz). Peak intensity (dB) in syllables was measured manually via Praat (Boersma and Weenink, 2007).

2.5 Analysis

There were in total 150 words in each syllable number set (2-syllable, 3-syllable, 4-syllable, 5-syllable data set): 5 words x 3 repetitions x 10 speakers =150 words. In 6-syllable set there were 60 words: 2 words x 3 repetitions x 10 speakers = 60 words

Mixed effect ANOVAs were used to statistically analyze the data for each syllable number set. The dependent variables were duration and F0 range. The fixed effects were syllable position (and tone or focus condition). The random effects were speaker (n=10) and word (n=5). Results on F0 was examined both quantitatively (F0 range) and qualitatively (F0 contour). The analysis specific to each variable is presented separately at the beginning of each result section.

3. Results

3.1 Results on acoustic correlates of nonfocus words

This section presents the results of acoustic correlates of nonfocus words.

Duration

A series of one-way mixed effect ANOVAs were conducted on each syllable position number set. The dependent variable was syllable duration (ms). The fixed effect was syllable and the random effects were speaker (n=10) and word (n=5). Tukey posthoc tests were carried out to determine the significant differences among levels of the main factor of syllable position (e.g., S1 vs. S2 in 2-syllable set, S1 vs. S2 vs. S3 in 3-syllable set, S1 vs. S2 vs. S3 vs. S4 in 4-syllable set, and so on). The result is presented in figure 1

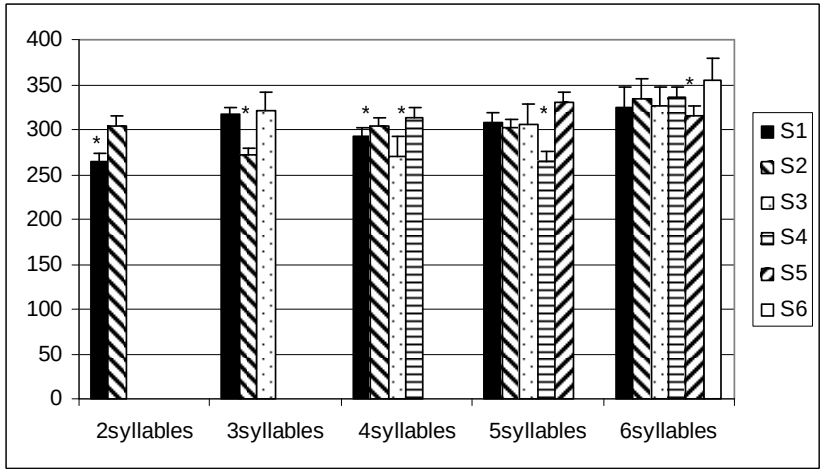


Figure 1. Mean syllable duration (ms) by positions and syllable-number sets. *: p<.01

One way ANOVA result on 2-syllable set shows a significant effect: $F(1, 273)=101.6$, $p<.0001$. The second syllable was significantly longer than the first syllable (S1<S2). There was a significant effect on syllable position

for the 3-syllable set: F(2, 382)=37.72, p<.0001. In three-syllable words, the first and the third syllables were significantly longer than the second syllable (S1~S3 >S2, p<.01). The ANOVA result on 4-syllable set showed a significant effect: F(3, 569)=24.74, p<.0001. Posthoc test showed that the second and the fourth syllables were significantly longer than the first and the third syllables (S2>S1, S4>S3, p<.01) and the last syllable was the longest, S4~S2>S1, S3, p<.01). There was a significant effect on syllable position for the 5-syllable set: F(4, 577)=19.68, p<.0001. Posthoc pairwise comparison among syllables showed that the fourth syllable was the shortest and the fifth syllable was the longest while there was no significant difference among the first three syllables (S5>S1~S2~S3>S4, p<.01)**.** The ANOVA result on 6-syllable set showed a significant effect F(5, 200)=4.06, p<.002. Posthoc comparison among the 6 syllables showed that the fifth syllable was the shortest and the sixth syllable was the longest while there were no significant differences among the first four syllables (S6>S1~S2~S3~S4>S5). Nevertheless, as shown in the graph, there is a tendency of even syllables being stronger, i.e. S2, S4, and S6 are longer than the counterpart odd syllables (S1, S3, S5).

In general, the results on syllable duration showed a strong right-headedness pattern with the last syllable of the word consistently having the longest duration and the penult syllable the shortest. There is also a strong tendency of syllable coupling headed by even syllables as shown in the 4-syllable and 6-syllable sets (1 **2**/ 3 **4**/ and 1 **2**/ 3 **4**/ 5 **6**/); i.e. the even syllables were longer (and thus more acoustically prominent) than the counterpart odd syllables. This is further supported by the results on onset duration which showed the effect of initial constituent strengthening and boundary marking as illustrated in figure 2 below. The odd syllable in 4-syllable set (the first and the third) and 6-syllable set (the first, the third and the fifth) tended to have longer onsets, suggesting a marking of initial boundary effect after a stress foot.

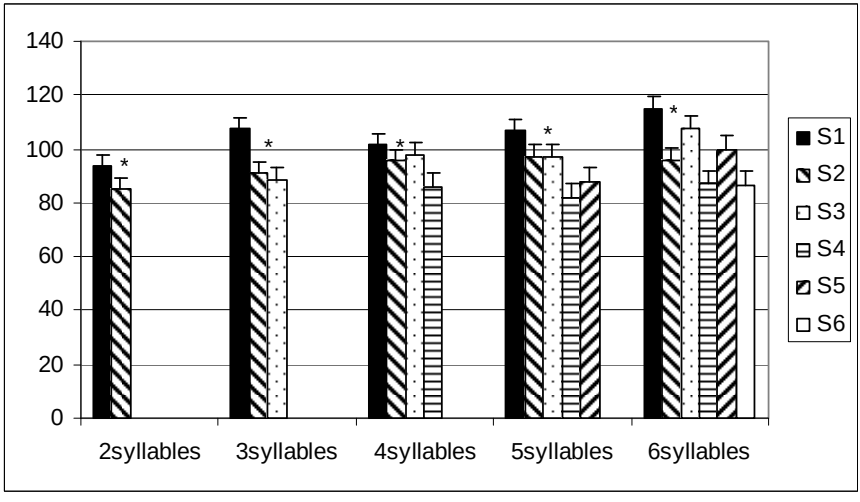


Figure 2. Mean onset duration (ms) by positions and syllable-number sets. *: p<.01

Intensity

One way ANOVA result on the 2-syllable set shows no significant effect: F(1, 273)=0.76, p=0.3 ns even though the second syllable had higher intensity mean than the first syllable (S1<S2). There was a significant effect on syllable position for the 3-syllable set: F(2, 412)=5.58, p<.01. In three-syllable words, the first and the third syllables had significantly higher intensity than the second syllable (S1~S3 >S2, p<.01). The ANOVA result on the 4-syllable set showed a significant effect: F(3, 588)=15.08, p<.0001. Posthoc tests showed that the third syllable had significantly lower intensity than the other syllables (S1~S2~ S4>S3, p<.01). There was a significant effect on syllable position for the 5-syllable set: F(4, 687)= 3.67, p<.01. Posthoc pairwise comparison among syllables showed that the fourth syllable had the lowest intensity (S1~S2~S3~S5>S4). The ANOVA result on the 6-syllable set showed a weak significant effect (F(5, 199)=1.43, p<.05). Posthoc comparisons among the 6 syllables showed that the fifth syllable had the lowest intensity and the first syllable had the highest intensity (S1>S2~S3~S4~S6>S5).

In general, the results on intensity showed a different pattern from duration results. The only similarity is that the penult syllables tend to have the lowest intensity, consistent with duration result: the penult syllable has the shortest duration. This suggests a right-headedness tendency with the weak penult syllable as a result of bisyllabic foot compensatory effect before the final strong syllable.

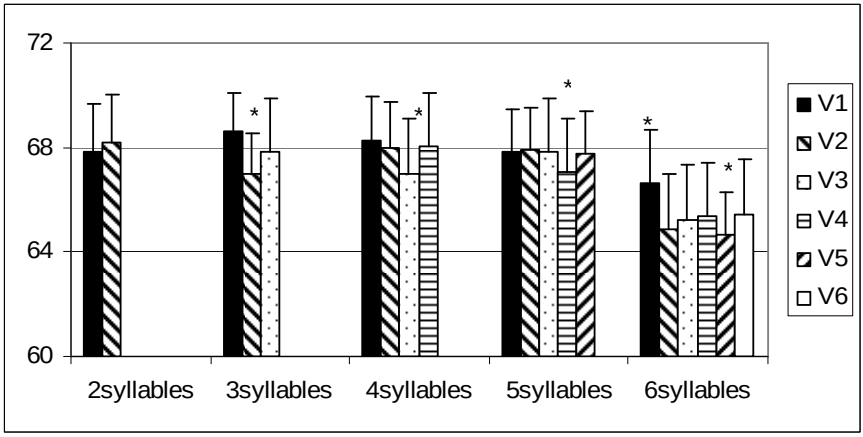


Figure 3. Mean intensity (dB) by positions and syllable-number set. *: p<.01 on top of a syllable that is significantly different from others.

F0 range

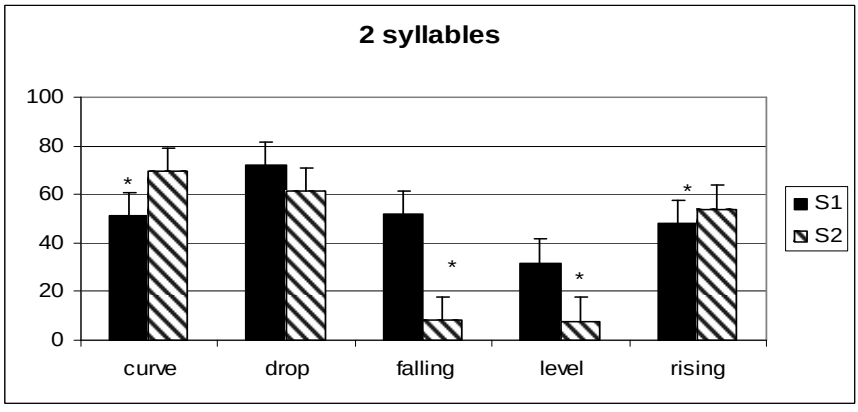
A series of mixed effect two-way ANOVAs were conducted on F0 range values. The fixed effects were syllable position and tone (curve, drop, falling, level and rising) and random effects were speaker (n=10) and word

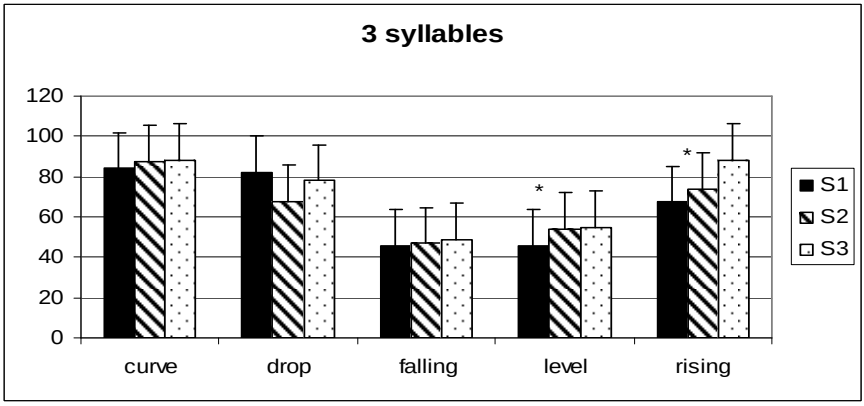
(n=5). The results showed significant effect for the main factor of syllable position and tone and the interaction of position x tone as summarized in table 1. Posthoc test by Tukey method was then conducted to examine the interaction position x tone effect. Results are presented in figure 4.

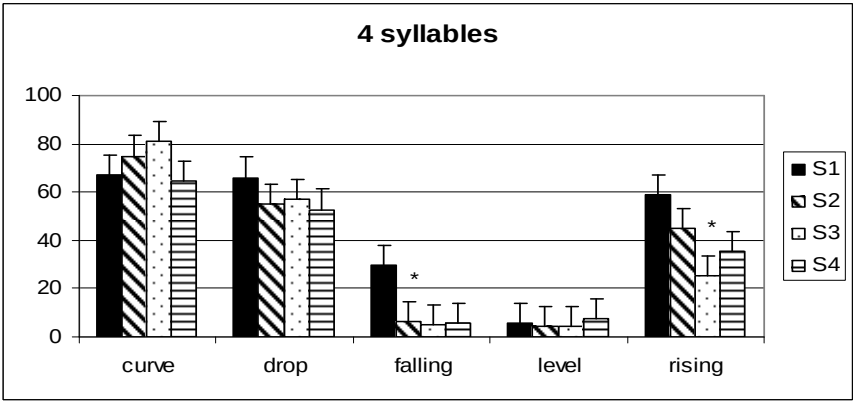
Table 1. ANOVA results on F0 range.

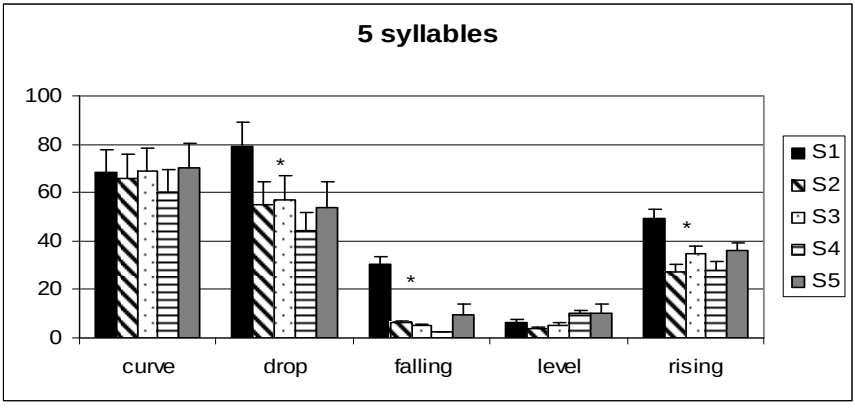
Syllable sets	Factor	Sig. level
2 syllables	position	F(2, 407)=6.21, p<.01
	tone	F(4, 407)=29.9, p<.0001
	position*tone	F(8, 407)=4.04, p<.0001
3 syllables	position	F(3, 552)=5.3, p<.01
	tone	F(4,552)=56.8, p<.0001
	position*tone	F(12,552)=2.5, p<.001
4 syllables	position	F(3, 552)=5.3, p<.01
	tone	F(4,552)=56.8, p<.0001
	position*tone	F(12,552)=2.5, p<.001
5 syllables	position	F(4, 671)=5.5, p<.001
	tone	F(4, 671)=91.9, p<.0001
	position*tone	F(16, 671), p<.01
6 syllables	position	F(5, 210)=26.9, p<.0001
	tone	F(1, 210)=21, p<.0001
	position*tone	F(5, 210)=22.6, p<.0001

As shown in figure 4 below, in 2-syllable words, the second curve and rising tones had a larger F0 range while the three falling tones (level, falling and drop) had a larger F0 range on the first syllable. In 3-syllable words, the last syllables seem to have the largest F0 range. In the 4-syllable set, the first and the last syllable of the falling, level and rising tones had a larger F0 range than the other syllables. A similar effect was found for the 5-syllable set that the first and the last syllable seemed to have a larger tone range than the other syllable. In the 6-syllable set, the first syllable had a larger tonal range than the other five syllables. In general, the F0 range result is not very strong and has no consistency with duration or intensity results.




3 syllables
120
100
80
60
40
20
0
curve
drop
falling
level
rising
S1
S2
S3
*
*


4 syllables
100
80
60
40
20
0
curve
drop
falling
level
rising
S1
S2
S3
S4
*
*


5 syllables
100
80
60
40
20
0
curve
drop
falling
level
rising
S1
S2
S3
S4
S5
*
*
*

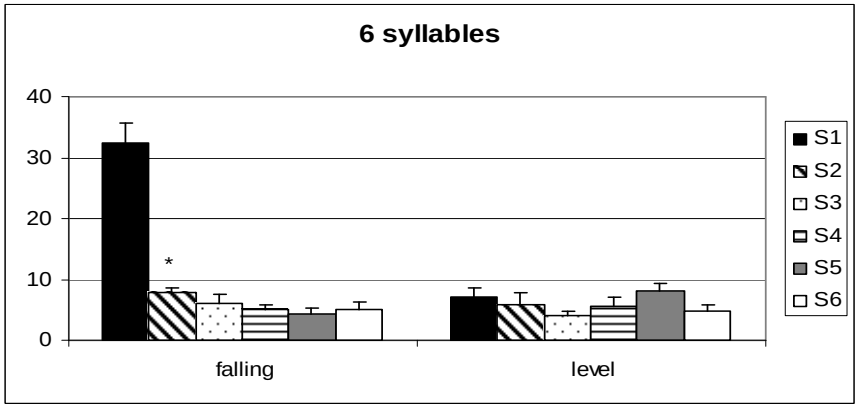
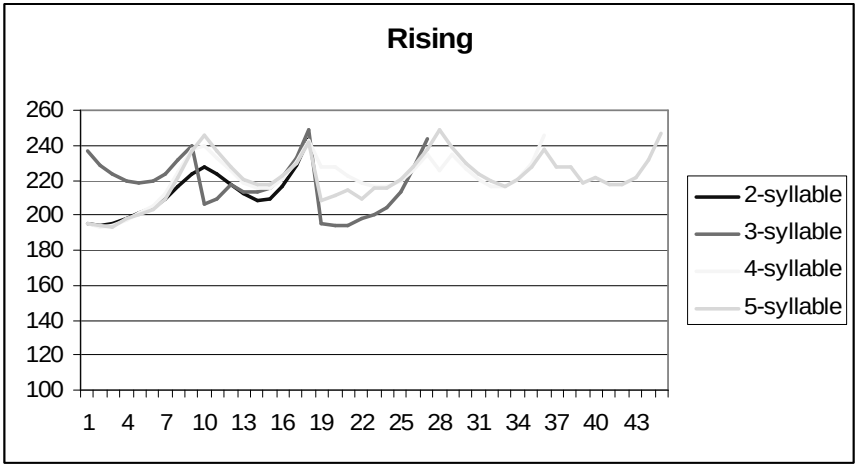
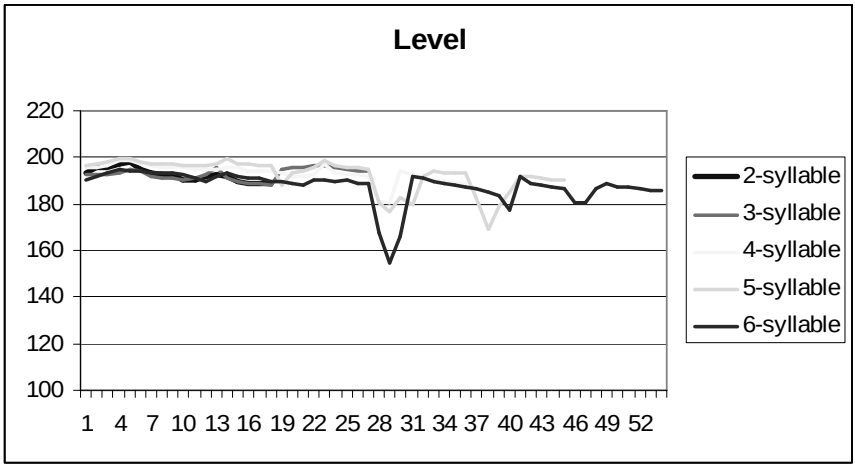


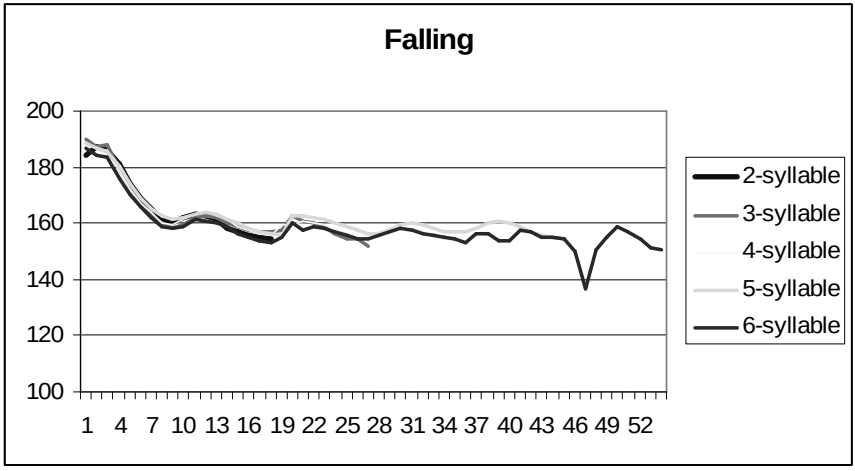
Figure 4. Mean F0 range (Hz) by tone and syllable position. *: sig. at p<.01

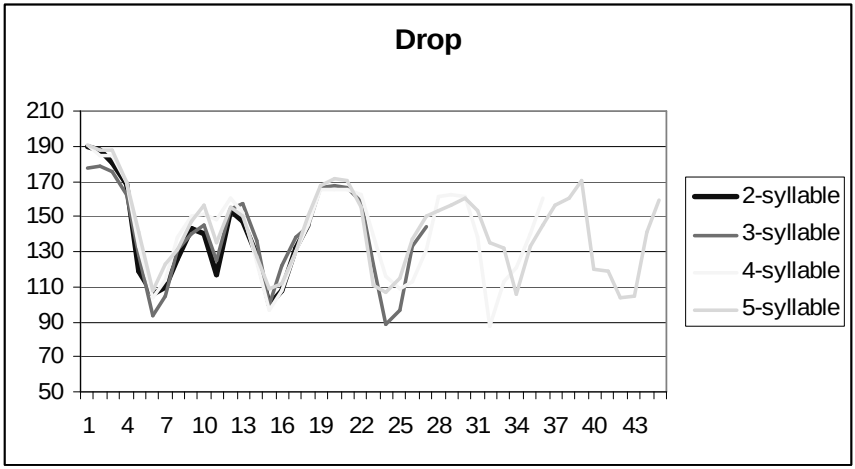
F0 contour

As shown in figure 5, generally the F0 contours do not show any consistent results with either duration or intensity in terms of alternating prominence effect but only a carryover tonal co-articulation effect such that the following tones tended to start from where the preceding tone ended (for the falling and level tones) or had a further lower or heightened F0(rising tone). For example, in sequences of rising tones, the following tone had higher F0 and rose higher. By contrast, in sequences of falling or level tones, the following tones had lower F0 and fell lower. In sequences of curve and drop tone, the following tones tended to have lower F0 valleys.




Level
220
200
180
160
140
120
100
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52
2-syllable
3-syllable
4-syllable
5-syllable
6-syllable


Falling
200
180
160
140
120
100
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43 46 49 52
2-syllable
3-syllable
4-syllable
5-syllable
6-syllable


Drop
210
190
170
150
130
110
90
70
50
1 4 7 10 13 16 19 22 25 28 31 34 37 40 43
2-syllable
3-syllable
4-syllable
5-syllable

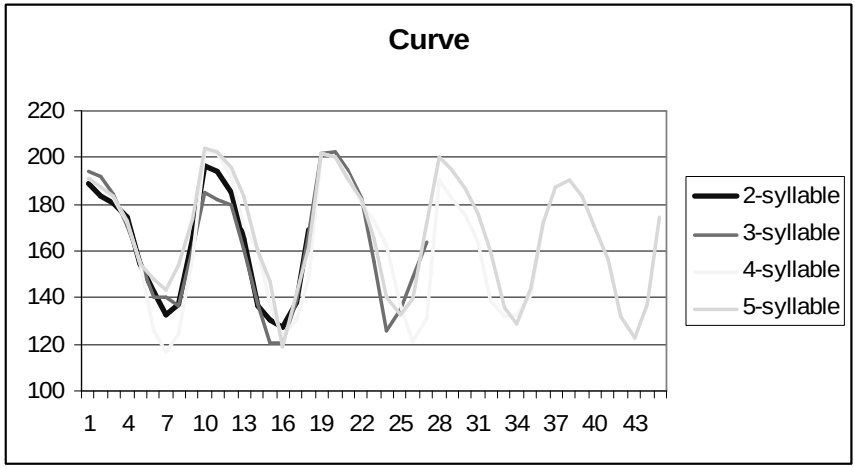


Figure 5. Mean F0 contour by tone types and syllable number sets. 1-9: first syllable, 10-18: second syllable 18-27: third syllable, 27-36: fourth syllable, 36-45: fifth syllable, 45-54: sixth syllable.

Discussion

In summary, the results on intensity and Fo of nonfocus words are not very strong and do not show any consistent patterns. However, the duration results show consistent alternating rhythmic patterns. This is consistent with Xu and Wang's results (2009) in Mandarin that syllable duration had the most consistent grouping-related patterns. That is, syllable grouping in Mandarin is encoded directly by syllable duration, while the corresponding variations in F0 displacement, vp/d ratio and velocity profile are the consequences of duration control. Maybe being a tone language in which F0 is involved in tonal contrasts, Vietnamese makes less use of Fo (and intensity) as cue to rhythmic stress. Instead, rhythmic stress is mainly realized by duration cue. For example, it is shown in Piraha, a language which relies on F0 to make lexical contrasts, that intensity and duration are the primary acoustic correlates of stress (Everett 1998). In Mandarin Chinese, duration was shown to be a more important cue to sentence level stress than intensity while F0 had no effect on stress perception (Shen 1993). The duration results of this study show two main things. First, there is a tendency for the syllable coupling with the even syllables (e.g. S2, S4, S6) being longer (and thus stronger) than the odd syllables (e.g. S1, S3, S5, figure 1), suggesting that mono-morphemic words in Vietnamese tend to be parsed into bi-syllabic iambic feet. When there is only one syllable such as in the case of odd syllable in three-syllable and five-syllable words (e.g., the first syllable in 3-syllable and the first and third syllable in 5-syllable words, figure 1), the odd syllable tended to be lengthened to fill the bi-syllabic foot template.

Second, Vietnamese polysyllabic words are right-headed with the last syllable always the longest syllable in the word, consistent with findings in bi-syllabic coordinative compound words (Nguyễn & Ingram 2007a) and in bi-syllabic reduplications (Nguyễn & Ingram 2007b) and support observations by previous researchers (Thompson 1965; Jones and Huỳnh 1960). The following rhythmic and prominent pattern can be predicted on the basis of duration patterns (the slash indicates the rhythmic beat, larger number indicates longer duration):

2-syllable words	1	**2/**					w	**s/**				
3-syllable words	**2/**	1	**3/**				**s/**	w	**s/**			
4- syllable words	2	**3/**	1	**4/**			w	**s/**	w	**s/**		
5-syllable words	2	**3/**	**3/**	1	**4/**		w	**s/**	**s/**	w	**s/**	
or	**3/**	2	**3/**	1	**4/**		**s/**	w	**s/**	w	**s/**	
6-syllable words	2	**3/**	2	**3/**	1	**4/**	w	**s/**	w	**s/**	w	**s/**

These predicted rhythmic patterns need to be investigated in a perception test which is the aim of the next experiment

3.2 Perception of rhythmic pattern

In order to test the rhythmic patterns predicted on the basis of the duration patterns, a small perception test was carried out. Five Vietnamese listeners (two male and three female listeners) listened to the whole sets of 3-syllable, 4-syllable, 5-syllable and 6-syllable words and put a slash after the syllable where they thought the rhythmic beat fell. For example

1. LA/ MA NA/
2. LA NA/ MA RA/
3. LA NA/ MA/ RA GA/
4. LA NA/ MA RA/ GA NHA/

The experiment was carried out in a quiet room. Each word embedded in its carrier sentence was played twice from a laptop computer. Listeners were instructed to pay attention to the rhythm within the target word. The result is presented in terms of the agreement rate among the five listeners, that is, the agreed patterns were those that all five listeners marked. The results are reported in table 2 below.

Table 2. Perception results: percentage of agreed rhythmic patterns by syllable number sets.

Syllable sets	Rhythmic patterns	percentage
3-syllable	**LA**/ NA **MA**/	82.70%
	la na/ ma/	17.3
	la/ na/ ma/	
4-syllable	LA **NA**/ MA **RA**/	96%
	la/ na ma/ ra/	4%
	la/ na/ ma ra/	
5 syllables	LA **NA**/ **MA**/ RA **GA**/	74%
	la/ na ma/ ra ga/	26%
	la/ na/ ma/ ra ga/	
	la na/ ma ra/ ga/	
	la na ma/ ra ga/	
6 syllables	LA **NA**/ MA **RA**/ GA **LA**/	90%
	la na ma/ ra ga la/	10%
	la/ na ma/ ra/ ga la/	

Table 2 showed the percentage of rhythmic patterns per syllable set as agreed by five listeners. 82.7% of the 3-syllable words have the rhythmic pattern of **s**/ w **s/**, consistent with the syllable duration pattern and the prediction. In four-syllable words, as predicted, the pattern of w **s**/ w **s**/ accounted for 96%. In five-syllable words, the pattern w s/ s/ w s/ accounted for 74%, while 26% were the other patterns. In six-syllable words, 90% of the words have the patterns of w s/ w s/ w s/. Generally, the perception results support the prediction and are consistent with the duration patterns, indicating that the rhythmic pattern in Vietnamese polysyllabic words realized by duration cues can be perceived by native listeners.

3.3 Results on corrective focus

Duration

A mixed effect two-way ANOVA was conducted on the syllable duration. The fixed effects were syllable position (S1, S2, S3, S4) and focus (on S1, S2, S3, S4, S1S2, S3S4, and word). The random effects included speakers (n=10) and word (n=5). The results showed a significant effect for the main effect of syllable position: $F(3, 1875)=120.8$, $p<.0001$ and the interaction position x focus: $F(21, 1875)=30.7$, $p<.0001$ while the main effect of focus was insignificant: $F(7, 1875)=0.73$, $p=0.6$, ns. A posthoc Tukey test was conducted to examine the interaction effect position x focus. The results are reported in figure 6.

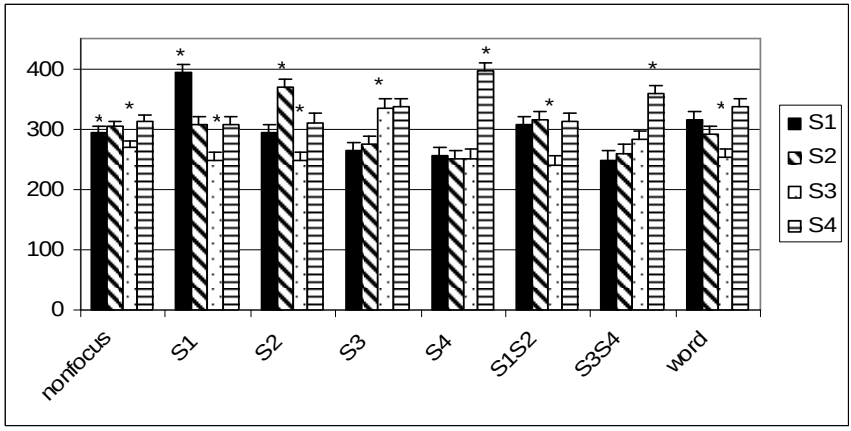


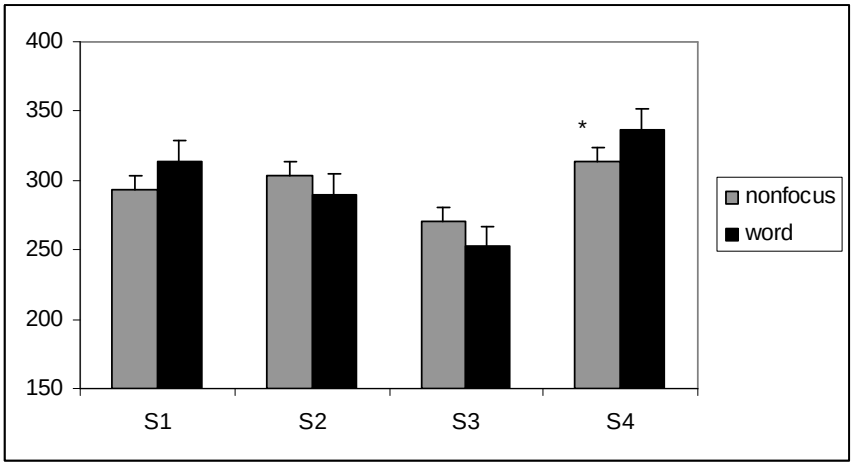
Figure 6. Mean syllable duration (ms) by syllable and focus condition. *: significant at p<.01 on top of syllable that is significantly different from others

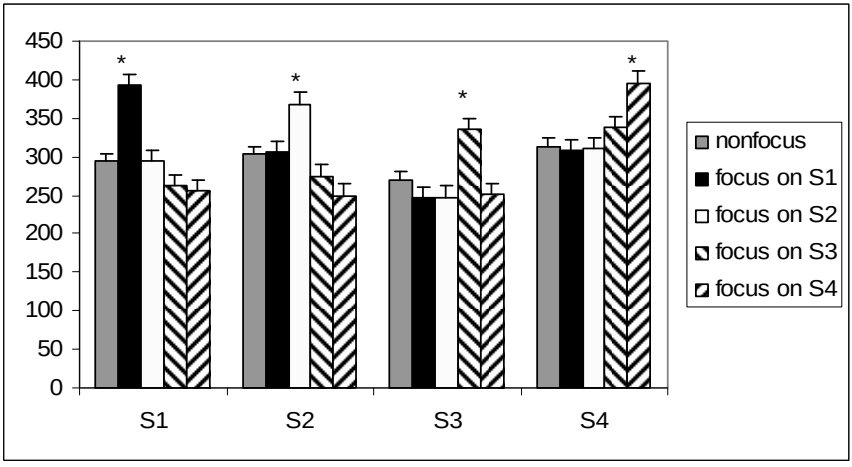
As shown in figure 6, when focus was on the first syllable (S1), the first syllable was lengthened and longest (S1>S2~S4>S3). When focus was on the second syllable, it was the longest syllable in the word (S2>S1~S4>S3). Similarly, the third and fourth syllables got lengthened and became the longest when they were focused (Focus on S3: S3~S4>S1~S2, focus on S4: S4>S1~S2~S3). When focus was on the first two syllables S1S2, both syllables were lengthened to the same extent in comparison with the last two syllables (S1~S2~S4>S3). When focus was on the last two syllable S3S4, both syllables were lengthened but the last syllable was the longest in the word (S4>S3>S1~S2). This is probably due to the compounded effect of accentual lengthening, final lengthening and right-headedness marking.

Compared with the nonfocus condition, when focus was on the whole 4-syllable word, the first and the last syllables got lengthened in comparison with the nonfocus condition (figure 7a). When focus was on individual syllables, only the syllables under corrective focus were lengthened and there is no spill-over lengthening effect; that is, only the syllable under focus was lengthened while its neighboring syllables (either prefocus or postfocus) stayed the same in comparison with the nonfocus condition (fig. 7b). When focus was on the first two (S1S2) or the last two (S3S4) syllables, both syllables under focus were lengthened (fig 7c).

Generally, the results in figures 6 & 7 showed three main things: (1) There is a strong right-headedness pattern in which the penult syllable was the shortest and the last syllable of the word the longest. This pattern is preserved even under focused condition. The tendency for the syllable coupling with the even syllables (e.g. S2, S4) being longer than the odd syllables (e.g. S1, S3) is also preserved, consistent with the results on duration of syllable grouping reported in the previous section and suggesting a right-headed retrograde or rightward rhythmic pattern in polysyllabic words. (2) In terms of within-word

syllable position, while both Vietnamese and Mandarin shared the pattern of the longest syllables in final position (Chen 2006), Vietnamese had syllable couplings with right-headedness pattern w s/ w s/ and the shortest syllable in the third position whereas Mandarin had the shortest syllable in the second position, and the first syllable was longer than the third position. (3) Only the constituent under focus was lengthened and there seemed to be no spill-over effect. This contrasts with Mandarin results where "there was corrective focus-induced lengthening not only on the pragmatically focused syllable but also on the neighboring, particularly immediately following syllable" (Chen 2006, p. 186), suggesting language-specific lengthening mechanisms for two different languages.





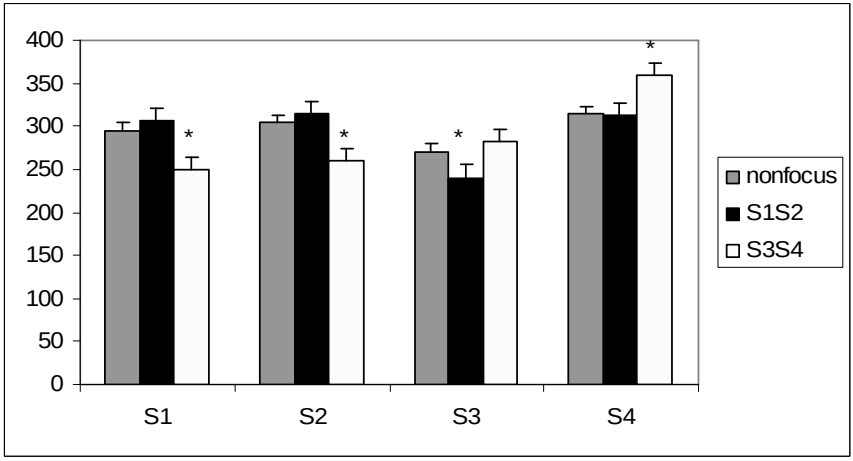


Figure 7. Mean syllable duration in comparison with nonfocus condition

Intensity

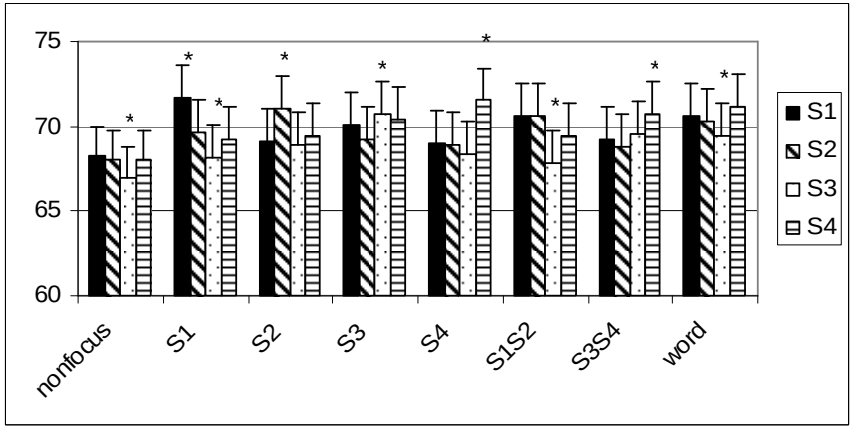


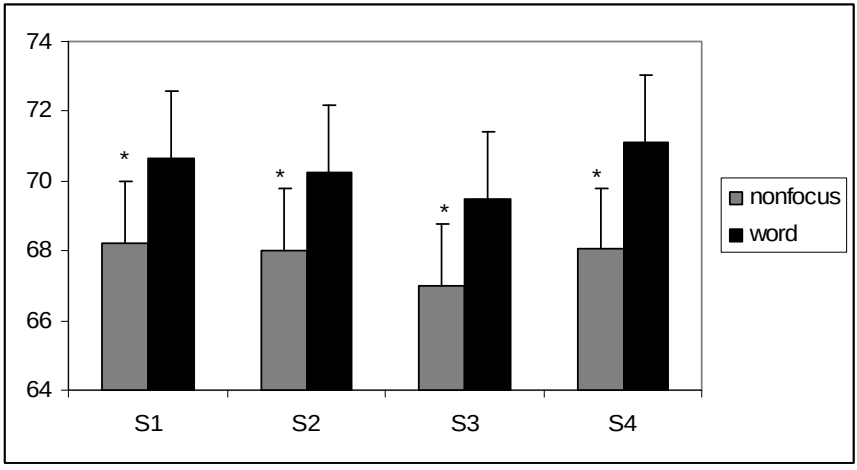
Figure 8. Mean intensity(dB) by syllable and focus condition. *: significant at p<.01 on top of syllable that is significantly different from others

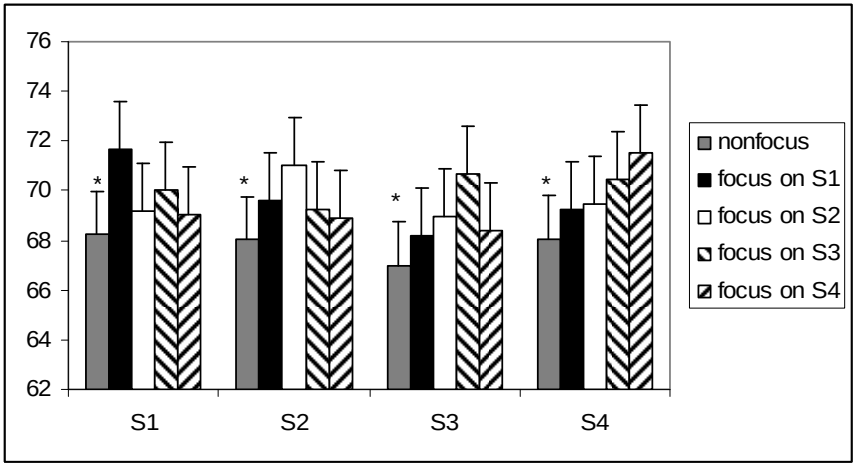
As shown in figure 8, when focus was on the first syllable (S1), the first syllable had the highest intensity (S1>S2~S4>S3). When focus was on the second syllable, it had the highest intensity in the word (S2>S1~S4~S3). Similarly, the third and fourth syllables had highest intensity when they were focused (Focus on S3: S3~S4>S1~S2, focus on S4: S4>S1~S2~S3). When focus was on the first two syllables S1S2, both syllables had higher intensity in comparison with the last two syllables (S1~S2~S4>S3). When focus was on the last two syllables S3S4, both syllables had raised intensity but the last syllable had the highest intensity in the word (S4>S1~S2~S3), which is

probably due to the compounded effect of accentual and right-headedness marking.

Compared with nonfocus condition, when focus was on the whole 4-syllable word, all syllables had raised intensity as shown in figure 9a. When focus was on individual syllables, the syllables under corrective focus had heightened intensity and there is a spill-over effect; that is, its neighboring syllables (either prefocus or postfocus) also had heightened intensity (fig. 9b and 9c).

Generally, the results in figures 8 & 9 showed that the intensity results closely mirror the duration results, but there is a spill-over effect for intensity such that when a syllable is under focus, it not only had raised intensity but its neighboring syllables also had heightened intensity while no spill-over effect was found for duration.





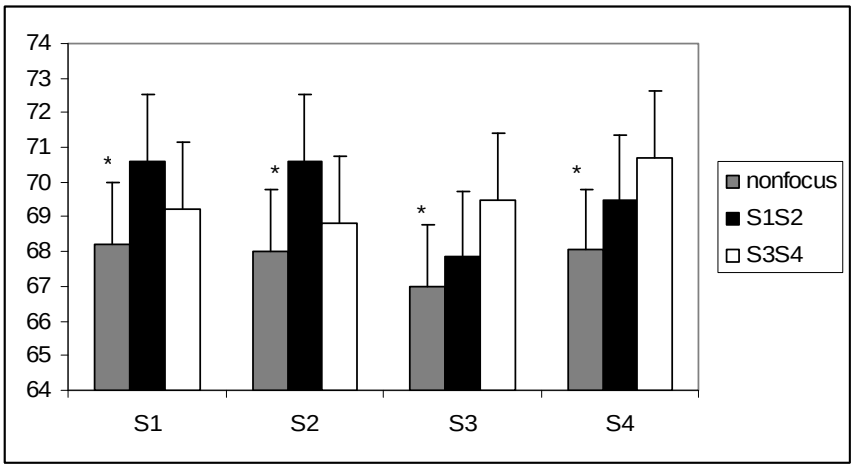


Figure 9. Mean syllable intensity in comparison with nonfocus condition

F0 range

A mixed effect three-way ANOVA was conducted on F0 range. The fixed effects consisted of syllable position (S1, S2, S3, S4), focus condition (nonfocus, S1, S2, S3, S4, S1S2, S3S4, and word) and tone (curve, drop, falling, level, rising). The random effect included speaker (n=10) and word (n=5). The results showed significant effects for the main factor position $F(3, 1778)=23.6$, $p<.0001$ and tone: $F(4, 1778)=141$, $p<.0001$ but no significance for focus: $F(7, 1778)=0.84$, $p=0.5$ ns. Nevertheless, there were significant interaction effects: position x focus: $F(21, 1778)=2.6$, $p<.0001$, position x tone: $F(12, 1778)=7.52$, $p<.0001$, position x focus x tone: $F(84, 1778)=1.6$, $p<.001$. A posthoc Tukey test was then conducted to examine the three-way interaction effect position x focus x tone. The results are presented in figure 10.

The results on F0 range generally showed that the constituents under corrective focus tended to have expanded F0 range. For example, when focus was on the first syllable, the first syllable of all tones tended to have the largest F0 range. When focus was on the second syllable, its F0 range was expanded in comparison to the nonfocus condition. Similar effects were found when focus was on the third and the fourth syllables. When focus was on the whole word, the first and the last syllables had more expanded F0 range than the other two syllables of the word. When focus was on the first two syllables of the word, both syllables had expanded F0 range. When focus was on the last two syllables, the last syllable had expanded F0 range,

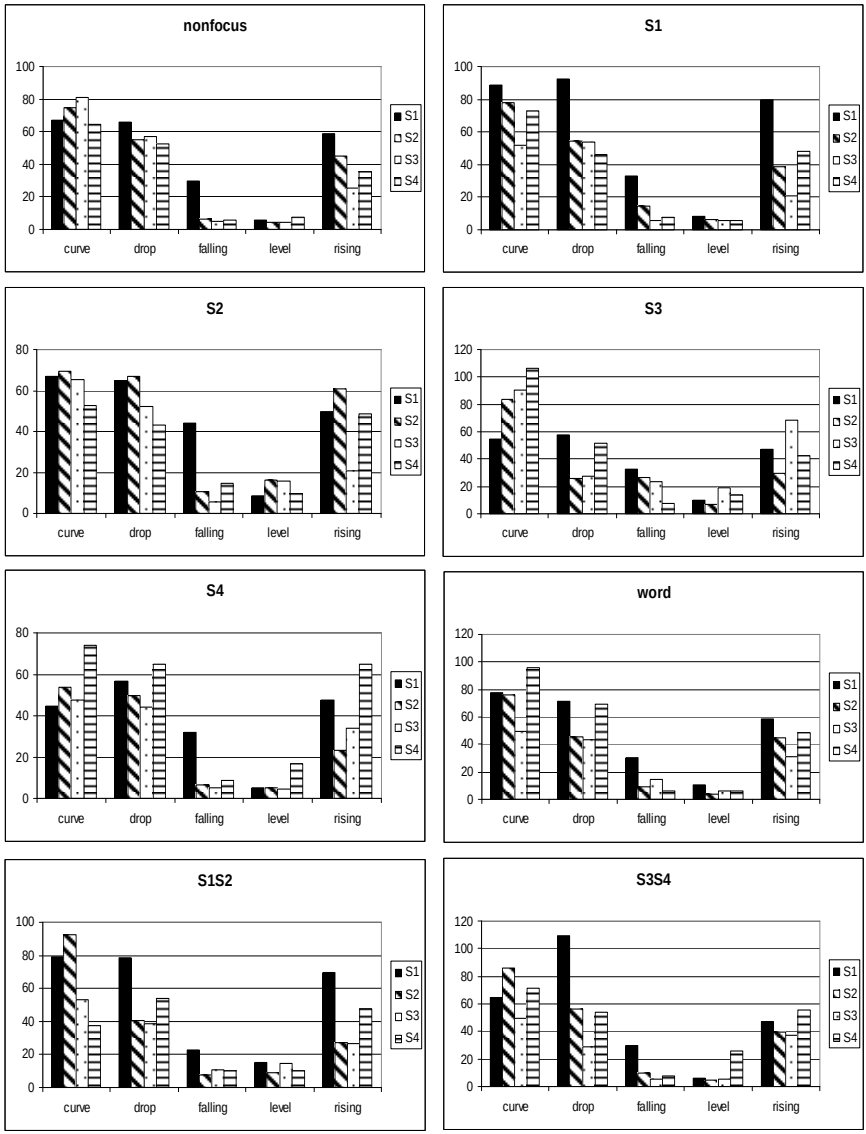
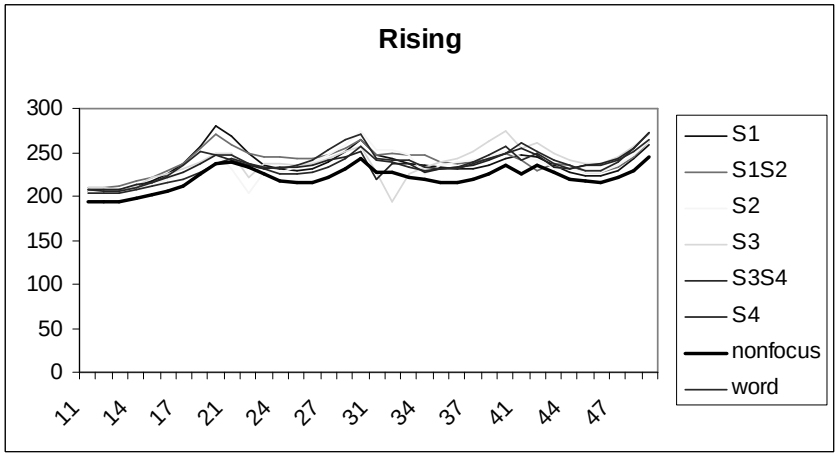


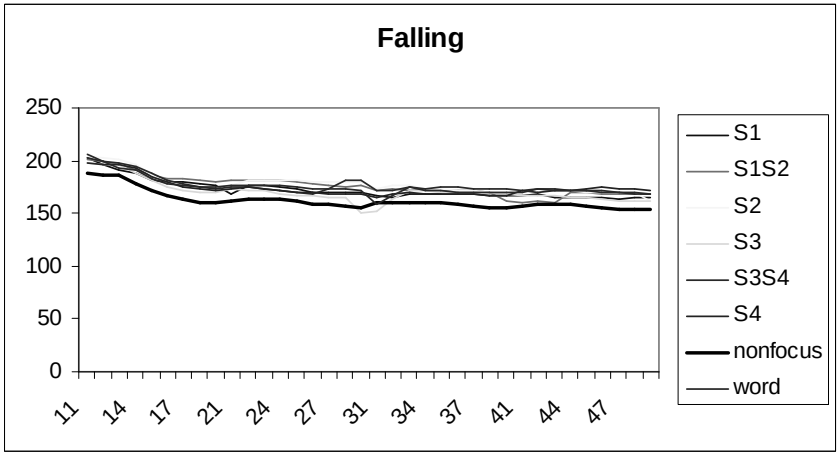
Figure 10. Mean F0 range(Hz) by tone and focus conditions.

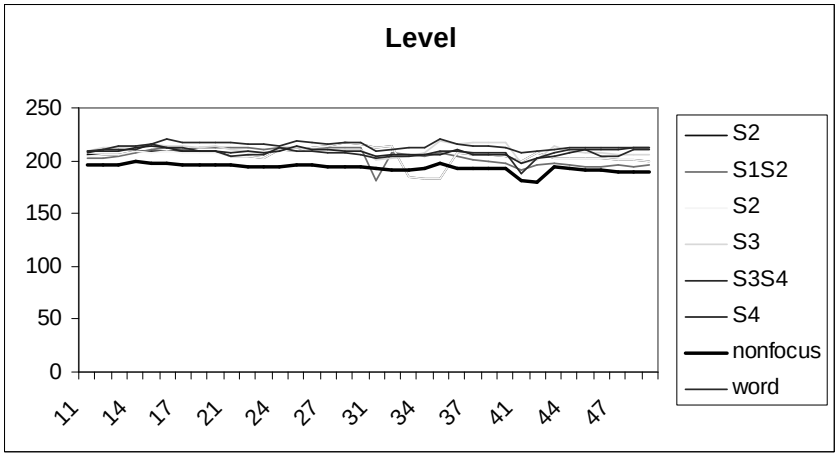
F0 contour

Apart from the expanded F0 range of the constituent under emphasis, corrective focus was also shown to heighten F0 of the whole F0 contour of the word in comparison with nonfocus condition as illustrated in figure 11. For example, the F0 contour of the nonfocus condition (the thick dark line) is always lower than those of all other focused conditions across all syllables of

the whole word and all tones. This is consistent with Michaud and Vũ's (2004) finding on "Fo register raising" of the drop tone under emphasis.







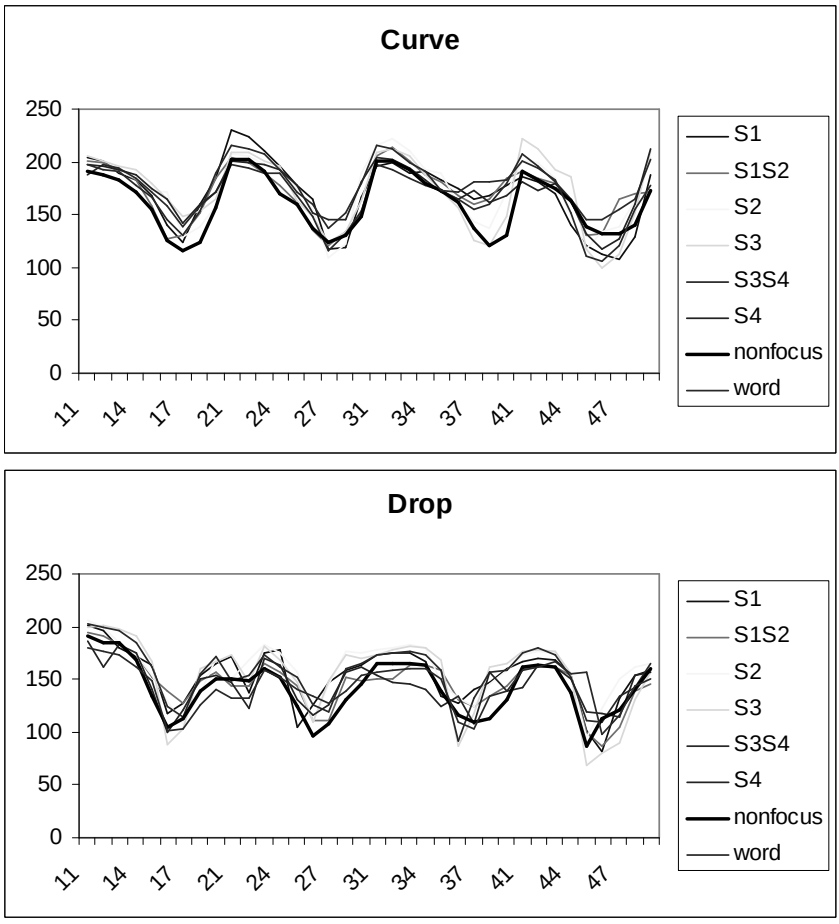


Figure 11. Mean F0 contour by focus conditions. 11-19: first syllable, 21-29: second syllable, 31-39: third syllable, 41-49: fourth syllable. S1: focus on S1, etc.

Discussion

In brief, the results of the focus experiment showed that Vietnamese used duration, intensity, F0 range and F0 height in signalling corrective focus, confirming observations on accent marking by previous researchers (Đỗ 1986; Chaudhary 1983; Hoàng & Hoàng 1975; Gsell 1980). On the other hand, rhythmic pattern tends to be manifested mainly by duration. The tendency for right-headed rhythmic pattern, the syllable coupling with the even syllables being longer than the odd syllables, supported by the compensatory weakening of the penult syllable before the last strong syllable is shown to be well preserved under different focus conditions. In contrast to other languages (English: Turk and White 1999; in English and Dutch: Cambier-Langeveld and Turk 1999; Mandarin: Chen 2006) where there was a spill-over lengthening

effect under contrastive focus, no spill-over effect was found for duration but there is a spill-over effect for intensity such that when a syllable is under focus, not only did it have raised intensity but its neighboring syllables also had heightened intensity, suggesting a language specific mechanism of accentual marking.

4. Conclusion

In summary, the results of this study showed several important things about Vietnamese rhythmic patterns. First, there is a tendency of syllable coupling indicated mainly by syllable duration and supported by the native listeners' perception results, suggesting that polysyllabic words in Vietnamese tend to be parsed into bi-syllabic iambic feet with a rightward or retrograde rhythmic pattern, consistent with Trần's (1969) observation. Second, when there is only one syllable such as in the case of the odd syllable in 3-syllable and 5-syllable words, the odd syllable tended to be lengthened to fill the bi-syllabic foot template, suggesting that bi-syllabic foot is a prosodic unit above the syllable. This is consistent with the large proportion of bi-syllabic words in Vietnamese. The bi-syllabic unit is also a prosodic template of syllable coupling in folk poetry and reduplications in Vietnamese. Third, Vietnamese polysyllabic words are right-headed with the last syllable always the longest syllable in the word, consistent with findings in bi-syllabic coordinative compound words (Nguyễn & Ingram 2007a) and in bi-syllabic reduplications (Nguyễn & Ingram 2007b) and confirming the observations by previous researchers (Thompson 1965; Jones and Huỳnh 1960) that last word of the phrase receives a strong stress. This right-headed pattern with the even syllable longer than odd syllable, supported by the compensatory weakening of the penult syllable before the last strong syllable is preserved under different focus conditions. Fourth, corrective focus in Vietnamese is realized by duration, intensity, F0 range and F0 height, supporting observations on accent marking by previous researchers (Đỗ 1986; Chaudhary 1983; Hoàng & Hoàng 1975; Gsell 1980).

Even though this study uses reiterant speech in the form of nonsense reduplicative words which allows the examination of the rhythmic pattern, the finding on the right-headed bi-syllabic foot template is well-supported by previous real-word data (reduplication: Nguyễn & Ingram 2007b, coordinative compound: Nguyễn & Ingram 2007a) and particularly patterns of syllable coupling in folk poetry. Therefore, it is predicted to extend to real-word spontaneous speech, which needs to be further investigated in future studies.

REFERENCES

Boersma P. & Weenink D. 2007. "Praat: doing phonetics by computer (Version 4.5.18) [Computer program]." Retrieved March 30, 2007, from http://www.praat.org/

Cambier-Langeveld, T. and Turk, A. 1999. "A cross-linguistic study of accentual lengthening: Dutch vs. English." *JP* 27:255–280.

Cassidy, S. 1999. "Compiling multi-tiered speech databases into the relational model: Experiments with the Emu system." In G. Gordos ed., *Proceedings of Eurospeech 99*, pp. 2239-2242. (European Speech Communication Association, Budapest, 1999).

Chaudhary, C. C. 1983. "Word stress in Vietnamese: A preliminary investigation." *IL* 44:1-10.

Chen, Y. 2001. "Contrastive focus in Beijing Mandarin (Abstract)." *JASA* 110 (5):2737.

Chen, Y. 2006. Durational adjustment under contrastive focus in Standard Chinese. *JP* 34:176-201.

Đỗ, T. D. 1986. "Elements pour une e'tude comparative de l' intonation en Francais et en vietnamien: L'accent de mots en vietnamien." Memoire de DEA (Universite' de Paries 3 ILPGA.

Duanmu, S. 2000. *The Phonology of Standard Chinese.* Oxford: Oxford University Press.

Everett, K. M. 1998. "The acoustic correlates of stress in Piraha." *JAL* 1:104-162

Feng, S. 1998. "On the Natural Feet in Chinese." *Zhongguo Yuwen* 1:40-47.

Gandour, J., Potisuk, S. and Perkins, J. 1997. "Using reiterant speech to study prosodic phenomena in Thai." In A.S. Abramson ed., *Studies in Linguistics in Honor of Vichin Panupong.* Bangkok: Chulalongkorn University Press.

Gsell, R. 1980. "Remarques sur la structure de l' espace tonal en Vietnamien du sud (Parler de Saigon)." Cahiers d'etudes Vietnamiennes, 4 (Universite' Paris 7).

Halle, M. and Vergnaud, J-R. 1987. *An Essay on Stress.* Cambridge, Mass: MIT Press.

Hayes, B. 1985. "Iambic and trochaic rhythm in stress rules." *BLS* 11:429–446.

Hayes, B. 1995. *Metrical Stress Theory: Principles and Case Studies.* Chicago: The University of Chicago Press.

Hirst, D.; Bouzon, C. 2005. "The effect of stress and boundaries on segmental duration in a corpus of authentic speech (British English)." In *Proceedings of Interspeech 2005*, Lisbon, Portugal, pp. 29-32.

Hoàng, T.; Hoàng, M. 1975. "Remarques sur la structure phonologique du Vietnamien. Essais Linguistiques." *Etudes Vietnamiennes* 40.

Jones, R. B. and Huỳnh, S. T. 1960. *Introduction to Spoken Vietnamese.* Washington, D.C

Kager, R. 1995. “The metrical theory of word stress.” In John A. Goldsmith ed., *The Handbook of Phonological Theory,* pp. 367-402. Cambridge, MA, and Oxford, UK: Blackwell.

Larkey, L. 1983. “Reiterant speech: An acoustic and perceptual validation.” *JASA* 73:1337-1345.

Liberman, M. 1978. “Modelling of duration pattern in reiterant speech.” In D. Sankoff ed., *Linguistic Variation: Models and Methods,* pp. 127-138. New York: Academic Press.

Liberman, M. and Streeter, L. 1978. “Use of nonsense-syllable mimicry in the study of prosodic phenomena.” *JASA* 63:231-233.

McCarthy, J. and Prince A. 1996. *Prosodic Morphology 1986.* RUCCS TR-32. Rutgers University.

McCarthy, J. and Prince, A. 1990. “Foot and word in prosodic morphology the Arabic broken plural.” *Natural Language and Linguistic Theory* 8:209–284.

Michaud, A. and Vũ, N. T. 2004. “Glottalized and nonglottalized tones under emphasis: open quotient curves remain stable, F0 curve is modified.” In Bernard Bel and Isabelle Marlien eds., *Speech Prosody-International conference, ISCA Archive,* pp.745-748.

Nakatani, L. and Shaffer, J. 1978. “Hearing ‘words’ without words: prosodic cues for word perception.” *JASA* 63:234-245.

Nakatani, L., O’Connor, K. and Aston, C. 1981. “Prosodic aspects of American English speech rhythm.” *Phonetica* 38:84-106.

Nespor, M. and Vogel I. 1986. Prosodic Phonology. Foris: Dordrecht.

Nguyễn, D. L. 1970. “A contrastive phonological analysis of English and Vietnamese (Pacific Linguistics Series, No 8).” Australian National University, Canberra.

Nguyễn, T.A.T. and Ingram, J.C. 2007a. “Acoustic and perceptual cues for compound - phrasal contrasts in Vietnamese.” *JASA* 112 (3):1746-1757

Nguyễn, T.A.T. and Ingram, J. C. 2007b. “Stress and tone Sandhi in Vietnamese reduplications.” *MKS* 37:15-39.

Nguyễn, T.A.T. and Ingram, J. C. “Perception of stress in Vietnamese compounds: evidence from reduplication.” *Unpublished*

Shen, X. S. 1993. “On the temporal dimension of Mandarin.” *ALH* 26:143-159

Shih, C. 2005. “Understanding phonology by phonetic implementation.” *Proceedings of Interspeech,* Lisbon, Portugal. Sept 4-8.

Sluijter, A. 1995. “Phonetic correlates of stress and accent.” Ph.D. dissertation, University of Leiden.

Thompson, L. 1965. *A Vietnamese Reference Grammar.* Honolulu: University of Hawaii Press.

Trần, H. M. 1969. “Stress, tone and intonation in South Vietnamese.” Ph.D. dissertation, Australian National University.

Turk, A. E. and Shattuck-Hufnagel, S. 2000. “Word-boundary-related duration patterns in English.” *JP* 28:397-440.

Turk, A.E. and White, L. 1999. "Structural influences on accentual lengthening." *JP* 27:171–206.
Xu, Y. 1999. "Effects of tone and focus on the formation and alignment of F0 contours." *JP* 27**:** 55-105
Xu, Y. and Wang, M. 2009. "Organizing syllables into groups—Evidence from F_0 and duration patterns in Mandarin." *JP***.** In press.
Yip. M. 1980. "The tonal phonology of Chinese." Ph.D. dissertation, MIT, IULC

Received: 30 September 2008

19 Prospect crescent
Forest Lake, QLD 4078,
AUSTRALIA
<nguyenthianhthu@email.com>

Body part terms in Kammu

Felix AHLNER[•]
Lund University, Sweden

Abstract

This article is a study of words for body parts in the Mon-Khmer language Kammu. It contains one descriptive part, presenting some 200 thematically ordered entries of words related to the body, with belonging elucidation and illustrations. In a more theoretical and comparative part, the data are used to relate Kammu to a wider debate concerning the semantic domain of body parts. Cross-linguistic studies from the late 1970's proposed several universals for the hierarchical structure of this semantic domain. These studies have since been criticized for being far too narrow, neglecting the variety of more or less hierarchical relationships that may hold between body parts in a language. Lately, comprehensive cross-linguistic studies has both corroborated this criticism and developed elicitation methods for accessing different relationships and structures in this semantic domain. The results show that Kammu follows most of the proposed universals, save for a few noteworthy exceptions. There are, however, also cases where different elicitation methods reveal slightly different hierarchical structures – a fact which supports many of the ideas put forward in recent studies.

1 Introduction

In terms of how we humans perceive our surrounding world and communicate about it, the human body is unique. We perceive other human bodies mainly through sight and feeling, but every one of us also has a body of our own, and it is perceived in a radically different way.

Humans use their bodies every day for vitally important tasks. Therefore every human culture and language needs to be able to talk about the body and, essential to this study, categorize it.

1.1 Anatomical partonomy

For the categorization of the human body and its parts (the 'body part terms' within the 'body part domain'), a concept called *partonomy* has been particularily popular. Partonomy is a hierarchical classification based on a *part*

[*]Acknowledgements: I am very grateful to Kàm Ràw for giving me the possibility to investigate his fascinating language in my own home town. I wish to thank Jan-Olof Svantesson, Niclas Burenhult and Arthur Holmer for their support and opinions. I also wish to thank the Language & Cognition group at the Max Planck Institute for Psycholinguistics in Nijmegen, Netherlands for letting me use their body drawings in this article.

of-relation (also called *meronomy*), i.e. that 'a finger is *part of* a hand' (Brown et al 1976).

One difference between a partonomical and a taxonomical classification (based on a *kind of*-relation: 'a hammer is a kind of tool' etc.) is the so called *transitivity* between different levels of the hierarchy. A taxonomy is fully transitive: a ball-peen hammer is a kind of hammer, a hammer is a kind of tool, a tool is a physical object, etc. Irrespective of how many levels there are in such a hierarchy, it is always possible to say that objects on the lowest level are 'a kind of' object on the highest level, in this case 'a ball-peen hammer is a kind of physical object'.

The partonomy of body parts seems to be only partly transitive. For example, "teeth are parts of mouths, mouths are parts of faces, but teeth are not parts of faces" (McClure 1975 in Enfield et al 2006:143).

1.2 Simplex and complex terms

Like most other surveys of the body part domain, this one makes a distinction between a *simplex term* and a *complex term*. The name simplex term denotes a body part term that consists of one synchronically unanalyzable morpheme (e.g. English *arm, eye, foot, tongue*). The complex term is usually a compound consisting of several morphemes, dividable into a *head* and a *modifier* (e.g. English *armpit, wing of the nose, ear lobe, knee cap, toenail*). The head of a complex term describes the body part's appearance, function or similarity to something else, and the modifier expresses something about the location of the body part, e.g. Finnish *sormenpää* 'fingertip' lit. 'head (*pää*) of the finger (*sorm-*)'. See Andersen 1978:355 for more examples.

This distinction is considered especially important in the survey of a partonomical system, since the modifier of a complex term supposedly bears some hierarchical information. A *knee cap* is therefore a 'cap related to the knee', and from a partonomical point of view this means that the knee cap is *part of* the knee (Brown et al 1976:74).

1.3 Partonomical surveys and universals

The largest survey in the domain is Cecil H. Brown's *General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature* from 1976. Brown examined the anatomical partonomy in 41 languages from different language families from all over the world. He also posited 'principles' about their arrangement and naming strategies (Brown 1976).

In 1978, Elaine Andersen published *Lexical universals of body-part terminology* in Joseph Greenberg's four tome *Universals of Human Language*. Andersen's study draws much of its data from Brown, but also from other studies on particular languages or language families. (Andersen 1978:346)

Andersen posits nine main "universals of categorization for the domain of human body-parts" (Andersen 1978:351ff), and she mentions that

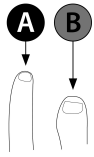
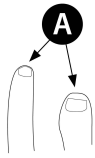
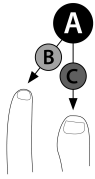
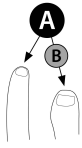
most of them are based on the eleven 'principles' posited by Brown (Brown 1976:404ff).

These universals are presented verbatim below, except for a few wordy explanations and examples that were left out or reworded.

The use of SMALL CAPS in this list indicates a *parton* (pl. *parta*), i.e. a part of the body "that may or may not be labeled in any given language" (Brown 1976:401), meaning that the parta mentioned below may not always correspond precisely to the English terms bearing the same name.

a) The BODY is labeled in all body-part partonomies.
b) Every language includes a term for HEAD in its lexical field of body-parts, and the term is always immediately possessed by BODY. Other categories which usually occur at the second level of the partonomy include TRUNK, ARM (plus HAND) and LEG (plus FOOT).
c) All languages label EYES, NOSE and MOUTH.
d) The upper limb, ARM (plus HAND), is named by a distinct term in all languages.
e) The categories FINGER and TOE are always labeled (by one of four general patterns):

Table 1. Labelling patterns for FINGER and TOE

Pattern	Graphic version	Examples
1. Different basic terms for each category		Finnish *sormi* and *varvas*. French *doigt* and *orteil*. Swedish *finger* and *tå*.
2. One polysemous basic term		*To emphasize either meaning, one can usually add either 'hand' or 'foot' to the polysemous term.* Czech *prst* (*prst na ruce, prst na noze*). Hebrew *etzba* (*etzba yad, etzba regel*).
3. Different terms derived from the same root		*B is usually the term for 'hand' or 'arm', C is usually the term for 'foot' or 'leg'.* Mayan *aal k'ab* and *aal ook*. Tagalog *daliri sa kamay* and *daliri sa paa*.
4. One basic unanalyzable term for FINGER, with TOE derived from it		*B is usually the term for 'foot' or 'leg'.* Hungarian *ujj* and *lábujj*. Malay *jari* and *jari kaki*. Tamil *viral* and *kālviral*.

f) All languages name (FINGER)NAIL and (TOE)NAIL by one of two patterns. Languages with the first pattern have one basic term applied to both categories, and the second pattern have different terms derived from a common root, like English 'fingernail' and 'toenail'.
g) A term for LEG implies a separate term for ARM.
h) A term for FOOT implies a separate, non-identical term for HAND.
i) Terms for INDIVIDUAL TOES imply terms for INDIVIDUAL FINGERS.

Finally, earlier studies in the categorization of other domains have found that folk taxonomies and "[b]ody-part partonomies rarely exceed five hierarchical levels and never exceed six" (Andersen 1978:348, Brown 1976:401). This is often referred to as the 'depth principle'.

1.4 Criticism

The main criticism against Brown and Andersen and their conclusions has concerned the focus on the partonomical relationship.

In 1985, Palmer and Nicodemus published an article with the telling name *Coeur d'Alene Exceptions to Proposed Universals of Anatomical Nomenclature*. In the introduction they state that

> [L]inguistic anthropologists […] have mistakenly assumed that the *part-whole relation* is universally the primary semantic dimension of anatomical domains, […] The purpose of this paper is to demonstrate that at least one other relation, that of *spatial orientation*, must be considered before cross-cultural or cross-linguistic comparisons can yield valid universals of folk anatomy. (Palmer and Nicodemus 1985:341ff, my emphases).

Their paper exemplifies this spatial relationship with a survey of the body part domain in the Salishan language Coeur d'Alene (Snčícuʔumšcn), spoken in the northwestern USA. Its inventory of body part terms contains for example *s-cin-čém-cən* lit. 'the surface below the mouth' for English 'neck', and *s-čiʔ-qṅ-šən* lit. 'forward part of the top of the leg' for 'knee'.

Such terms "cannot easily be arranged in a taxonomic hierarchy, even one based upon part-whole relations" and consequently "[c]ross-linguistic comparisons which presume the universality of part-whole hierarchies must fail". The authors' conclusive opinion is that

> the part-whole model is inappropriate to folk anatomical domains based on nontransitive spatial relations […] Linguistic anthropology needs not one model of folk classification, but several (Palmer and Nicodemus 1985:353f).

Similar ideas were expressed in a 2006 special issue of the journal *Language Sciences* (hereafter abbreviated *LS*), dedicated entirely to descriptions and analyses of the body part domain in languages from all over the world (Majid et al. 2006).

All of the researchers found it difficult to analyze the entirety of their findings as a partonomical system. Many of the investigated languages simply have no word corresponding to English 'body part' or even 'part'. This forces the researcher to use possessive constructions like 'X has Y' or locative constructions like 'X is located on/at/near/etc. Y' – none of which actually contain any partonomical information (Meira 2006:275, Levinson 2006:233ff, Wegener 2006:357f, Terrill 2006:320, Gaby 2006:218, Burenhult 2006:178f). In some languages, e.g. Punjabi, different body part terms can have different relations to each other: partonomical 'the leg is part of the body', possessive 'the finger has a nail' or locative 'the nose is on the face' (Majid 2006:255ff).

Enfield found a similar system in Lao and ends his article with a conclusion that could have summarized the entire issue:

> The semantic relations which pertain between terms for different parts of the body not only include part/whole relations, but also relations of location, connectedness, and general association. Calling the whole system a 'partonomy' attributes greater centrality to the part/whole relation than is warranted. (Enfield 2006b:199)

This notion of different coexisting relational systems offers a possible explanation to why the body part partonomy is not always fully transitive: there simply is no continuous path of part-of relations in the partonomical hierarchy between the nail and the body, for instance. The arm might be a *part of* the body, and even the finger a *part of* the hand, but as long as the nail is *on* the finger and the hand is *connected to* the arm, it is impossible to say that the nail is *part of* the body.

2 Kammu

Kammu (also known as Khmu) is a Khmuic language in the Mon-Khmer branch of the Austro-Asiatic language family. The language has about 500,000 speakers, mainly in Laos but also in Thailand, Vietnam, China and Burma.

Kammu is a highly isolating language with no inflectional morphology. There is however an abundance of derivational affixes, e.g. an instrumental infix in *prnɔ̀* 'broom', from *pɔ́* 'to sweep'. Kammu is head-initial and the basic word order is SVO. English adjectives correspond to stative verbs: *saŋ háan* can mean 'dead pig' or 'the pig died'.

My informant speaks the Yùan dialect, which distinguishes high and low tone. These are marked on vowels with ´ and ` respectively. Long vowels are written double. The rest of the alphabet is identical to the IPA, except for <ñ> representing [ɲ] and <y> representing [j]. <ʔw> and <ʔy> are laryngealized approximants (Svantesson 1983).

Kammu has been in contact with the neighboring but genetically unrelated Tai-Kadai language Lao for several centuries, borrowing thousands of words and expressions. Many of these words have become parts of the everyday language, whereas others are used only in more restricted contexts, e.g. prayers and sayings.

Over the last 30 years many of the Kammu populations have been relocated to cities in Laos and to ethnically mixed villages. My informant came to Sweden before this, so he speaks a language that retains words that are no longer used, reflecting the Kammu spoken during his upbringing in the 1940's.

2.1 Simplex and complex terms in Kammu

There are two types of simplex terms in Kammu. *màt* 'eye' is a basic underived noun, the most common kind of simplex term in the body part vocabulary. *srŋèɛk* 'area above wrist or ankle' is a derived noun, in this case derived through prefixation from the verb *ŋèɛk* 'to be thin in the middle'.

These two types both differ from complex terms like *khúul màt* 'eyelash', a compound with the head *khúul* '(body) hair' and the modifier *màt* 'eye'.

3 Aims

The main aims of this study is 1) to survey the inventory of body part terms in Kammu, and 2) to analyze the relations between body parts terms and attempt to categorize them in one or more systems. The results will then be compared with the universals posited by Andersen and Brown.

4 Method

The data for this paper was gathered during several elicitation sessions with a native speaker of Kammu at the department of linguistics at Lund University, Sweden. I used many of the methods explained in the *LS* articles *Elicitation guide on parts of the body* (Enfield 2006a) and *Body colouring task* (van Staden & Majid 2006). Some of the methods were

developed for situations with multiple informants, but their general design was still applicable for me.

The *Kammu Yùan-English Dictionary* (Svantesson et al, forthcoming) served as a secondary source. I used it to compare translations and to look for unusual terms. I noted these and asked about them during the next session.

Once the collected inventory of body part terms had reached a substantial size, I began constructing sentences that could reveal how the different terms relate to each other. Seeing that all the surveys in *LS* support the notion of coexisting categorizational systems, it was necessary to also investigate the *possessive categorization* and the *spatial categorization* of body part terms.

4.1 Partonomical categorization

Early during the elicitation phase it became clear that Kammu has no word for 'part', precluding questions like 'Is X part of Y?'. *Elicitation guide on parts of the body* attends to this rather common situation and suggests that the researcher find "[l]anguage-specific expressions relating various body parts to each other", like "An arm must have a hand" (Enfield 2006a:156). But again, such phrases do not necessarily convey any partonomical information.

In his article about Lao, Enfield mentions the *entailment test*. The basic method is to attribute mutually exclusive properties to two different body part terms, e.g. as a statement 'This man has a mosquito bite on his X, but not on his Y' or as a question 'Can my X be unscathed if my Y is bleeding?'. Enfield means that this method can provide evidence that one term is *part of* another term or not (Enfield 2006b:197). Two typical 'frames' for this test in Kammu were:

(1) ò àh òh tàa X tɛ̀ɛ tàa Y ò pə́ə àh
1SG have wound at X but at Y 1SG NEG have
'I have a wound on my X but not on my Y'

(2) múuc pók X kə̀ə tɛ̀ɛ pə́ə pók Y kə̀ə
ant bite X 3SG but NEG bite Y 3SG
'An ant bit his X but not his Y'

4.2 Possessive categorization

The original frame for statements about the possessive categorization was:

(3) X (ò) àh Y
X (1SG) has Y
'(my) X has Y'

This frame was not always adequate, since some terms can only be possessed if they occur in pairs (see 6.2 for further detail). The subscript WP in the following frame indicates that the two terms must belong together in a word pair.

(4) X (ò) àh Y_{WP} àh Z_{WP}
X (1SG) has Y_{WP} has Z_{WP}
'(my) X has Y_{WP} and (has) Z_{WP}'

4.3 Spatial categorization

There was never any established sentence frame for investing the spatial categorization. Once my informant understood what I was looking for, I would just mention a body part term and he came up with acceptable sentences which described its relations to other terms. Their structure was usually:

(5) X yèt PREPOSITION Y
X located PREPOSITION Y
'X is located PREPOSITION Y'

See 6.3 for an overview of the prepositions included in the analysis.

5. Results

5.1 Inventory of body part terms

The inventory of body part terms is divided into categories and presented in tables, then discussed and exemplified. These tables consist of three columns showing 1) the Kammu term, 2) an English translation and 3) possible comments on any particular term and always a gloss for complex terms. Some of the tables are accompanied by drawings showing the scope of different body part terms, based on the results from the body part coloring task.

Throughout this paper the physical extension of a body part term will be called its *scope*, and the corresponding verb is *to cover*.

It is important to notice that in these tables the glosses for complex terms are translated lexically but not syntactically. E.g. 'nostril' in Kammu is *hntú mùh*, glossed here as 'hole nose', while the translation with English word order would be 'nose hole'. However, in the running text all glosses are translated both lexically *and* syntactically.

The complex terms all consist of isolated morphemes, but some of them are nevertheless unanalyzable ('cranberry morphemes'). Such morphemes are simply glossed with the Kammu word. If some part of a complex term is translated into English as several words, these will have periods between them, e.g. *plé plɔ́ŋ* 'fruit lower.leg'.

5.2 Major areas and parts of the body

Kammu has a number of encompassing complex terms for areas of the body, similar to English 'upper/lower body'. These are created with *lɨ̀aŋ* 'side, direction' (from Lao *luaŋ*) and they occur in contrastive pairs (upper ↔ lower, left ↔ right etc.).

Table 2. Encompassing terms for areas of the body

Kammu	**English**	**Details (*lɨaŋ* plus …)**
lɨaŋ tí / kmpóŋ	upper body	arm / head
lɨaŋ cɨaŋ / tɔ̀	lower body	leg / bottom
lɨaŋ knáaŋ	front side	front
lɨaŋ kntrɔ̀ɔŋ	back side	back
lɨaŋ káal	in front of, before	in front of
lɨaŋ knní	behind, after	behind
lɨaŋ wè	left side	left
lɨaŋ hám	right side	right
lɨaŋ nɔ̀ɔk	outer (of arm, leg)	outside (Lao *nɔɔk*)
lɨaŋ klúaŋ	inner (of arm, leg)	inside
lɨaŋ trɨak	one side	side

The pair *lɨaŋ kmpóŋ* ↔ *lɨaŋ cɨaŋ* corresponds to English upper body ↔ lower body, and the two meet at the waist, *kùaŋ*. As a more easily translatable example, the Kammu say that a bed has *lɨaŋ kmpóŋ* 'head-end' and *lɨaŋ cɨaŋ* 'foot-end' (see 5.6 for a discussion about *cɨaŋ*).

Another indication of which body parts the Kammu consider most significant is the belief in twelve 'body souls' (*hrmàal*) living in every human being. The torso, head, nose and mouth house one such soul each. The pairs of ears, eyes, arms and legs house two souls each (Lundström & Svantesson 2006:144ff). A person is healthy as long as these souls are present, but they will leave the body if they become scared or are mistreated. Their absence can result in anything from stumbling and headache to blindness, illness and death. In order to retrieve a lost soul a shaman and the family of the suffering person perform elaborate rituals spanning over several days (Tayanin et al. 2006:37ff).

Several times, my informant also suggested a division of the body into 'the three important parts': *kmpóŋ* 'head', *lɔ̀h* 'torso (and arms)' and *cɨaŋ* 'leg(s)'.

5.3 The head and the neck

Table 3. The head and the neck

Kammu	English	Details
Simplex terms		
kmpóŋ	head	
kntɨ́ar ~ kntúar	neck	
tk'lòk	nape of the neck	
rmpòh	face	
ktáh	forehead	
káap	chin	
pɔ́ɔm	cheek	
kéep	side of lower jaw	
stmàat	temple	
kntùur	crown	cf. *kntùur pnɨ̀m* 'top of a termitary'
klə	hair on the head	
khúul	body hair	
màt	eye	
hrməəy	ear	
mùh	nose	
tnɔ́h	mouth	
ráaŋ	tooth	
hntáak	tongue	
kltak	palate	
hɨ́l	gums	
tróoŋ	throat, voice	
Complex terms		
krwèey klə	bare spot in the center of a whorl of hair	'left.unsown hair.on.the.head'
hntú mùh	nostril	'hole nose'
hntú kntùur	fontanel	'hole crown'
hntú hrməəy	ear canal	'hole ear'

hntú tnɔ́h	oral cavity	‘hole mouth’
khúul hmpíir	eyebrow	‘hair pumpkin’
khúul màt	eyelash	‘hair eye’
khúul káap	beard	‘hair chin’
khúul stmàat	hair at the temples	‘hair temple’
khúul kràw	whiskers, sideburns	‘hair whiskers’ (Lao *gaw*)
khúul tñmùuñ	moustache	‘hair dirty.around.mouth’, from *mùuñ* ‘dirty face’
hntá kmñàam	hair at the neck	‘tail cricket’
màt klɔ́ɔk	white of the eye	‘eye white’
màt yíaŋ	iris and pupil	‘eye black’
klɔ́ɔŋ kók	Adam’s apple	‘stone hogplum’
klɔ́ɔŋ màt	eyeball	‘stone eye’
lá hrməəy	ear-conch, pinna	‘leaf ear’
rŋkɔ́ɔŋ mùh	bridge of the nose	‘mountain-range nose’
hmpúur màt	eyelid	‘skin eye’
hmpúur tnɔ́h	lips	‘skin mouth’
ráaŋ tùl	wisdom tooth	‘tooth stick.up.through’
ráaŋ túut	molar tooth	‘tooth bottom’
ráaŋ tal	front tooth	‘tooth top’
ráaŋ cntràas	front tooth	‘tooth lightning’
ráaŋ òm pù	milk-tooth	‘tooth liquid.breast’ (= ‘tooth milk’)
tróoŋ yèl	uvula	‘throat *yèl*’
tíal hntáak	tip of the tongue	‘top tongue’
hɨ́l laŋ pə̀h	upper gums	‘gum side up’
hɨ́l laŋ tál	lower gums	‘gum side down’

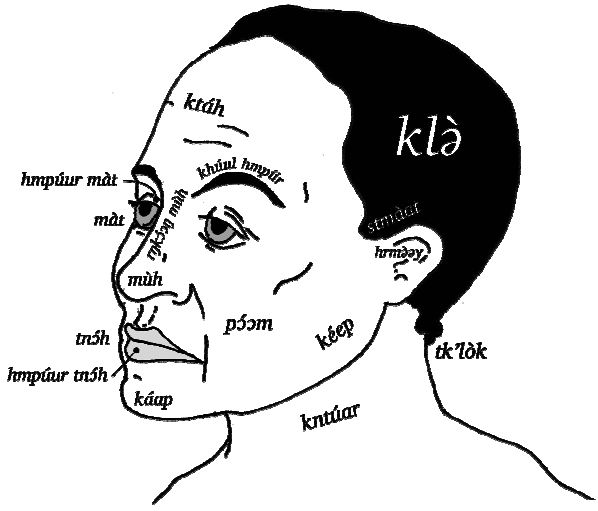


Figure 1. The head and neck (drawing by Jonas Ahlner)

Both *khúul* '(body) hair' and *hntú* 'hole' are common as heads in complex terms all over the body. *khúul* refers to any hair except for the hair on the top of one's head, which is *klə*. The meaning of a complex term with *khúul* as its head is easily inferable ('shin hair', 'armpit hair' etc.) and will not be listed. There are however two complex terms for facial hair with slightly different modifiers.

The first term is *khúul hmpíir* 'eyebrow', lit. 'pumpkin hair', a complex term whose origin and meaning defies explanation at this point. The second term is *khúul tñmùuñ* 'moustache' where the modifier is originally a derived form of the root *mùuñ* 'dirty face'.

(6) tnɔ́h mèe cə̀ tñmùuñ yɔ̀h ràa kə̀ə
mouth 2SG dirty dirty.face go wash 3SG
'Your mouth is dirty, go and wash it!'

The teeth are divided into 'back' and 'front' like in many other languages, but the front teeth also have the alternative name *ráaŋ cntràas* 'lightning teeth'. My informant suggests a metaphoric comparison of how both lightning and front teeth strike down and split things. The term for wisdom tooth, *ráaŋ tùl*, is less poetic and simply states that these teeth 'come up through'.

One example of a culturally (or maybe 'genetically') based term is *màt yíaŋ* 'black of the eye' covering both the pupil and the iris. The reason for this conflation is simply that all Kammu have dark brown iris with no

perceptually salient border to the pupil. When I showed my informant how my own green iris contrast with the pupil he surprised me by asking "Do you think my (black of the) eye has two parts too?". This question clearly shows that the 'black of the eye' constitutes one conceptual whole in Kammu.

5.4 The torso

Table 4. The torso (back and front)

Kammu	English	Details
Simplex terms		
lɔ̀h	body, torso	see discussion about polysemy below
kntrɔ̀ɔŋ	back	
plà	shoulder	(possibly from Lao *ɓa*)
srnàat	shoulder blade	
kntìip	upper back	
kàam	lower back	
lùuy	stomach, abdomen, belly	
póoŋ	area below the navel	
kùaŋ	waist	
àk	chest	
trɨ́ak	sides	approx. from armpit to waist
lmpía	flank	
klʔèk	armpit	
pù	breast	
tɔ̀	rump, behind	
Complex terms		
(hntú) kntìiŋ	navel	'(hole) navel'
cmə̀ kntìiŋ	navel string	'thread navel'
póom tɔ̀	buttocks	'protuberance rump' (Lao *pom*)
trɨ́ak càar	lower part of side	'side spaced', reflecting the space between ribs

tŕak cìik	upper part of side	'side close', reflecting the space between ribs
prlè pù	nipple	'fruit breast'
cnóor pù	nipple	'summit breast'
pɨt tríik	inward bend above buttocks	'behind frog'
hntú ʔyíak	anus	'hole shit' (impolite term)
hntú klʔùŋ	anus	'hole hipbone' (preferred term)
hntú pɨt	anus	'hole behind' (preferred term)

There is no unique term for 'torso'. *lɔ̀h* is the common term for the entire body, but its scope usually shrinks when it is contrasted with other terms. This behavior is similar to that of English 'body' in phrases like 'she has scars on her face and her arms but none on her body'. As mentioned above, *lɔ̀h* is contrasted with *kmpóŋ* 'head' and *cɨaŋ* 'legs' as one of three 'important parts' of the body. In this case, *lɔ̀h* obviously covers the torso, or at least 'the entire body minus the head and legs'.

Similar to English 'stomach', *lùuy* refers to the visible part on the torso as well as to the internal organ. You can have a *lùuy nám* 'big belly' as well as a *lùuy kràŋ* 'hard (constipated) stomach', and a pregnant woman is said to *àh lùuy* lit. 'have stomach'.

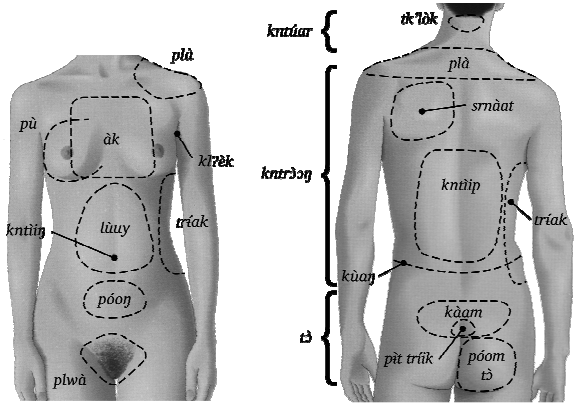


Figure 2. The torso (drawings from Enfield 2006a)

The term *pɨt tríik* 'frog behind' is one of two complex terms that compare a structure of the human body to that of an animal. *pɨt tríik* is the little inward bend right above the buttocks, similar to a structure on a frog's back where the tadpole's tail once was attached. The second term is *hntá kmñàam* for neck hair in the shape of a 'cricket tail'.

5.5 The arm and the hand, measurements

Table 5. The arm and hand

Kammu	English	Details
Simplex terms		
tí	arm	from shoulder-joint to finger tips
kíaŋ	upper arm	
sɔ́ɔk	forearm	(Lao *sɔɔk*)
Complex terms		
ktɔ́ tí	(flesh of) forearm	'bud arm'
tlmɨaŋ tí	segment of the arm	'segment arm'
srŋèɛk tí	area just above the wrist	'inward.bend arm', from *ŋèɛk* 'bent inwards'
krlèey tí	wrist	'joint arm'
krlèey plà	shoulder-joint	'joint shoulder'
krlèey kíaŋ	elbow	'joint upper.arm'
krlèey crkùul	knuckle, finger joints	'joint digit'
séŋ tí	knuckle, finger joints	'joint digit' (less common)
crkùul tí	finger	'digit arm'
mà tí	thumb	'mother arm'
kɔ́ɔn (tàam) tí	little finger	'child (outgrowth) arm'
tmmɔ́ɔŋ tí	fingernail	'nail arm'
làay tí	wrinkles in hand	'stripes arm'
ktáak tí	palm of the hand	'palm arm', see below for further definition
tɨal crkùul	fingertip	'top digit'
kɔ́ɔn kné	biceps	'child mouse'

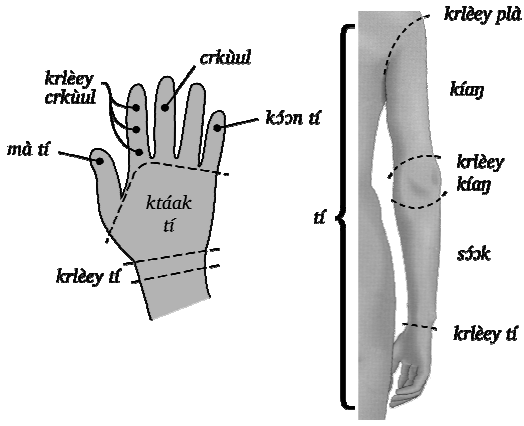


Figure 3. The hand and the arm (body drawings from Enfield 2006a)

The term *tí* covers the entire upper extremity from the shoulder joint to the fingertips. The existence of this encompassing term does not prevent terms that cover smaller parts like 'upper arm' and 'forearm', but there is no separate term covering only the hand.

When *tí* is used together with certain verbs and adjectives its meaning, however, often corresponds to the English use of 'hand'. The hand is very versatile, capable of assuming more shapes and performing more actions than any other part of *tí*. Therefore it will be referred to more often and used together with a large variety of other words. A few examples are *rwàac tí* 'shake hands', *nòom tí* 'clench one's fist' and *sntúk tí* 'hold hands interlacing each other's fingers'. The verbs by themselves translate to 'grab', 'squeeze' and 'entwine' respectively.

The term *ktáak tí* is translated above as 'palm of the hand', but it could also be translated as 'area between knuckles and wrist'. It is for example possible to refer to the back of the hand as the 'top of *ktáak tí*'.

(7) ò ành màk hlŋəəy ktáak tí
ISG have tattoo top palm arm
'I have a tattoo on the back of my hand'

The term *crkùul* corresponds to 'digit' and is the head of the complex terms for both *crkùul tí* 'finger' and *crkùul cìaŋ* 'toe'. As a simplex term the primary meaning of *crkùul* is 'finger'. Even in a context involving the foot, the full term *crkùul cìaŋ* is preferred for 'toe'. This is likely due to the fact that we use our fingers for more purposes than we use our toes.

Since *crkùul* is ambiguous on its own, *tmmɔ́ɔŋ* 'nail' is modified by either *tí* or *cɨaŋ*. My informant stresses that *tmmɔ́ɔŋ* is always located on *crkùul*, therefore a complex term such as **tmmɔ́ɔŋ crkùul* would be uninformative.

Some complex terms are however modified by *crkùul*, e.g. *krlèey crkùul*, a term that covers the knuckle as well as the finger's other joints. The reason that the joints of the finger are sufficiently described as 'digit joints' is probably because they are mentioned much more often than the joints of the toe. Similarly, English 'knuckle' without any modifier never refers to the foot.

If a Kammu speaker wants to refer to the 'toe joints' or the 'toe-tip' after all, both *tɬal crkùul* and *krlèey crkùul* can be additionally modified by *cɨaŋ* in a three-part compound.

5.6 The lower body

Table 6. The leg

Kammu	English	Details
Simplex terms		
cɨaŋ	leg, foot	see discussion about polysemy below
plɔ́ɔŋ	lower leg	from knee to ankle
krwàac	lower leg	from knee to toe, less common
plù	thigh	
knúun	knee	
pltàk	hollow of the knee	
prcòl	heel	
prcèr	area above the heel	
ŋɔ̀ɔŋ-ŋɔ̀ɔŋ	shin	unanalyzable reduplication
Complex terms		
tlmɨaŋ plù	thigh	'segment thigh'
tlmɨaŋ plɔ́ɔŋ	lower leg	'segment lower.leg'
plé plɔ́ɔŋ	calf muscle	'fruit lower.leg'
klɔ́ɔŋ lìap	knee cap	'stone lìap.fruit'
póom róoy	ankle protuberance (medial malleolus)	'protuberance spirit' (Lao *pom*)

trpɔ̀k plɬam	area where foot meets ankle	'passing leech', see below
ktáak cɬaŋ	sole of the foot	'palm leg'
crkùul cɬaŋ	toe	'digit leg'
krlèey plù	hip-joint	'joint thigh'
krlèey cɬaŋ	ankle	'joint leg'
srŋɛ̀ɛk cɬaŋ	area just above the ankle	'inward.bend leg', from *ŋɛ̀ɛk* 'bent inwards'
tmmɔ́ɔŋ cɬaŋ	toenail	'nail leg'
mà cɬaŋ	big toe	'mother leg'
kɔ́ɔn (tàam) cɬaŋ	little toe	'child (outgrowth) leg'

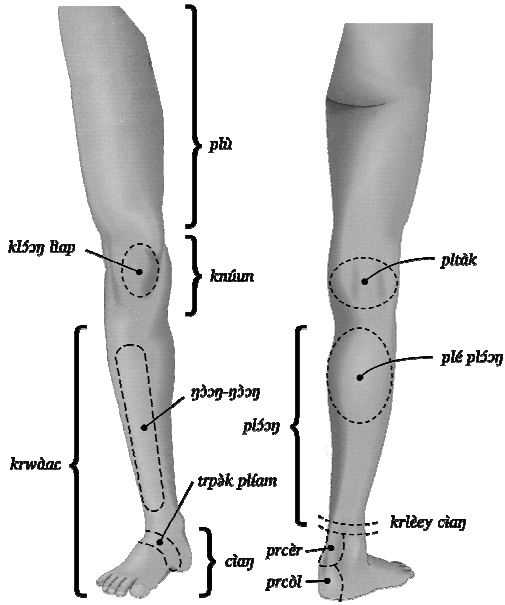


Figure 4. The leg (body drawings from Enfield 2006a)

At first sight, the divisions of the upper and the lower extremities are very similar: they both have two main parts and three main joints, with *ktáak*, *crkùul* and *tmmɔ́ɔŋ* at the tips. However, the term *cɬaŋ* is elusive and behaves differently from the supposedly parallell term *tí*.

When my informant was presented with *tí* and *cɨaŋ* as context-less isolated words in the body coloring task, he marked them both as covering the entire upper and lower extremity, respectively. But as soon as the words were mentioned in a context that involves a specific location, e.g. 'he has an ant-bite/a mole on his *tí/cɨaŋ*' or 'his *tí/cɨaŋ* hurts/burns', the scope of *cɨaŋ* shrunk to cover only the foot. Similar contextual use did not affect the scope of *tí*.

In another kind of drawing task I asked my informant to mark a few antbites on a man's *cɨaŋ*. He began marking a few dots on the foot but stopped at the ankle. At this point I still thought of *cɨaŋ* as the entire leg, so I pointed to the calf and asked if that was not *cɨaŋ* as well, but he answered "no, that's his *plɔ́ŋ*". Again, this 'contextual narrowing' applies only to the lower extremity. Antbites on a man's upper arm can be called antbites on his *tí* if no further precision is needed, but antbites on a man's thigh are always antbites on his *plù*. Antbites on a man's *cɨaŋ* are located below his ankle, and *cɨaŋ* cannot be 'expanded' to cover the thigh in these contexts. Similarly, 'to spread one's legs' is *cltàh plù*.

A telling example of the difference between *tí* and *cɨaŋ* is their use together with the word *tlmɨaŋ* 'segment'. The *tí* can be divided into two *tlmɨaŋ tí* 'arm segments', i.e. the upper arm and the forearm. There is, however, no **tlmɨaŋ cɨaŋ*. Instead, the thigh and the calf are segments of their own: *tlmɨaŋ plù* and *tlmɨaŋ plɔ́ŋ*. They are not segments of anything else.

cɨaŋ can refer to the entire leg only as long as there is no explicitly mentioned location that could cause a 'contextual narrowing' (e.g. place of antbite/mole/pain etc.). The polysemy of *cɨaŋ* is one of *vagueness*. This means that *cɨaŋ* can be ascribed one quality that activates the 'contextual narrowing' and at the same time another quality that applies to the entire leg:

(8) cɨaŋ ò cú dɛ́ɛ skár cŋnòŋ dɛ́ɛ
leg 1SG hurt also straight stiff-legged also
'My *cɨ̀aŋ* hurts and is (also) stiff'

Some of the other terms found on the leg also deserve further mentioning. For example the complex term *trpə̀k plɨ́am* that covers the perceptually inconsiderable area where the foot meets the ankle. The term translates to 'leech passing', simply because leeches often attach themselves to this area.

Table 7. Genitalia

Kammu	English	Details
Simplex terms		
plwà	crotch	
lòk, tlé	penis	
ké, kàn	vulva, vagina	
rám	vulva, vagina	*Rɔ̀ɔk* dialect
kláa	scrotum	
héel	(protruding) labia	
Complex terms		
któŋ kláa	testicle	'egg scrotum'
khúul lòk / ké	pubic hair	'hair penis / vulva'
kmpóŋ lòk	head of penis	'head penis'
kmpóŋ ké	clitoris	'head vagina'
hmpúur lòk	foreskin	'skin penis'
hmpúur ké	labia	'skin vagina'
pràŋ plù	crotch	'between thighs'
tè cmkín / cmrɔ̀	privates	'of woman / man'

Terms for genitalia are used in everyday language and do not always carry the same strong sexual connotations that they often do in English.

> [A story] certainly does not become risky, just because the intimate parts of the body are mentioned by name. […] There seems to be no [equivalent] set of completely neutral, everyday words in English pertaining to sexual life. (Lindell et al 1980:25f)

Nevertheless, many of these words can be used in expressions of frustration e.g. *tlé mɘ́h?* 'what the dick?' or insults *ké mà mèe!* 'your mother's pussy!'.

5.7 Internal organs

Table 8. Internal organs and skeleton

Kammu	English	Details
Simplex terms		
ríaŋ	intestines, bowels, guts	
hrñɨam	heart, mind, breath	
klán	kidney	
hntáŋ	brain	
ktúh	stomach, abdomen	
kmnùum	urinary bladder	from *nùum* ‘urine’
lŋàr	bone marrow	
páaŋ	spleen	(Lao *paaŋ2*)
cʔáaŋ	bone	
krnɔ̀ɔn	womb, uterus	from *kɔ́ɔn* ‘child’
Complex terms		
cʔáaŋ kmpóŋ	cranium	‘bone head’
cʔáaŋ klʔùŋ	hipbone	‘bone hip’
cʔáaŋ kntrɔ̀ɔŋ	spine	‘bone back’
cʔáaŋ àk	collar-bone, breast bone	‘bone chest’
cʔáaŋ kràp	gristle	‘bone crispy’, from *krùp* ‘crunching sound’
cʔáaŋ trak	ribs	‘bone side’
(cʔáaŋ) rŋsɔ́ɔŋ mùh	nose cartilage	‘(bone) *rŋsɔ́ɔŋ* nose’
hmpúuy kɔ́ɔn	placenta	‘nest child’
tlɔ́ɔm (ràŋ)	liver	‘liver (*ràŋ*)’ (Lao *raaŋ2*)
tlɔ́ɔm túus	lungs	‘liver foam’, see below
ríaŋ cə̀	large intestine, colon	‘intestine dirty’
ríaŋ sít	large intestine, colon	‘intestine end’ (Lao *sut*)
ríaŋ tŋkà	appendix	‘intestine forked’

ríaŋ plìa	small intestine	‘intestine clean’
tlŋòk hrñìam	heart muscle	‘stalk heart’
tlŋòk prcáŋ	bile duct	‘stalk bile’
hmpúuy prcáŋ	gall bladder	‘nest bile’
rmmé nám	large sinew	‘sinew large’ (e.g. Achilles’ tendon)
rmmé nè	small sinew	‘sinew small’
rmmé ñèer	thin sinew	‘sinew thin’
rmmé màam	blood vessel	‘sinew blood’
rmmé cŋáar	blood vessel	‘sinew green’ (covers Eng. ‘blue’)

The Kammu do not see their own internal organs more often than other people do, so the knowledge about them comes mainly from observations of animal organs from hens, pigs, buffaloes and other animals. When an animal is killed, all of the organs are usually eaten, except for the urinary bladder which can be saved and used as a ball. Some of the inner organs are carefully divided and eaten during certain sacrificial ceremonies (Tayanin 2006:60ff).

The term *hrñìam* is widely used in expressions describing emotion and attitude, either as the head of an adjectival phrase or as an argument to a verb. Some of these expressions can be translated verbatim into English, perhaps with a poetic ring to them, e.g. *hrñìam kəə lə* ‘his heart is good’. Others have a different meaning than their English translation, e.g. *nám hrñìam* ‘big heart’ meaning ‘brave’, and *hrñìam kəə rɔɔn* ‘he has a hot heart’ meaning ‘he is impatient’.

Another aspect of *hrñìam* concerns breathing. This conceptual link is found in many languages, for example between the two originally Latin words *spirit* and *respiration*. The common word for ‘to breathe’ is *tóh hrñìam* lit. ‘to push out breath’, and ‘to hold one’s breath’ can be expressed as *ìt hrñìam* lit. ‘to stop (one’s) breath’.

The heart muscle is called *tlŋòk hrñìam* ‘heart stalk’, likening the heart to a (fruit with a) stalk. This stalk provides the entire body with *hrñìam*, allowing us to move, think and live. The gall bladder, on the other hand, is called *hmpúuy prcáŋ* ‘bile nest’ with *tlŋòk prcáŋ* ‘bile stalk’ referring to the bile duct, which leads bile to the digestive system.

The word *tlɔ́ɔm* means ‘liver’ on its own (as do its cognates in other Mon-Khmer languages), but it is sometimes used with an unanalyzable modifier *ràŋ* in order to contrast the liver with the lungs, *tlɔ́ɔm túus*. The modifier *túus* is understood as ‘foam’, but this meaning is likely a result of folk

etymology. A more plausible origin is that *túus* comes from Proto-Mon-Khmer **tɔh* meaning 'breast'. A similar form is found in the closely related Waic languages (Shorto 2006:508).

5.8 Bodily fluids and emissions

Table 9. Bodily fluids and emissions

Kammu	English	Details
Simplex terms		
màam	blood	
hɨ́al	vomit	
plúŋ	pus	
ʔyíak	shit, residue	
rùk	dirt	e.g. behind the ear
kmháak	phlegm in the throat	
Complex terms		
ʔyíak ráaŋ	plaque	'shit tooth'
ʔyíak hɨ̀r	ear wax	'shit whirl'
èɛc pèɛc	eye matter	'*èɛc* eye.matter'
hmpúut kmpóŋ	dandruff	'shedded.skin head'
òm mùh	runny snot	'liquid nose'
òm nùum	urine	'liquid urine'
òm màt	tears	'liquid eye'
òm hmlù	sweat	'liquid sweat'
òm ɨ̀n	sweat	'liquid damp' (Lao *ʔɨn*)
òm tʔá	saliva	'liquid saliva'
òm kcɔ́ɔr	saliva (running), spittle	'liquid saliva'
òm kcúh	spittle	'liquid spit'
òm prcáŋ	bile	'liquid bile', possibly from *cáŋ* 'bitter'
òm ké	vaginal secretion	'liquid vagina'

òm ñìir	amniotic fluid	'liquid slippery'
òm tlé	semen	'liquid penis'
òm pù	milk	'liquid breast'

òm is probably the most common head in complex body part terms. It is used about all kinds of fluids, from various kinds of honey to the many rivers and streams meandering through the traditional Kammu homeland.

Some of the compounds with *òm* have a body part as their modifier, but others are seemingly pleonastic combinations of *òm* plus a modifier that translates to a bodily fluid on its own. This might be another case of analogical leveling similar to the one with *khúul* suggested earlier.

6. **Categorization of body part terms**

My informant often pointed out that the phrases used for this investigation sound strange because of their content. Their syntactical structure is the same as in fully acceptable everyday questions and statements like 'where is your knife?' and 'my basket is in the house', but the whole scenario is implausible: why would anyone be uncertain about clearly visible traits of something so well-known as our own human body?

Nonetheless, some of these phrases are perceived as wrong and others as correct, allowing me to use them as a basis for the following categorizations.

6.1 Partonomic categorization

As already mentioned, Kammu has no word corresponding to the English 'body part' or even 'part'. This in itself cannot disprove the existence of a body partonomy, but it does make it impossible to prove that other methods of investigation could access the partonomic categorization. The 'entailment test' (see 4.1) has been suggested as one such 'access way', but in Kammu this method seems to access a possessive categorization. Of course it could be the case that both the entailment test and the possessive test access a partonomic hierarchy, but again: it is impossible to prove. These two categorizations will therefore be described in the section below instead.

6.2 Entailment-possessive categorization

During the analysis of data from the different kinds of tests it soon became clear that the entailment test and the possessive test are merely two ways of accessing the same categorization. Accordingly, this will be called the *entailment-possessive categorization*.

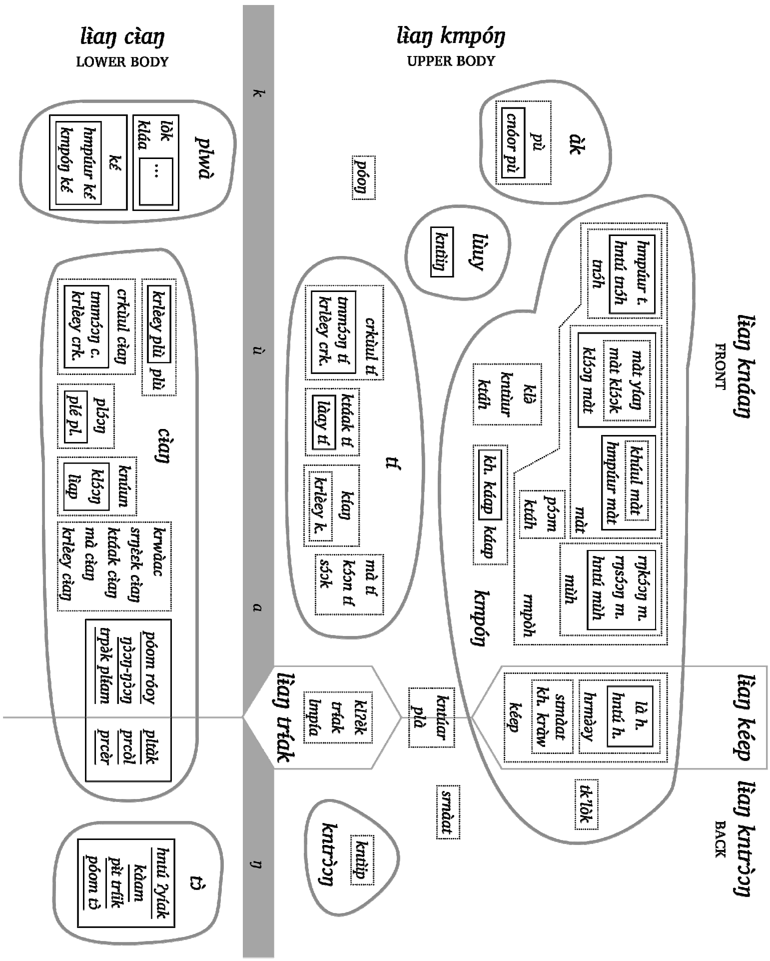


Figure 5. Graphic representation of the possessive hierarchy

The main difference between these ‘access ways’ is that the entailment test allows the categorization to be more transitive. The sentence *‘his finger joints are hurt but his *tí* is not’ is perceived as wrong, indicating a transitivity all the way from the finger joints to *tí*, but the equivalent possessive phrase ?‘my *tí* has finger joints’ sounds odd.

Figure 5 is a graphic representation of the partonomic-possessive categorization. The thin black lines spanning the entire figure are divisions between the upper/lower body and front/back, with special areas for the side of the torso and the side of the head. The waist, *kùaŋ*, divides the upper and lower body and is therefore placed in the middle, pertaining to neither half.

Terms that are possessed by *lɔ̀h* and themselves possess other terms are written in a larger typeface inside areas with smooth grey borders. These constitute the first level of the hierarchy.

The boxes inside the grey borders contain the further levels of the hierarchy. The content of a box is always accessible to the term immediately possessing it. Beyond that, boxes with dotted borders ('dotted boxes') are transitive, and those with solid borders ('solid boxes') are intransitive.

For example, *hrmàəy* is immediately possessed by *kéep* which is immediately possessed by *kmpóŋ*. Since both *hrmàəy* and *kéep* are in transitive dotted boxes, *hrmàəy* is also possessed by *kmpóŋ*. On the other hand, *lá hrmàəy* is placed in a solid box, showing that it is possessed only by *hrmàəy* and not by *kéep* nor *kmpóŋ*. Abbreviated heads in these figure are mainly *kh.* for *khúul* and *kr.* for *krlèey*.

The investigation of the possessive system in Kammu is somewhat obstructed by a phenomenon called *kàm kùu* 'word pair'. Its many interesting features could easily fill an entire thesis, but the aspect that is most relevant for this paper is that some terms can only be possessed in pairs! Mentioning only one of the two (or mentioning both but in reverse order) will make the phrase sound incomplete and even ungrammatical. My informant compares it to *táəm hóŋ háac* 'singing without completing the rhymes'. The closest equivalent in English is idiomatic pairs such as 'leaps and bounds' or 'high and dry'.

The surprising result is that *any* term is immediately possessed by *lɔ̀h* as long as it occurs in a word pair. My informant smiled at the sentence 'a body has nails' since it made him think of a person with nails all over his body 'like a scaly ant-eater'. But he then added 'finger' to get the perfectly acceptable 'a body has fingers, has nails', *lɔ̀h àh crkùul àh tmmɔ́ɔŋ*.

As an isolated word pair, one could think of this as 'my body has fingers that have nails', but far from all word pairs consist of such adjacent terms. In fact, the other main category of body word pairs consists of spatial or conceptual antonyms such as *kmpóŋ* 'head' ↔ *tɔ̀* 'rump'. There are even a few pairs in which the second word is a 'buffer word' with no meaning of its own, e.g. *tnɔ́h* 'mouth' ↔ *kñòor*.

Apart from *lɔ̀h*, the other body part terms do not have the same need for word pairs. It is perfectly acceptable to say 'a head has ears' or 'a foot has a heel'. Some terms do not occur in any word pair and can therefore not be possessed directly by *lɔ̀h*. These terms are underlined in figure 5.

The entailment-possessive system shows a great deal of categorization for the head and the limbs, but the torso is more scattered with many isolated terms encompassing very few other terms.

Figure 5 clearly shows that the naming of complex terms very often reflects their categorization. If the modifier of a complex term is another body part term, that term immediately possesses the complex term. The modifier is usually omitted: *hrmàəy àh lá (?hrmàəy)* 'the ear has a (ear) conch'.

This principle does have its exceptions. For example, fingernails are named *tmmɔ́ɔŋ tí*, but my informant sees them as possessed by the fingers rather than by *tí* (this is still in line with the discussion about the nails in 5.5, since the non-ambiguous *crkùul tí* is used here). Similarly, wrinkles in one's hand, *làay tí*, are possessed by the palm, *ktáak tí*. The eyelashes, *khúul màt*, are possessed by the eyelid, *hmpúur màt*, rather than by *màt*.

6.3 Spatial categorization

Figure 6 shows the spatial relations between a subset of external body part terms, superimposed on a cartoon-like human body.

Table 10. Graphic representation of spatial relations in Figure 6

Relation	Kammu reading	English reading
X —o Y	*X yèt tàa Y*	X is located on/at Y
X ←| Y	*X yèt tal Y*	X is located at the tip/end of Y
X —| Y	*X yèt túut Y*	X is located at the base of Y
X ↓ Y	*Y yèt kntrùum X*	Y is located below X
X ↑ Y	*X yèt prwàəy Y*	X is located above Y (if both 'below' and 'above' apply to a pair of terms, the two arrows are combined into one)
Z X —+— Y	*Z yèt trtí X yʌ̀ʌ Y*	Z is located between X and Y

Finally, the three terms inside the oval 'mouth' are expressed as located *klúaŋ tnɔ́h* 'inside the mouth'.

In order not to clutter the overview more than necessary, a few relations are expressed through underlining. The underlined terms on the arm can all be expressed as being located *tàa tí*, and the underlined terms on the head as *tàa kmpóŋ*. The marking of *tí* at the shoulder joint is just a technicality – *tí* is an encompassing term. Once again, the fingernails do not have any immediate relationship to *tí* but rather to the fingers.

6.4 Differences between the categorizations

The spatial categorization clearly shows a narrowing of *cìaŋ* compared with the possessive categorization. The ankle can be expressed as being located between *cìaŋ* and the calf, for example. *cìaŋ* possesses the knee, but it is wrong to say that the knee is located on *cìaŋ*.

The spatial system contains less hierarchical boundaries, e.g. the navel is located on the belly, but both the navel and the belly can be expressed as being located above *póoŋ*. Similarly, *kàam* 'lower back' is located both on the waist and on the behind.

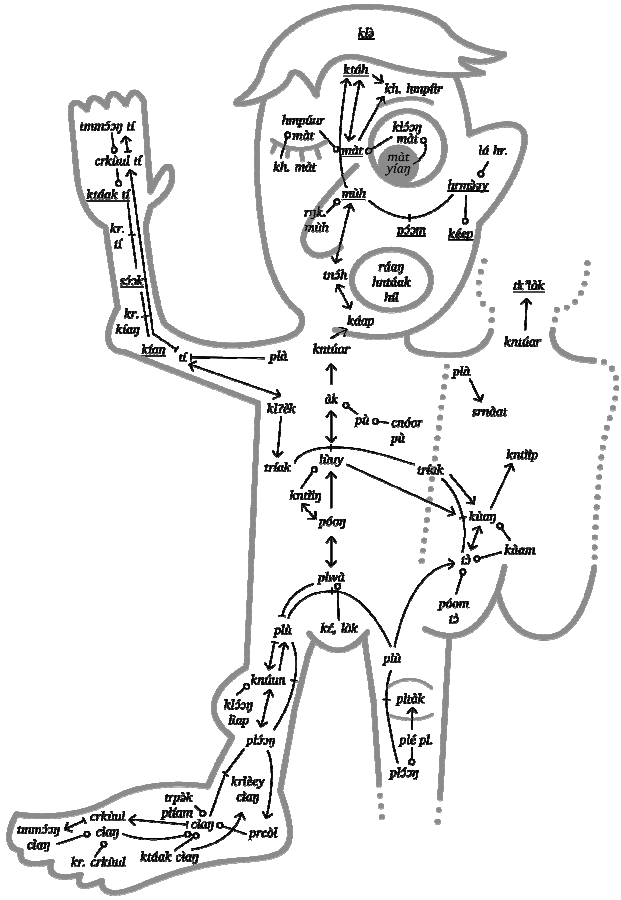


Figure 6. Spatial relations between a subset of external body part terms

6.5 Universals

With the body part terms listed and their categorization analyzed, it is now possible to see how the body part universals apply to Kammu.

a) The BODY is labeled in all body-part partonomies.
Yes. The term *lɔ̀h* covers the entire body, see more under b).

b) Every language includes a term for HEAD in its lexical field of body-parts,
Yes. The term *kmpóŋ* covers the entire head.
…and the term is always immediately possessed by BODY.
Yes. The term *kmpóŋ* can be expressed as possessed by *lɔ̀h*.
Other categories which usually occur at the second level of the partonomy include TRUNK,
No! There is no unique term for 'trunk' or 'torso'. The term *lɔ̀h* is polysemous, meaning both 'body' and 'torso'. However, it is not possible to say that *lɔ̀h* 'body' possesses *lɔ̀h* 'torso'.
…ARM (plus HAND) and LEG (plus FOOT).
Yes. The terms *tí* and *cɨaŋ*, respectively.

c) All languages label EYES, NOSE and MOUTH.
Yes. The terms *màt*, *mùh* and *tnɔ́h*, respectively.

d) The ARM (plus HAND) is named by a distinct term in all languages.
Yes. The term *tí*.

e) The categories FINGER and TOE are always labeled (by one of four general patterns).
Yes. But it is unclear whether or not the Kammu system can be restricted to only one of these four patterns. The existence of both *crkùul tí* and *crkùul cɨaŋ* suits pattern 3: "Different terms derived from the same root". But *crkùul* on its own can also mean 'finger', whereas the modifier *cɨaŋ* is always preferred for 'toe'. This suggests pattern 4: "One basic unanalayzable term for FINGER, with TOE derived from it". It seems that many languages which follow pattern 3 can use the 'root' term on its own to refer to a finger, so the line between patterns 3 and 4 may be too arbitrary.

f) All languages name (FINGER)NAIL and (TOE)NAIL by one of two patterns. Languages with the first pattern have one basic term applied to both categories, and the second pattern have different terms derived from a common root, like English 'fingernail' and 'toenail'.
Yes. But again the distinction between the patterns is very vague. *tmmɔ́ŋ* 'nail' *cɨaŋ* be used on its own if the context is clear, but if a distinction is important, the modifiers *tí* or *cɨaŋ* are added.

g) A term for LEG implies a separate term for ARM.
Yes. *cɨaŋ* can cover the entire leg, and the term *tí* covers the entire arm.

h) A term for FOOT implies a separate, non-identical term for HAND.
No! There are contexts where *cɨaŋ* covers the foot and nothing else, but the scope of *tí* is never narrowed in the same way.

i) Terms for INDIVIDUAL TOES imply terms for INDIVIDUAL FINGERS.

Yes. Fingers and toes have the same level of detail, with separate terms only for the thumb/big toe and little finger/little toe.

Concerning the 'depth principle' there is indeed no documented hierarchical chain stretching beyond six levels. The few terms that are found on this sixth level are the white and the 'black' of the eye, and the eyelashes.

7 Summary

This is the first systematic study of the body part domain in the Kammu language. The relations between body part terms can be ordered in a hierarchical possessive system, similar to the partonomic system proposed in earlier literature. Since Kammu has no word for '(body) part' it is not possible to show that this system is a partonomy. But at the same time, it is not unreasonable to suggest so.

This paper also describes a spatial system in which body part terms relate to each other through prepositions. This supports the more recent claims of several coexisting relational systems.

The modifier in complex body part terms often reflects the term's position in the possessive system. The inconsistencies that do exist show that modifiers can convey semantic information other than 'hierarchical subordination'. A hierarchy cannot be deduced solely from a word list, but requires an investigation of its own.

Kammu adheres to most of the posited body part domain universals. One of the more unusual deviations is that the leg can be divided with more detail than the arm. There are, in other words, situations where the term that usually covers the entire leg only covers the foot, but the term for the arm never shrinks its scope to cover only the hand.

An interesting subject for future studies would be to investigate in what aspects, and to what degree, the ever-increasing Lao influence has changed the Kammu body part domain over the last 30 years.

REFERENCES

Andersen, Elaine S. 1978. "Lexical universals of body-part terminology." In Greenberg, Joseph H, ed., *UHL*, pp. 335-368. Vol. 3: Word structure. Stanford, California: Stanford University Press.

Brown, Cecil H. 1976. "General principles of human anatomical partonomy and speculations on the growth of partonomic nomenclature." *AE* 3(3):400-424.

Brown, Cecil H.; John Kolar, Barbara Torrey, Tipawan Trương Quang, Phillip Volkman. 1976. "Some General Principles of Biological and Non-Biological Folk Classification." *AE* 3(1):73-85.

Burenhult, Niclas. 2006. “Body part terms in Jahai.” In Majid et al, eds., pp.162-180.
Enfield, Nick J. 2006a. “Elicitation guide on parts of the body.” In Majid et al, eds., pp.148-157.
Enfield, Nick J. 2006b. 1991. “Lao body part terms.” In Majid et al, eds., pp.181-200.
Enfield, Nick J.; Asifa Majid, Miriam van Staden. 2006. “Cross-linguistic categorisation of the body: Introduction.” In Majid et al, eds., pp.137-147.
Gaby, Alice R. 2006. “The Thaayorre ‘true man’: Lexicon of the human body in an Australian language.” In Majid et al, eds., pp.201-220.
Levinson, Stephen C. 2006. “Parts of the body in Yélî Dnye, the Papuan language of Rossel Island.” In Majid et al, eds., pp.221-240.
Lindell, Kristina; Jan-Öjvind Swahn, Damrong Tayanin, eds. 1980. “Folk Tales from Kammu II: A Story-Teller's Tales.” London and Malmö: Curzon Press.
Lundström, Håkan & Svantesson, Jan-Olof, eds. 2005. “Kammu - om ett folk i Laos (Kammu - about a people in Laos).” Lund: Lunds universitetshistoriska sällskap.
Majid, Asifa. 2006. “Body part categorisation in Punjabi.” In Majid et al, eds., pp.241-261.
Majid, Asifa; Enfield, Nick J.; van Staden, Miriam, eds. 2006. “Parts of the body: cross-linguistic categorisation.” Special Issue of *LS* 28(2-3):137-360.
Meira, Sérgio. 2006. “Tiriyó body part terms.” In Majid et al, eds., pp.262-279.
Palmer, Gary B. & Nicodemus, Lawrence. 1985. “Coeur d’Alene exceptions to proposed universals of anatomical nomenclature.” *AE* 12(2):341-359.
Shorto, Harry L. 2006. “A Mon-Khmer Comparative Dictionary.” In Paul Sidwell, Doug Cooper, and Christian Bauer, eds., *PL* 579. Canberra, Australian National University.
Svantesson, Jan-Olof. 1983. “Kammu Phonology and Morphology.” Travaux de L’Institut de Linguistique de Lund XVIII. Lund: Liber Förlag.
Svantesson, Jan-Olof; Damrong Tayanin, Kristina Lindell, Håkan Lundström (forthcoming) “Kammu Yùan–English Dictionary.”
Tayanin, Damrong; Kristina Lindell (editor), Arthur Holmer (assistant editor). 2006. “Where do our souls go after death? Kammu traditions, rituals and ceremonies.”
Terrill, Angela. 2006. “Body part terms in Lavukaleve, a Papuan language of the Solomon Islands.” In Majid et al, eds., pp.304-322.
van Staden, Miriam & Majid, Asifa. 2006. “Body colouring task.” In Majid et al, eds., pp.158-161.
Wegener, Claudia. 2006. “Savosavo body part terminology.” In Majid et al, eds., pp.344-359.

Received: 25 January 2008

Box 117 S-221 00 Lund,
Lund University, SWEDEN
<felix.ahlner@gmail.com>

Copula of identity in Old Khmer

Pogibenko TAMARA
Institute of Oriental Studies, Russia

Abstract

This article is part of my work "Copula of identity in Old Khmer and Modern Khmer". The subject of this article is the word *gi* (*giy*, *gui*) in Old Khmer - the language of the inscriptions of Cambodia from the early seventh to the middle of the fourteenth centuries A.D. In Old Khmer inscriptions this element is frequently used, but its functions are sometimes difficult to determine. It is found that in Old Khmer *gi* is used in two functions: as a deictic word with the meaning 'this', 'here' and as a copula of identity, which is to be distinguished from the copula of predication *jā*. In Modern Khmer the first function was lost and the second function had further developments. So, Old Khmer is a language where qualifying or characterizing sentences are distinguished from equational or identifying sentences on formal grounds and where the copula of identity is the result of grammaticalization of a demonstrative.

Three different usages of the copula of identity *gi* are analyzed: in equational sentences, where it is used for identification, in clauses of consequence, where it functions as a conjunction, and in sentences with rheme shift, where it is used as a marker of rheme shift. It is shown that the function of the conjunction of consequence is the result of restructuring of sentences with a more complex structure where *gi* was initially used as a copula if identity. It is also asserted that in sentences with rheme shift *gi* was initially motivated from the semantic level, the semantic representation of such sentences being an equation. It is suggested that three types of sentences with rheme shift found in Old Khmer, i.e. sentences without syntactic restructuring, sentences with *gi*-movement to the rhematic NP, and cleft sentences, represent a scale of restructuring that leveled the asymmetry of the syntactic and the semantic structure. Contextual analysis of some examples shows that in Old Khmer *gi* was a sophisticated tool that made it possible to express subtle nuances in Old Khmer discourse.

Introduction

This article deals with the copula of identity *gi* (*giy*, *gui*) in Old Khmer. This element is frequently used both in Old Khmer and Modern Khmer and plays an important part in organising discourse structure.

Generally in languages a copula of identity is found in equational or identifying sentences of the type A = B, where A and B are different nominations of one and the same object (*This man is Socrates*). They are to be contrasted with and distinguished from qualifying or characterising sentences, i.e. bi-nominal sentences formed by two NPs with the second NP being a predicate (*This man is a philosopher*). In languages these two sentence types

can have identical morpho-syntactic structure, in Khmer they are formally contrasted: the copulas *gi* (Old Khmer), *kuuu* (Modern Khmer) are used in equational sentences, the copulas *jā* (Old Khmer), *ciə* (Modern Khmer) - in qualifying sentences. So languages like Khmer provide good evidence that equational and qualifying sentences can be distinguished on formal grounds, and the two types of statements - statements about identity of two objects and statements attributing some quality to an object are formally contrasted.

The copula of identity *gi* (*giy*, *gui*) presumably originates from a word with deictic type of meaning 'this', 'here'. The original deictic usage of *gi* is preserved in Old Khmer. The copular of identity in Old Khmer is used in the function of identification and in sentences with rheme shift. Thus, in (Long Seam 2000) all these functions are mentioned: "particule à valeur explicative, démonstrative et locative; à savoir; s'emploie aussi dans l'énumération". However, in (Sak-Humphry 1993) *gi* is treated as a copula of predication and its other uses are neglected, even the evident cases of deictic uses.

Deictic function of *gi*

In deictic function *gi* can be used as a noun head or a modifier. As an NP head *gi* is found in argument and locative NPs. It is also often found in temporal NPs. Consider the examples from the inscription K 259 (VIII AD), where *gi* is found in all the above mentioned deictic functions: gi_1 - in NP_{TEMP}, gi_2 - as a head in NP_{OBJ}, gi_3 - as a modifier in NP_{OBJ}, gi_4 - as a head in NP_{LOC}.

(1) *loḥ ta gi rājya vraḥ kamratāṅ añ çrījayadevī _ _ _ ājña vraḥ kamratāṅ añ çrījayadevī oy dana ta vraḥ kamratāṅ añ*
'During the reign of V. K. A. Srijayadevi (*lit*. 'during *ta* gi_1 reign') _ _ _ the order of V. K. A. Srijayadevi to give a donation to God ' (K 259)

(2) *ge ta cap gi ge ta sak gi ge ta soṃ gi sot ge ta pre roḥ gi neḥ bhūmidā ta nai vraḥ*
'Those who seize this (=gi_2), those who rob this (=gi_2), those who demand this (=gi_2), as well as those who use this (=gi_3) property, this land of God' (K 259)

(3) *voṃ jā anak hau prativeça ta gi voṃ jā anak cap dai ta gi _ _ _ kñuṃ tmur krapi rddeḥ _ _ _ ta gi dai ge ta thve talaṅ ta gi ge tve yatheṣa ta gi*
'It is prohibited for people to call up people from neighboring lands here (=*ta* gi_4), it is prohibited for people to capture other people here (=*ta* gi_4), _ _ _ slaves, bulls, buffalo, carts _ _ _ here (=*ta* gi_4), those who destroy here (=*ta* gi_4), those who act at will here (=*ta* gi_4)' (K 259)

Equational sentences: Identification

Equational sentences are used for identification in two cases: for *specifying* or *interpreting* identification and for *clarifying* identification (Shmelov 2002: 187).

Specifying or interpreting identification is found in sentences A *gi* B, where A and B are two different nominations of one and the same object and where the nomination representing the first component of equation is not informative enough to enable the listener to locate the referent in the relevant denotative domain. In this case the nomination represented in the second component of the statement of equation promotes the required felicity conditions of the speech act and allows the addressee to establish the reference.

Clarifying identification - that is when the nomination represented in the first component of equation is informative enough to enable the addressee to locate the referent in the relevant denotative domain, but this localization remains ineffective as long as the addressee is unable to relate the newly-acquired information to some information from his previous experience. This localization cannot be complete due to the fact that in the addressee's memory there exists another "mental dossier" on this same object (Shmelov 2002: 187). Without the knowledge of a certain identity the addressee is unable to make a comparison of "qualities, features, facts, etc., given in direct observation or received through information channels, with the information and impressions coming in from past experience" (Arutunova 1998: 11). In that case the equational sentence provides the required identity, promoting the felicity conditions of the speech act, and the nomination represented in the second component of the statement of equation allows the addressee to establish the identity of the object.

Equational sentences used for identification are very common in Old Khmer texts. Their frequent usage reflects the specific character of Old Khmer discourse represented in the inscriptions. In texts describing acts of donations to God a great precision was required when identifying "What is what" and "Who is who". Consequently messages carved in stone had to be informative enough to enable the addressee to correctly locate the referents in the denotative domain; therefore a lot of equational sentences with *gi* were used to promote the required felicity conditions of the speech act. The same is true of the texts representing the king's edicts and the texts, regulating property rights and duties. Identifying objects mentioned in the inscription had to be an integral part of logical and communicative structure of such discourses.

In Old Khmer the most common case of specifying identification is when the first component is a generalizing nomination. In that case identity relation connects the generalizing nomination and the nomination or the set of nominations disclosing its content. For example:

(4) *srū man dār sin tloṅ 10 dmār gui poñ cpoṅ poñ vreṅ*
'Rice that has been received - 10 measures tlong, the recipients (are) (*gui*) Ponh Chpong and Ponh Vreng' (K 79)

(5) *jnāhv gui canlek aṃval vlaḥ*
'the price (is) (*gui*) a set of clothing amval' (K 79)

(6) *piṇḍa gui sre tloṅ 9 je 1*
'Altogether (*gui*) ricefields 9 tlong and 1 je' (K 49)

The components of equational sentence A *gi* B, where B is an interpreting nomination, can have the order B *gi* A, which is quite natural and results from the symmetry property of equation (A = B is the same as B = A). Thus, for instance, in the inscription K 56A the generalizing nomination follows a long list of objects donated.

(7) *gui caṃmāṃ roḥh kalpanā yajamāna*
'(these are) (*gui*) the supplies of the donator's construction' (K 56A)

It appears that Coedès failed to take into account the possibility of the inversion of the components of equation since in a footnote to this sentence he remarks: "On voit mal quelle relation *gi* sert à établir avec ce qui précède" (IC VII: 16).

What is more, the order *gi* A B is also possible, for example:

(8) *gui neḥ _ ndāna tāṅ añ*
'(*gui*) this _ donation of Tang Anh' (K 98)

One of the common cases of specifying or interpreting identification is the usage of the copula *gi* before quantitative and parametric descriptions:

(9) *sre mā vaṃraḥ gui 20*
'rice field (measuring) (one) ma and 20 vamrah (*lit.* 'vamrah *gui* 20' (K 926)

(10) *aṃruṅ gui mā je 2*
'the size (of the rice field) (*gui*) (one) ma and 2 je' (K 79)

Equational sentences are also used when the first component is an NP with the word "edict" in head position - a nomination which is not informative enough to enable the addressee to absorb the full scope of the message. The nomination represented in the second component of equation allows the addressee to understand the content within the required scope. In this case the semantic relation that holds between the nominalisation of a speech act and the component specifying its content is labeled as an identity relation. In such sentences the word *ni* 'about, concerning' can also be present. Examples:

(11) *ājñā vraḥ kamratāṅ añ ni gui puṇya pu caḥ añ ratnabhānu*
'The king's edict concerning (=*ni gui*) pious accomplishments of Pu Chah Anh Ratnabhanu' (K 49)

(12) *ājñā vraḥ kamratāṅ añ ni giy aṃras mratāñ anaṅga phoṅ giy ka tel mratāñ oy phoṅ*
'The king's edict concerning (=*ni giy*) the servants of Mratanh Ananga, that (=*gui ka tel*) Mratanh Ananga gave' (K 502)

The last sentence provides an interesting example of copula resumption: *giy* is resumed before the relative clause linked to the second component of the identity by relativisers *ka* and *tel*. The pluraliser *phoṅ* is resumed in the relative clause as well.

Consider an example from the inscription K 940. This inscription represents the king's edict ordering to deliver some amount of salt to the place Tirthagrama (a village with a quay). The salt was to be delivered on boats belonging to different temples. In this example the copula of identity is used in three different types of identification:

gui$_1$ - is used for specifying or interpreting identification and puts into identity relation the nomination "king's order" and the nomination describing the content of the order;

gui$_2$ - is also used for specifying identification and puts into identity relation the nomination "one delivery of salt" and the quantitative description "three (measures) knanh";

gui$_3$ - is used for clarifying identification and puts into identity relation the description of the content of the order and some "mental dossier" - information coming from the previous discourse "king's order concerning allocation of responsibilities between the temples";

gui$_4$ - defies unambiguous interpretation, one of the possible interpretations is the same function as the function of *gui*$_5$;

gui$_5$ - has a deictic function - head of NP_{OBJ}.

(13) *ājñā vraḥ kamratāṅ añ ni gui dok vraḥ kamratāṅ añ çrīpiṅgaleçvara moyy aṃpel ple gui knañ 3 dok kpoñ kamratāṅ añ kamratāṅ slot moyy aṃpel ple gui knañ 3*
'King's order concerning (=*ni*): (*gui*$_1$) the boat (of the temple) V.K.A. Sripingaleshvara - one delivery of salt, *(gui*$_2$*)* 3 (measures) knanh; the boat (of the temple) Kponh K. A. Kamratang Slot - one delivery of salt, *(gui*$_2$*)* 3 (measures) knanh' (...)

gui tel ājñā vraḥ kamratāṅ añ pre pek āy tīrthagrāma gui cuḥ lah tleṅ lah ge ta ckop gui ge ta dap gui ge cmer ājñā ge daṇḍa
'That is (=*gui*$_3$) what the king's order charges to distribute at Tirthagrama (*gui*$_4$), increases or decreases, who levies tax upon this (=*gui*$_5$), who hampers this (=*gui*$_5$), who transgresses the king's order will be punished.' (K 940)

As it seems, Coedès does not always distinguish the different usages of *gi*. Thus, for instance, in the previous sentence he translates *gui*$_1$ as a pronoun: '*Voici l'ordre de V.K.A.*' (IC V: 74).

Equational sentences in Old Khmer are used for clarifying identification too. One of such cases is to be found in the previous example.

Clauses of consequence

Both in Old Khmer and Modern Khmer the copular of identity is found in clauses of consequence. However, in Old Khmer *gi* is used in combination with *pi*, the latter has a variety of functions: it introduces clauses of purpose, reason, time, is used as a deontic modal of obligation, and has some other functions. Consider the examples with clauses of consequence with the conjunctions *kɯɯ* (Modern Khmer) and *gi pi* (Old Khmer):

(14) *nàw pèel dael yəəŋ baək kaa daan ciik prʊlaay daoy phdʊl ʊŋkʊʊ pĕəlĕəkam ciə ʊmnaoy rɔbʊh sʊmdac hun saen niəyɔk rŏəthaʔmɔntrɛy tii pii nuh kɯɯ miən prʊciəpɔlrŏət rɔhout dʊl tàw* 1000 *nĕək mɔɔk pii khum kʊh khael prɛɛk ʊmbəl khpɔɔp nɯŋ tɯk vɯl baan smak cət ciik prʊlaay làək tumnup nɯŋ phləw lɔm*
'When we started digging canals giving rice as reward for labour, which was provided by His Excellency co-Prime Minister Hun Sen, (*kɯɯ*) nearly 1,000 citizens came from khums Kohkhael, Prekambil, Khpop and Tikvil to willingly participate in digging canals, building dams and roads.' (Evening News)

(15) *pi* [a]*yat kvan cau ley syaṅ ta tāc santāna dau phoṅ gi pi vraḥ pāda kamrateṅ añ çrīudayādityavarmmadeva oy prasāda bhūmi sruk stuk rmmāṅ nu khñuṃ upāya noḥ phoṅ ta dhūli jeṅ vraḥ kamrateṅ añ çrījayendrapandita leṅ jā nai vraḥ āy stuk ransi*
'Since (=*pi*) (they) did not have children or grandchildren and their whole clan disappeared, (*gi pi*) His Highness Sri Udayadityavarmadeva granted by His grace the land Stuk Rmang, as well as the slaves and all livelihood of this land to His Highness V.K.A. Sri Jayendrapandita to consolidate it with the land Stuk Ransi' (K 219)

It appears that *gi pi* in old Khmer and *kɯɯ* in Modern Khmer, as conjunctions of consequence constitute a paradigm case of a grammaticalization cline, the starting point of which can be found in the following example:

(16) *(...) mel man āçrama ta dai ti [a]val ta çūnya [a]yat _ _ _ ley gi hetu pi thvāy saṃmvat svaṃ vraḥ ālakṣana pi cār*
'(...) saw that all other temples are deserted and do not have _ _ _ so (= *gi hetu pi*) he asked to issue an edict to be carved' (K 829)

This example suggests that sentences with clauses of consequence with the conjunction *gi pi* originated from sentences with a more complex structure, where *gi* was used as a copula of identity and *pi* as a conjunction of purpose. The original complex sentence had the structure P_1 *gi* "reason" (= *hetu*) *pi* P_2, where P_1 represented the state of affairs causing the state of affairs P_2. In this original sentence the copula of identity put into identity relation two nominations: the description of situation P_1, and its taxonomic nomination "reason" (=*hetu*), so that the semantic structure was: P_1 that is (=*gi*) the reason for (=*pi*) P_2. The next stage of grammaticalization was induced by the omission of *hetu* 'reason' - the second component of the equation. As a result, restructuring and reassignment of meaning followed: *gi* and *pi*, structurally unmotivated, lost their original meaning becoming one complex conjunction *gi pi*, where both elements were asemantic. The final stage was the loss of one of the asemantic components, which happened to be *pi*.

Sentences with rheme shift

At first sight, the use of the copula of identity as a marker of rheme shift might seem puzzling. In fact, this usage is well motivated, since the semantic representation of sentences with rheme shift *is* an equational sentence used for identification: *It is Ivan who loves her* = "The person who loves her is Ivan". The first component of this equational sentence contains a relative proposition. So it should not be surprising that in Old Khmer both the copula of identity *gi* and relativiser *ta* are found on the surface level in different constructional types of sentences with rheme shift.

In Old Khmer inscriptions sentences with rheme shift are very common. Their frequent usage reflects the specific character of Old Khmer discourse: the lapidary character of the inscriptions imposed rigid requirements for unambiguity of the communicative structure. The following three types of sentences with rheme shift are found:

a) Sentences which do not undergo syntactic restructuring, where the constituents remain in-situ, i.e. in exactly the same syntactic position that they occupy in pragmatically neutral sentences, and *gi* or *gi ta* are placed before the verb phrase. Sentences of this type allow rheme shift on NP_{SUBJ} and on the assertive status of head predication.

b) Sentences with *gi*-movement, where *gi* is placed with the NP on which the rheme is shifted. They allow rheme shift on NP_{SUBJ} and NP_{OBJ}, the latter can have different syntactic roles.

c) Cleft sentences. They have the structure *gi* NP + Relative clause.

Morphosyntactic structure of sentences with rheme shift can be different, but they all have a formal marker - the copula of identity *gi*.

Sentences with rheme shift without syntactic restructuring. As discussed above, the appearance of the copula of identity in sentences with rheme shift is well motivated, since their semantic structure is an equational sentence with the first component containing a relative clause. What *is* surprising, though, sentences with rheme shift in Old Khmer may undergo no syntactic restructuring whatsoever, *gi* being the only marker of rheme shift placed with the head predicate. Even more surprising is the fact that along with *gi* the relativiser *ta* can also be used. Does that mean that *gi* and *ta* operate as fully grammaticalised rheme shift markers on the syntactic level or, rather, they are markers of equation and relativisation of underlying semantic structure? The latter is more likely since the restructuring yielding "b" and "c" types looks like targeting to map the semantic structure onto the syntactic level.

Sentences with rheme shift on subject NP are most common within this type. Consider the following example:

(17) *aṃmoy mratāñ kumārasvāmi ai ta vrah kamratāṅ añ tilakeçvara tāñ gui ta yajamāna ta vraḥ*
'The gifts of Mratanh Kumarasvami to V.K.A. Tilakeshvara: (it is) Tanh, (who) (*gui ta*) sacrificed to God' (K 664)

With sentences of this type, where *gi* and *ta* are the only markers, discourse structure can provide an explanation for the rheme shift. Thus, according to Coedès, in this inscription Tanh is a subordinate of Mratanh Kumarasvami, the donator, and he was assigned the task of presenting the gifts to God on behalf of his master (IC V: 69). If this interpretation is correct, the rheme shift is highly motivated here.

Consider the following example from an inscription where pedigree in maternal line of a prominent family is described:

(18) *aṃvi kāla vraḥ pāda stac dau parameçvara gi ta nām [a]ji yeṅ mok aṃvi bhavapura pañket kvan pvan [a]nak si piy [a]nak kantaiy*
'After His Majesty the King was gone to Parameshvara, (it was him who) (*gi ta*) brought our ancestors from Bhavapura, (here he) begot children: four males and three females' (K 956)

Consider other examples where *gi* or *gi ta* placed with the head predicate are the only markers of rheme shift:

(19) *sthāna jā āçrama vraḥ vleṅ kalpaṇā ta smiṅ candrapura rlām purohita vraḥ neḥ gi ta paripāla āçrama*
'The place which is (=*jā*) the sanctuary of the Sacred Fire, is the foundation of the Chandrapura Rlam servicer, (it is) purohita of this God, (who) (*gi ta*) guards this sanctuary' (K 691)

(20) *kulapati āçrama naiṛtīya gi pujā vraḥ kamrateṅ añ puṇḍarikākṣa*
'(It is) kulapati of the South-West temple, (who) (*gi*) will perform puja to God Pundarikaksha' (K 56)

The motivation for rheme shift in these examples is to be found in the discourse structure. They are both concluding sentences of discourses of one particular kind, where rheme shift in the concluding sentence is a canonical "figure of discourse". The first inscription informs that a Lonh Dan founded a sanctuary, gave a certain number of slaves for the period of waning moon and a certain number of slaves for the period of waxing moon, the slaves were to work the rice fields and supply a certain amount of rice per day to be sacrificed, this foundation was the sanctuary of the Sacred Fire, the tributer was the priest of Candrapura Rlam and, *lastly*, purohita of this God was the guardian. The second example is from an inscription which describes the things to be done at a certain "ceremony of tomorrow", it lists the amount of deliveries for different ceremonial events on that day and allocates responsibilities of the people involved, *lastly*, it mentions that kulapati of the South-West temple performs the puja ceremony. Discourse structure of such inscriptions includes a closed set of points relevant to the situation in question $\{a_1, a_2, \ldots a_n\}$ (for example: a_1 - type of foundation, a_2 - donations, a_n - the guardian). The speaker says something about each of them, point after point, until he comes to the last point - a_n. Since the last element from a closed set cannot be "new information", rheme shift in the last sentence is required.

Consider an example from Modern Khmer, where rheme shift has a similar motivation (the activity of carpentry is described):

(21) *kaa pɯt chaarat krŏən tae cuəy heek hɒɒk ləək kdaa hoc aoy kèe nɯŋ cuəy aa khlah khlah ponnɒh ae rɯəŋ vŏəh vɛɛŋ dɒmrəm dɒmrɒɒy aoy trəw kbuən khnaat kɯɯ srac ləə cah cah tĕəŋ buən nĕək nuh*
'To tell the truth, Charat only helped to lift and carry the planks for the others and gave a hand to saw only a few, as for measuring and marking according to the pattern (he) (*kɯɯ*) relied on the four old men' (Kong Ponchien)

There are contexts where the technique of marking rheme shift was absolutely indispensable, they are situations where major distributions and redistributions of land occurred. Interestingly, we can even find examples where in sentences with rheme shift referring to such critical acts other morphosyntactic tools are used as a "support group" -- this is done to attach more importance to the described state of affairs. The following example is taken from King Udayadityavarman's edict by which he gave the land Stuk Rmang to Jayendrapandita and his family in maternal line, the former owners of this land having no heirs.

(22) *leṅ santāna dhūli jeṅ vrah kamrateṅ añ çrijayendrapandita ti matṛpakṣa gi nu ka ta mān bhūmyupāya noḥ phoṅ pradvan*
'Let the family of His Highness V.K.A. Sri Jayendrapandita in the maternal line (*gi nu ka ta*) own this land and all sustenance henceforth' (K 219)

Here we find *gi ta* rheme shift markers "reinforced" by the enunciation particle *nu* and by *ka*. The functions of *ka* are still to be studied in detail, but it is clear that they considerably overlap with those of *ta* (Long Seam 2000). In the inscriptions there are many examples of *ka* used in combination with *gi* (*ka gi* [K 502]) and the relativiser *tel* (*ka tel* [K 502]).

Among the sentences where *gi* and *gi ta* are placed with the verb phrase we find those where rheme shift operates *within* the verb phrase, its goal being the affirmative status of the proposition. Consider the example:

(23) *mratāñ çrī sātyayuddha nu mratāñ çrīripumatha kvan mratāñ çrīprathivinarendra gi ta oy sruk neḥ ta vyar*
'Mratanh Sri Satyayuddha and Mratanh Sri Ripumatha, children of Mratanh Sri Prathivinarendra, *did* (=*gi ta*) give those two lands' (K 956)

The narration of the passage this sentence is taken from runs as follows. Mratanh Sri Prathivinarendra, a prominent personality of the period, is given the lands Rdval and Sratac by the king (for conducting the ceremony for liberation from Java). He gives the lands to his two wives, who were sisters. Then the two sisters' brother respectfully addresses the king asking to grant these lands anew to himself and his sisters, lest the family of Mratanh Sri Prathivinarendra should take them. So, Mratanh Sri Satyayuddha and Mratanh Sri Ripumatha, children of Mratanh Sri Prathivinarendra, *do* give these two lands. As Coedès points out, it is not clear from the text why this request on the part of the sisters' brother should be made since Prathivinarendra had already given those lands to the two sisters, so he presumes that after Mratanh Sri Prathivinarendra's death his family laid a claim to these lands (IC VII: 133). Actually, the rheme shift on the affirmative status of the proposition in this sentence clearly and formally indicates that that was exactly the case. Due to this grammatical device additional senses are generated, i.e. that the family of the late Mratanh Sri Prathivinarendra did not want to give the land in question.

Sentences with gi movement. Sentences where *gi* and *gi ta* are placed with the head predicate are highly asymmetrical in respect of their syntactic and semantic structures, the semantic structure being an equational sentence with a relative clause. It is natural, that there should emerge restructured variants targeting to level this asymmetry. The initial stage of restructuring represented in Old Khmer sentences with rheme shift is *gi*-placement with the subject NP on which the rheme is shifted.

Variants of this type are often found in opening sentences with enunciation particle *nu*, they have the structure NP_{TEMP} *gi nu* NP_{SUBJ} VP, where NP_{SUBJ} often refers to the king. Consider the following examples:

(24) *971 çaka aṣṭami roc phālguṇa gi nu vraḥ pāda kaṃmrateṅ añ çrīudayādityavarmmadeva svey vraḥ dharmmarājya*
'In 971 of the era shaka on the eighth day of the waning moon of the month Phalguna (*gi nu*) His Majesty the King Sri Udayadityavarmadeva came to the throne' (K 219)

(25) *925 çaka gi nu vraḥ tmo vvak āmātya driṅ jauv bhūmi lveṅ tvaṅ _ _ _ ta vāp ju nūv kaṃvaṅ ryyap saṅ āçraṃ _ _ _ duk caṃmāṃ kalpanā ta kaṃmrateṅ jagat çrījalāṅgeçvara saṅkkrānta raṅko je 1 pratidina liḥ 1 thlvaṅ _ _ _*
'In 925 of the era shaka (*gi nu*) Vrah Tmo Vvak, councilor, bought the land Lveng Tvang _ _ _ from Vap Ju from Kamvang Riep, founded a temple _ _ _ fixed tribute for the God Sri Jalangeshvara: 1 (measure) je of rice for the new year, 1 (measure) lih thlvang _ _ _ daily' (K 87)

Rheme shift on NP_{SUBJ} can be found in other contexts as well. Thus, the following sentence comes after enumeration of donations and here we find *gi* in two different functions: the first one looks like deictic *gi* and the second one is a marker of rheme shift on NP_{SUBJ}:

(26) *ta gi roḥ _ _ _ rājya vraḥ pāda dhūli jeṅ vraḥ kamrateṅ añ ta stac dau _ _ _ gi añ ta jmaḥ loñ kṣetra duñ sre neḥ _ _ _ dravya añ duñ sre srū 20 vodi 1_ _ _ mas*
'those (=*gi*) things _ _ _ (during) the reign of His Majesty the King, who has gone to _ _ _ (it is) (*gi*) me, named Lonh Ksetra, (who) bought this rice field _ _ _ for 20 (measures) of rice, one vodi _ _ _ gold' (K 259)

Now that *gi*-placement with rhematic NP is allowed, it becomes possible to mark rheme shift on object NPs as well. Consider the following example, where the rheme is shifted on instrument NP and *gi* is placed *after* the marker of instrument *nu*:

(27) *chley neḥ dravya ta roḥh neḥh yeṅ yok ta mratāñ khloñ ti yeṅ thve vraḥ rājakāryya nu gi saṃmall yeṅ aras*
'(They) answered: this property we (indeed) took from Mratanh Khlonh, we paid royal duty with it, on (*gi*) the rest we lived' (K 257N)

In inscription K 956 we find a thrilling example of rheme shift on a comitative NP referring to the sellers of the land:

(28) *man srāc thvāy dravya ta roḥh neḥh jā nai kaṃmrateṅ jagat dep dau duñ bhūmi nu gi ta [a]nak thkval nu [a]nak pralāy slā vraḥ jaṃnvan varṇa vraḥ kṣīra*
'After having given all those valuables to be possession of God, (he) went on to buy land both (*gi*) from (=*ta*) people in Thkval and people in Pralay Sla, guardians of sacred donations of the category of (guardians of) sacred milk' (K 207)

There are several discourse-related factors that account for the rheme shift in this sentence. Firstly, the decision to buy land for God was already mentioned in the previous sentence: "Having a desire to build a place of God, he further made a foundation providing for daily deliveries and *donated the following valuables to be given as the price for the land* (further goes a list of valuable things)". That alone could account for the rheme shift on the NP referring to the sellers of the land in the sentence that follows. The other factor accounting for the rheme shift here is that the sellers are the main topic of the following part of the discourse, so rheme shift could be used here for thematic prominence, to highlight the thematic importance of the referents for the subsequent discourse. Actually, this rheme shift arouses suspicion that there is something important concerning these sellers, which is to be read between the lines. Thus, according to Coedès, it looks like the sellers or their relatives, despite of the fact that they were all given their share of valuables for the land, continued to use it to their benefit (IC III: 17). One cannot agree more with Coedès's opinion, because the remaining 49 lines of the inscription give an exaggeratedly full account of the details of the transaction. The most telling is the vow that the sellers gave: "This land that V.K.A. from Palapriya has bought from us, with the view of building a place of God and providing for the deliveries for daily services for God, which he prescribed to our relatives who are *mulas* to guard henceforth, we shall not commit any fraud or deny our awareness. Should a future king withdraw our relatives from (the domain of) God, we shall instruct our descendants (...) not to withdraw this land from (the domain of) God". All in all, one finds enough justification for the rheme shift in this sentence. This rheme shift makes it possible to guess that there really is some intrigue that should be read between the lines. This example proves that in Old Khmer rheme shift marker could be a powerful tool in organising discourse structure and content.

Cleft sentences. Clefting is another example of reconstructional readjustment driven by the demand for symmetry of syntactic and semantic structures of sentences with rheme shift. Clefting makes it possible to represent the sentence with rheme shift as a complex one in surface structure too. In Old Khmer cleft sentences have the structure *gi* NP + Relative clause. Interestingly, there are cases, where the relativiser is not promoted to initial position but remains with the head predicate. Consider the following example, where the rheme is shifted on NP_{LOC}, the relativiser remains with the head predicate, and there are several *gi*: gi_1 is used in deictic function, gi_2, is the rheme shift marker:

(29) *ge ta sak gi ge ta caṃlāk gi lāṅ vraḥ avicīnaraka mahāraurava raurava kumbhipāka vaitaraṇi kālasūtra taptaka druma vāluka açītimukha gi nā ge tel kaṃvoṅ doṅ ge pitāmātā ge*
'Those who commit pillage here (=gi_1), those who engrave here (=gi_1) and blot out, the hells avicīnaraka, mahāraurava, raurava, kumbhipāka, vaitaraṇi, kālasūtra, taptaka, druma, vāluka, açī timukha, that is where they stay (*lit.* 'gi_2 there they that (=*tel*) stay') with their fathers and mothers' (K 728)

In conclusion, consider an example of a cleft sentence which might suggest that the technique of clefting in Old Khmer was still in the process of perfecting.

(30) *gui man gui noḥ sre tel oy ta vraḥ poñ rudrabhava*
'(It is) (*gui man gui*) this field, that (=*tel*) (I) gave to God Ponh Rudrabhava' (K 79)

This example is rather controversial due to the fact that both *gui* and *man* have several functions. Thus, Jenner (1992) distinguishes six different functions of *man*, one of them is that of a relative pronoun. Actually, in Old Khmer we find the following relativisers and their combinations: *ta*, *ka*, *tel*, *ti*, *man*, *ta ti* (K 598B), *ta tel* (K 726), *ka tel* (K 502), *ta man* (K 49), *man ti* (K 702B), *ti man* (K 388B). So, it is quite acceptable to assume that the previous sentence contains the configuration of relativisers *tel* and *man*. A significant fact, though, is that this configuration has been subjected to splitting: *man* is placed with the object NP *noḥ sre* 'this rice field', *tel* is placed in the relative clause. The more peculiar fact is that both the relativiser *man* and the NP *noḥ sre* 'this rice field' have the rheme shift marker *gui*. This analysis would yield the following construction: "It is (= *gui*) which (= *man*) it is (=*gui*) that rice field, which (= *tel*) I gave to God Ponh Rudrabhava". It looks like either the writer of the inscription does not have a good command of clefting technique or he is trying to be too sophisticated.

Concluding remarks

The resource of the copula of identity in the Khmer language is diverse. This article is just a first attempt to describe its functions. Some of the analysis presented here might prove to be erroneous on further consideration; some will be supplemented by a more thorough analysis. As pointed out in Partee (1998), copula sentences continue to be an inviting domain for cross-linguistic studies in syntax, semantics and information structure. Equational sentences are common in languages; they play an important role in organising discourse structure and content.

The scope of semantic and communicative resources of identity sentences in Khmer is vast. To illustrate this point, I should like to cite one example taken from Judith Jacob's article where Sanskrit borrowings in pre-Angkor Khmer are analysed (Jacob 1977: 152):

(31) *ge ta tve viptya gi saptapitā saptamatā pañcamaharau ravanarakāḥ patañti*
'Those who do any damage will, with seven generations of male and female ancestors, fall into the hell, Raurava, of the five great [crimes]' (K 127)

This sentence is one of the many variants of the warning formula, placed at the end of many inscriptions. It contains only four Khmer words: *ge* 'marker of plurality', *ta* 'relativiser', *tve* 'to do', *gi* 'copula of identity', the rest is borrowed from Sanskrit. As Judith Jacob suggests, the use of Sanskrit for warnings was inspired by the wish to have a good ending (p. 156). This example is also cited in Sak-Humphry 1993: 32 where *gi* is translated as a copula of predication 'is'. In my opinion, both scholars failed to grasp the full sense of this utterance. I suggest that here *gi* is the copula of identity used in an equational

sentence. It puts into identity relation two things - elements of two sets from two different worlds: a type of crime in this world and a type of punishment in another world. If it were not for the copula of identity, this formula would be a sentence with predication (we find many such variants of the warning formula in the inscriptions) and it would not be as meaningful as it is supposed to be. It would be a mere prediction made by a mortal, the utterance would be a judgment bearing a subjective sentiment. Unlike a sentence with predication, the equational sentence is devoid of this subjective sentiment, it is an impartial basic assumption of the way things relate between the worlds. It establishes a fact that has a proverbial flavour, which is not to be challenged by the addressee or some other party; it has a strong inferential character and asserts that one state of affairs is closely tied to another with the implication of a cause-effect relationship (Reintges 1998). Analysing syntactic adaptation of such borrowings, Judith Jacob comes to the conclusion: "The only way in which Sanskrit was modified in these passages was by an increased admixture of Khmer words, producing some particularly difficult sentences" (p. 161). Hopefully, after all that has been said about the Old Khmer copula of identity in this article, one would not regard *gi* in this sentence as "a complexity", one would perceive it as a refinement of the borrowed Sanskrit idiom.

REFERENCES

Arutunova, N. D. 1998. *Yazyk i Mir Cheloveka.* Moscow: Yazyki Russkoi Kultury.

Coedès, G. 1942-1966. *Inscriptions du Cambodge.* BEFEO vol. 2-8. Hanoi (vol. 2), Paris (vol. 3-8).

Jacob, Judith. 1977. "Sanskrit loanwords in pre-Angkor Khmer." *MKS* VI: 151-168.

Jenner, Philip N. 1992. "Observations on Old Khmer *man*." *MKS* XX:1-10.

Long Seam. 1992. *Slovar Drevnekhmerskogo Yazyka*. (po nadpisyam Kambodgi VI - XIV vv.). Unpublished ms. pp. 2194.

Long Seam. 2000. *Dictionnaire du Khmer Ancien.* Phnom Penh: Phnom Penh Printing House.

Partee, Barbara H. 1998. "Copula inversion puzzles in English and Russian." In Katarzyna Dziwirek, Herbert Coats, and Cynthia Vakareliyska eds., *Formal Approaches to Slavic Linguistics: the Seattle Meeting 1998*, pp. 361-395. Ann Arbor: Michigan Slavic Publications.

Reintges, Chris. 1998. "The syntax and semantics of the Coptic cleft constructions." *Electronic Working Papers Series of the Dep. of Linguistics and English Language.* University of Sussex, pp.15.

Sak-Humphry, Chhaney. 1993. "The syntax of nouns and noun phrases in dated pre-Angkorian inscriptions." *MKS* XXII:1-126.

Shmelov, A. D. 2002. *Russki Yazyk i Vneyazykovaya Deistvitelnost.* Moscow: Yazyki Russkoi Kultury.

Received: 24 July 2008

Rozhdestvenka 12
Institute of Oriental studies, Moscow, 107996
RUSSIA
<languages@bk.ru>

Metaphors in the Lavɨəʔ Ləsɔm ʾlɛ poetry

NIRAMON Suwangard
Mahidol University, Thailand

Abstract

A number of metaphoric expressions commonly appear in Lavɨəʔ Ləsɔm ʾlɛ poetry of the Lavɨəʔ of Ban Papae, Mae Sarieng District, Mae Hongsorn Province in Thailand. In this paper, metaphors and metonymies are highlighted to show the people's thought process where an entity as a source model is projected to another (a target model) through their similarities and contiguity under the influence of cultural patterns, society and surroundings. Poetic lines are grouped in basic level categories which provide the predominant attributes for super-ordinate and event categories. These attributes are viewed as source models that facilitate the cognitive process for abstract categories in the target models, showing relationships and blending between salient perceived features. The category 'GIRL' may be mapped from various source models such as beauty, weakness, innocence, a follower, an orchid, a young squirrel, a pheasant chic, a taro-sprout, a calf, a fish, an axe or a pestle; the 'POWER' is associated to a rain cloud ; or a wild cock is mapped onto 'SPIRIT' depending on qualities and characteristics of the entity.

Introduction

"We feel that no metaphor can ever be comprehended or even adequately represented independently of its experiential basis…" (Lakoff & Turner 1989: 19)

Metaphors which are used frequently throughout Ləsɔm ʾlɛ poetry are defined as "The expression of an understanding of one concept in terms of another concept where there is some similarity of correlation between the two."

This paper is taken from the author's doctoral thesis titled "The Analysis of Lavɨəʔ Oral Poetry: the Ləsɔm ʾlɛ". Ləsɔm ʾlɛ oral poetry illustrates a number of figurative language patterns which at the same time present the beauty of the language using similar techniques and devices for composing poetry as those of literate societies. In the poetry, metaphoric expressions commonly occur. The data collected are in Lavɨəʔ dialect of Ban Papae, Mae Sariang district, Mae Hongsorn province, Thailand.

The Lavɨəʔ, also referred to as Lawa, speak a language classified in the Palungic/Waic subgroup of Mon-Khmer. They live along the Southwest Bo Luang plateau in Chiangmai province and the southern part of Mae Hongsorn province between Hod and Mae Sariang districts. Premsrirat (2007) reported that there were less than 10,000 Lavɨəʔ people in 2007. Most of the Lavɨəʔ

still follow the traditional pattern of life. Their economy is based on subsistence agriculture, with rice grown on terraces following a rather sophisticated rotation system. The Lavɨəʔ are animists but many of them have adopted Buddhism or Christianity.

Rattanakul (1979, 1996) suggested that the function of Ləsɔm ʾlɛ poetry in the Lavɨəʔ tradition related to the coming of age of Lavɨəʔ adolescents. Boys and girls must learn to recite the Ləsɔm ʾlɛ and must recite it in funeral ceremonies. The ability to recite the poetry was viewed as a sign of maturity; that one was ready for courtship and marriage and for taking up social responsibilities. In addition, the poetry is sung on other occasions such as wedding ceremonies, new house cerebrations and is also sung for entertainment during fieldwork to lighten the drudgery of their work.

However, the Ləsɔm ʾlɛ style of oral poetry is now gradually dying since it has not been sung for many years (Ratanakul, 1979). This precious art of an obscure ethnic group was apparently valuable more than twenty years ago. Analyzing how metaphors appear in this oral tradition may reveal facts and folklore that are valuable resources for future generations.

Some metaphors and metonymies in the Lavɨəʔ Ləsɔm ʾlɛ poetry

The discussion follows a semantic analysis based both on linguistic metaphors as well as conceptual metaphors.

Metaphor and metonymy are powerful cognitive tools for our conceptualization of abstract categories. The existence of conceptual metaphors which illustrates the move from language to thought is based on the abundant and systematic presence of metaphors in language (Gibbs and Steen, 1997: 1).

The author attempts to group basic level categories which provide the predominant attributes for superordinate and event categories and views them as source model to facilitate the cognitive process for abstract categories. The lines of metaphors and metonymies are highlighted to show thought processes from concrete meanings in the source domain to the abstract one in the target domain, showing relationships between salient perceived features taken from pervasive metaphors in the poems. Metaphorical lines were grouped in semantic domains which relate to various categories (Concrete aspects) and concepts (Abstract aspects) such as girl, man, pleasant life, hard life, big family, power, baby, danger, dying young, feeling, spirit, rain, words or promise, sincerity, etc. Some examples are presented as follows:

1. Category GIRL

The metaphors imply attitudes of the Lavɨəʔ towards girls who are viewed as a flower, a possession, a victim, and as a follower in relation to men. The girls' qualities and conditions: beauty and skillfulness in working are required for marital life. However, sometimes a girl is viewed innocent and needs teaching.

Items 1.1 and 1.2 imply a girl's beauty in manner and in form while item 1.2 illustrates an immature girl.

1.1 kɛh cʰɔm səlit **ləkoʔ** pʰit **ⁿdəom**
really beautiful manner sticky rice steamed soak
'Having a really beautiful manner (like), the soaked sticky rice steamed.'

That is, 'the girl is beautiful with a tender manner.' The ləkoʔ pʰit ⁿdəom 'the soaked sticky rice' is compared to a beautiful manner or tenderness because before cooking (steaming) the sticky rice, it is necessary to soak the rice over night to make the rice grains softer.

1.2 joʔ ˀRɛh ˀRiəŋ ˀRiəŋ **ʔɔʔ** **cɨəŋ** **tʰu**
look overacting girl Exp.-Int-very bamboo beside top of the mountain
'Looking at the girl extremely overacting, (reminding me of) the bamboo beside the top of the mountain.'

The 'ʔɔʔ cɨəŋ tʰu' is compared to the overacting girl because the bamboo at the top of the hillside is burnt repeatedly by strong sunshine on the same side which causes the bamboo canes to become stunted.

That is 'The reciter will not choose the overacting girl to be his wife.'

1.3 **ciək** **ᵐbaɨ** tʰiə: **ʔɨəŋ** cʰok mɨəŋ tai
pick wear flower orchid blossom town Thai
'Pick an orchid blossom from Khon Meuang town to wear behind my ears.'

In 1.3 the flower mentioned refers to 'the girl' and the verb *ciək* 'pick' and ᵐ*baɨ* 'wear' mean 'to select or to love'. Thus all the lines express the intention of the reciter that though there are a lot of girls, the only one he loves is one particular girl.

1.4 meʔ pɛn **kuən** **ˀləi** kəi
you be kid squirrel trail horizontally around the hill

ʔəiɲ pʰit
return drop grains into a hole
'You are a young squirrel who drops grains into the hole.'

The kuən ˀləi 'young squirrel' refers to the girl.' In Lavɨəʔ traditional rice planting, the wife following the husband, who makes holes with a stick with a sharp end, drops a few rice grains into the hole and covers the hole lightly with one foot. This reflects a girl's responsibility.

Items 1.5, 1.6 and 1.7 imply "a girl" as a victim or a weaker.

1.5 ʾɲuət ʾɲuət **kuən tuə** ləlaɨŋ **cʰuəŋ hleʔ**
trembling Exp.-Int. chick pheasant in the middle of drop rain
'The trembling pheasant chick in the midst of rain drops.'

In 1.5, the kuən tuə 'pheasant's chick' refers to 'the girl' and the cʰuəŋ hleʔ 'rain drops' refers to 'the husband's parents and siblings.' The line means that the reciter promises the girl that after marriage, he will take care of her very well in his parents' house.'

1.6 ʾɲuət ʾɲuət **taŋ ʔəic ʾɲuət** taŋ ʔəic
trembling Exp. sprout taro trembling sprout taro with fear
'The taro sprout is trembling with fear.'

1.7 kʰɨəh təo **ləic** loh kloh lələoŋ
give not pig pull out hole plain
'I will not let the pig dig the hole to pull out the taro sprout.'

In 1.6 and 1.7 the taŋ ʔəic 'taro sprout' and the taŋ haɨn 'starchy root sprout' refer to 'the girl' and the ləic 'pig' refers to 'danger', here is another man.

This personification implies the meaning that the reciter will take care of the girl very well if she becomes his wife.

It is noticeable that the word 'ʾɲuət' is repeated, and this renders the feeling of deserving pity and sympathy to the girl as a wife among strangers (the reciter's family and cousins) at the beginning of her marital life.

Items 1.8 and 1.9 reflect a girl's routines

1.8 **ŋgjaɛ ʾmɔɛ** miəŋ kʰəiʔ saop ⁿjɔʔ joʔ
lose axe cut firewood seek not see
'The axe which is used for cutting firewood is lost, (the father) tries to find it but he fails.'

1.9 **ŋgjaɛ ŋgəiʔ** toh hŋɔʔ ləkʰo ⁿjɔʔ pʰɨn
lose pestle pound paddy grasp not can
'The pestle which is used for pounding rice has been lost, (the parents) cannot grasp it.'

In 1.8 and 1.9 the ʾmɔɛ-'axe' and the ᵑgəiʔ-'pestle' refer to the girl. This is the description of the parents' feelings when their daughter gets married and has to leave home to live with her husband. The girl's routines are cutting firewood and pounding rice grains to remove husks. That is, the parents feel sad and disheartened.

In 1.10, the kaʔ 'fish' is compared to the girl and noŋ 'log' is compared to the young man. In 1.11, the ʔoŋ 'wasp' refers to the girl.

The lines describe the reciter's desire to have the girl as his follower.

1.10 sə ᵐbuət ʔmaɛ ʔeʔ **kaʔ** **huət** **noŋ**
follow with I/me fish follow log
'(You) follow me (as) a fish follow a log.'

1.11 sə ᵐbjoŋ ʔmaɛ ʔeʔ **ʔoŋ** **ʔəiɲ** **liək**
accompany with I/me wasp return hive
'(You) accompany me (as) the wasps return in group to their hive.'

1.12 mah ʔlaŋ kuən məok **ʔɛh** ʔlaŋ **taŋ**
be still kid cow take still saddle
'(The girl) is still a calf who carries a saddle on its back.'
The ʔɛh 'take' means to be trained and the taŋ 'saddle' means teaching.

That is, the girl still needs teaching.

Here, the girl is viewed as a follower who is still innocent and needs teaching and training

The kuən ʔləi 'young squirrel' refers to 'the girl.' In Lavɨəʔ traditional rice planting, the wife follows the husband, who makes holes with a stick with a sharp end, drops a few paddy grains into the hole and covers the hole lightly with one foot.

2. Category MAN

Also metaphors show male social status and feelings such as hardship, power and loneliness. These characteristics are expressed in metaphors and appear in rituals and customs as can be seen in the following examples.

In 1.14 and 1.15, the reciter expresses his patience in searching for the girl. This expression is expressed when referring to the marriage ceremony while the bridegroom's procession walks a long way on a hilly trail up and down for possibly a whole day to the bride's house, or when the reciter goes outside his village to a faraway village to court the girl.

1.14 həok ᵐblacŋ Rɔŋ Rɔŋ **ᵐbjɔŋ** **klɔh** **tiəŋ**
go up mountain steep Exp.-very horse hoof single
'The single-hoofed horse goes up the steep mountain.'

The 'ᵐbjɔŋ klɔh tiəŋ-single hoofed horse' is compared to the reciter.

That is, 'I (the reciter) go up steep mountains to find you (the girl).'

1.15 ləih ⁿjɨ: ʔŋiəŋ ʔŋiəŋ **saŋ** **ʔŋiəŋ** **teʔ**
go down mountain short Exp.-very elephant short hand

‘The elephant with the very short trunk goes down the mountain.’

The ʾŋiəŋ ʾŋiəŋ saŋ ʾŋiəŋ teʔ ‘elephant with the very short trunk’ is compared to the reciter.

That is, ‘I go down the mountain to find you with difficulty and hard effort.’

This implies that the reciter has tried to find the girl for a long time while going on a difficult journey.

1.16	paŋ	ʔo	mah	**kʰəiʔ**	ʾɲok	meʔ	**miəŋ**
	though	not	be	firewood	appropriate	you	cut

‘Though I am improper firewood for you to cut.’

The word ‘kʰəiʔ’ refers to ‘the reciter’ and ‘miəŋ’ means ‘to speak to’. The reciter says modestly that he is not worthy for the girl to speak to.

3. Category POWER

The Lavɨəʔ are farmers, thus clouds, especially rain clouds, can impact their lives and this results in thinking of powerful people as clouds.

1.17	ŋoi	ᶯgɨm	kam	tiəŋ	ləlaŋ	**cʰut wu**
	stay still	under	words	assigned	among	clouds

‘Keep still under the order of the clouds.’

The cʰut wu ‘cloud’ refers to the husband and his parents whose power is compared to the height of the clouds.

That is, (the girl) must be under the control of her husband and his parents.

4. Category LONELINESS

Loneliness is compared to human and animal behaviors.

1.18	**lə puən**	nɨŋ	ᶯgloʔ		ᶮjɔʔ	kʷɔɛ	teʔ
	small frog	at	small muddy hole		not	have	hands

‘(I am) a small frog in a muddy hole, who has no one to help him’.

The reciter compares himself to a small frog which lives alone without (anyone) to care for him.

Small frogs are left to grow alone without help from parents just as the reciter has no wife to help him.

5. Category DANGER

A wild pig sometimes destroys plants in the field, thus it is seen as a danger, a boar, here.

1.19	liək	**ləic**	ʾma	mɔk	ŋgoh	hmɔŋ	klaɨk
	enter	pig	field	cut	hit to make a loud noise	hear	a tool made of bamboo

'If a pig enters the cultivated field, (I) hit the 'klaɨk' loudly.'

The ləic 'pig' refers to 'stranger or danger'.

That is the reciter will protect the girl from danger. The 'klaɨk' is a tool made from a big hollow bamboo cane. When hitting the 'klaɨk', a loud noise disperses over the village to make it known to the villagers that a serious accident has occurred which means someone needs help from every villager.

Many things and natural phenomena are construed as a baby, a spirit and concepts and many other categories in metaphorical lines.

6. Category BABY

A baby for parents is tenderly compared to a small plump frog that needs careful bringing up.

1.20	lɔn	me	sə ʔiək	(kuən) ŋak	**ləpuən**
	if	you	hold against the waist	little child	small plump frog

'If you hold the little child, the small plump frog (The baby will sleep well.).'

The ləpuən 'small plump frog' is compared to a lovely small child.

The comparison illustrates an image of a small plump child which is carried on his/her parent's hip.

Here, the reciter encourages the wife by telling her that he will help her take care of the children.

7. Category SPIRIT

A wild cock, a wild bird, a prawn and a porcupine represent the suitable behavior that a dead spirit should follow. In other words, the spirit must go away from the village and must not return to the family.

1.21	koh	pəo	**ʔɛ**	pʰjeʔ	cap	nɨŋ	(ʔɔʔ) pjɔh
	get up	fly	cock	wild	stop	at	dead bamboo clump

'The flying wild cock perches on the dead bamboo clump.'

The 'ʔɛ 'cock' is compared to the dead spirit.

For the Lavɨəʔ, life after death is seen separately from living people. The dead spirits have their own place to live deep under the earth. Influenced by this belief, the Lavɨəʔ bury the corpse with his/her property so that the spirit can use it in the other world.

Here the reciter (shaman) orders the spirit (wild cock) to stop and perch at a fallen dead bamboo clump (and not to go into the village).

8. Category RAIN

Rain brings prosperity to the Lavɨəʔ and this can be seen in the comparison.

1.22 kʰɨəh cap juh hɔic **kʰjaɛ** mɨəŋ luəŋ
give at the right time do arrive gold capital (from the)
'(I wish) the gold from the capital would arrive at the right time.'

The kʰjaɛ 'gold' refers to 'the rain.'

That is, the reciter wishes that the guests will have rain at the right time of planting rice.'

9. Category PROMISE (WORDS)

Some metaphors in Ləsɔm ʾlɛ poetry imply the exact meanings of specific sayings under the characteristics of the entities and the conditions and the use of those in everyday life as follows:

1.23 meʔ paŋ ⁿjaim klaɛh **taih** ʔit **lɔn**
You though want feeling of hesitation mushroom already rotten
'Though you want to change, the mushroom is rotten and cannot become fresh.'

The reciter compares the agreement of marriage to the mushroom which is unchangeably rotten. That is, whatever happens, the girl must marry him.

1.24 paŋ ⁿjaim **ləbe** səme hoic **pleiʔ**
though want return into the same state (of the branch) Indian gooseberry already fruit
'Indian gooseberry branches which are already bent down by a heavy load of fruit cannot fling themself up again.'

This physical character is compared to the agreement of marriage that cannot be changed. In fact, if the girl accepts presents from the boy, this means she is ready to be his future wife but she can change this agreement by returning with two double gifts. The ləbe səme 'the branch of heavy fruits' refers to the agreement.

This means that 'though you desire to go free, it is impossible, just as branches laden with lots of gooseberries cannot spring up again.

1.25 som **ᵐblɨəh** **kɔɛ** **ʾməo** **həo** kiəŋ kiəŋ
Will let cotton roll go about-Exp.
'Let the cotton thread unroll endlessly.'

The verb ᵐblɨəh 'unroll' means to speak and the kɔɛ məo 'cotton roll' refers to words. That is (the girl) talks incessantly without communicating anything important. The meaning is that the girl cannot avoid marriage with the reciter.

10. Concept PLEASANT LIFE

In the Lavɨəʔ view, rich people are ones who have lots of cattle. A cow, a bull, a buffalo or a horse is expensive to buy. Buffalos are used for farming and as sacrifices to the spirits. Oxen and bulls help carry heavy things when they go trading.

1.26 **moʔ** **kjak** ɲɨəʔ ʾjiək kʰjoiɲ cʰuəŋ ʾRɔŋ
cord buffalo house brother every foot post
'My house has buffalo-cords for every post.'

The 'ʾjiək' means 'the reciter'. That is, the reciter has a lot of buffaloes.

In item 1.26, the reciter promises the girl a pleasant life with a lot of buffalos to work for her. However, this is only hyperbole because in real life, it's the girl who does house chores.

11. Concept HARD LIFE

The Lavɨəʔ metaphors reflect that girls and women work hard for their families. The following shows the reciter's attempt to convince the girl of a convenient and comfortable life after marriage.

1.27 juh təo nɨŋ meʔ **hlaʔ** **lə ᵑgɨə:**
make not at, to you leaf a tree which has large leaves
'(I) do not allow you to be a leaf for wrapping things.'

1.28 juh **hlaʔ** **lə ᵑgɨə:** kʰɨə: toʔ taɛh
make leaf a tree which has large leaves wrap meat give away
'(I) do not allow you to be a leaf which is used to wrap meat offered to the guests in a ritual ceremony.'

Big leaves of trees and bananas are used to wrap things for other people at special occasions. For example, when visiting the sick, the guest brings areca nuts and betel leaves and salted fish wrapped in a leaf. This is to bless the sick. If someone can hunt a deer, the meat wrapped in leaves is given away to neighbors. Besides, wild leaves are used much in rituals. It can be considered by the Lavɨəʔ that leaves carry a heavy burden as in lines 1.27 and 1.28, the hlaʔ lə ŋgɨə: 'leaf for wrapping things' refers to working hard. That is, the reciter promises the girl that she will not work hard if she gets married to him.

12. Concept BIG FAMILY

As farming is the main activity for living, a great deal of labor is required to help with planting and havesting. In the Lavɨəʔ society, not only people in their own family help by working in the field but also their relatives and neighbors make a contribution to the success of farming.

1.29 tɔk nɨŋ wɨəh paŋ ʾmaɛ kʰoʔ **cʰɨəʔ**
desire at, to large tree clump with tree a kind of big tree
'(I) desire a big tree clump of 'cʰɨəʔ.'

In 1.29 the wɨəh paŋ ʾmaɛ kʰoʔ cʰɨəʔ 'a big tree clump' refers to a big family with a lot of cousins. So the meaning is that the reciter desires to get married to the girl and consequently, he will have many more cousins and relatives.

13. Concept DYING YOUNG

In the Lavɨəʔ view, some plants that have a short life and easily rot, or natural phenomena like hail and rain, or sprouts which are cut before growing to maturity, are compared to human life especially one who dies at a young age.

1.30 meʔ paŋ jɨə lɔic tɔk **ləmaɛ**
you though easy die, finish like starchy tubers of a kind of wild yams
'Though you died easily like sweet potatoes.'

1.31 paŋ pəʔ yɨəŋ ŋgjaɛ tɔk **pyɛ** təom
though you easy lost like **hail** come
'Though you got lost easily like hail which has come and gone away quickly.'

14. Concept FEELINGS

Feelings are compared to natural things, animal behavior and natural phenomena and imply feelings and emotions as in the following examples:

The reciter suffers painful feelings from the girl's stiff words when she responds to his words. However, the comparison is rather exaggerated in order to ask for sympathy from the girl. In certain occasions, happy emotions are expressed using hypoboles with natural and familiar events.

15. Concept PAINFUL FEELING

1.32 **kʰɨ:** chuəŋ **ᵐbɔ** toh loh sə kaɨ
rub the end of the lever supporter where people step on to raise the pestle when pounding grains big mortar pound tear off awning
'The awning is rubbed and torn off by the lever end when pounding paddy (until it falls down).'

The reciter compares his seriously painful feeling to that of the awning which is repeatedly rubbed by the end of the pestle of a big wooden mortar during the paddy pounding activity.

Lavɨəʔ people have big wooden mortars with levels of pestles set beside their houses where there is an elongated awning from the roof to protect the area from rain and sunshine.

16. Concept HAPPINESS

1.33 maic (lə) pʰaom **kuən** **ʾlaɨ** nɨŋ wɨəh ŋgloʔ
good mind kid frog at large marsh.'
'Good mind (like) a tadpole in a big marsh.'

Tadpoles can be seen only in the rainy season when the marshes and hollows are full of water. The manner of tadpoles when they swim actively reminds the viewers of cheerfulness and happiness.

The line means that the reciter feels happy like a tadpole in a marsh.

We can see from the examples above that the entities in the Lavɨəʔ surroundings extremely influence the Lavɨəʔ way of thought. However, the same entities can be referred to as different concrete and abstract categories depending on their qualities and characteristics in the view of the Lavɨəʔ. Metaphors and metonymies mentioned above express Lavɨəʔ thought process which are relevant to their society and surroundings.

A 'girl' is associated with 'beauty, weakness, innocence and being a follower' and so is compared to an orchid, a young squirrel, a pheasant chick, a taro-sprout, a calf, and a fish. She is an orchid for her beauty. She is a squirrel when working in a rice field, and an axe and a pestle when she does her chores at home. She is a pheasant chick or a taro-sprout when she is in danger. She is

a calf because she is innocent and needs teaching or training. She is a fish when a man takes her on as a follower. If a 'girl' has good manners, she is steamed sticky rice, but becomes a bamboo burnt by the sun on one side, when her behavior is not allowed for a girl of her age. This attitude supports the inferior status of a woman in a Lavɨəʔ family and society in that it assigns her to working hard, being responsible for all chores such as providing firewood, cooking, pounding rice, weaving clothes for everyone in her family and planting. On the contrary, a 'man' is supposed to be a lame horse or elephant which travels a long way and with hardship to court a girl expecting to marry her. A 'man' is compared to a cloud, which is high in the sky, because he has the most power in his family and plays a dominant social role. He is a breadwinner, he leaves the village to earn money by trading his labor for income after the harvesting season. 'Loneliness' relates to a small frog who naturally feeds himself alone. 'Danger' is compared to a pig who destroys crops. A 'baby' is a plump frog showing deep affection to a beloved child who is always lovely in his parents' thought. A 'dead spirit' is compared to a wild cock which should fly away from living people and live deep under the earth. This influences the way the Lavɨəʔ treat the deceased person, in terms of what property will be buried.

'Rain' is valuable for the Lavɨəʔ especially during the cultivation season and so is compared to 'gold'. A 'promise' is compared to a rotten mushroom or a bent Indian gooseberry branch loaded with fruit which cannot return to its original condition. On the other hand, incessant talking is compared to enless unrolling cotton thread. A 'pleasant life' is linked to wealth with lots of cattle, and comfort means a woman has a servant to work for her while 'hard life' is compared to the big tree leaves which are usually used to wrap things. A 'big family' is compared to a big tree clump. 'Dying young' is like being a starchy tuber or hail, both of which have very short lives. Painful feeling of an awning which is rubbed repeatedly and so violently that it is torn off from the roof is personified to become the 'painful feeling' of a person. But 'happiness' is like the cheerful manner of a tadpole in a marsh or a hollow which is full of water.

It is obvious that Lavɨəʔ poets have the capacities to use literary techniques, metaphors in general such as imagery, figures of speech, personification and hyperbols. The metaphoric lines reflect the Lavɨəʔ's knowledge of natural sciences about animals and plants, traditional rites and rituals. It is extremely worthwhile for young Lavɨəʔ to raise awareness of the value and wisdom implied, to conserve the poems and to develop their high self-esteem and identity.

REFERENCES

Diffloth, Gerard. 1972. "Notes on expressive meaning." *Paper from the 8th Regional Meeting.* Chicago Linguistic Society, Chicago.

Dirven, Rene and Ralf Porings (eds.). 2003. *Metaphor and Metonymy in Comparison and Contrast.* Berlin, New York: Mouton de Gruyter.

Freeman, Magaret H. 2003. "Poetry and the scope of metaphor: Towards a cognitive theory of literature." In Antonio Barcelona, ed., *Metaphor and Metonymy at the Crossroads*: A cognitive Perspective. Berlin: Mouton de Gruyter.

Gibbs, JR, Raymond W. and Gerald J. Steen. 1997. *Metaphor in Cognitive Linguistics.* Selected paper. Amsterdam: John Benjamins Publishing Company.

Goddard, C. 1998. *Semantic Analysis***.** Oxford: Oxford University Press.

Grady, Joseph E. 1997. "A typology of motivation for conceptual metaphor correlation vs. resemblance." In Raymond W. Gibbs, Jr., ed., *Metaphor in Cognitive Linguistics***.** Amsterdam: John Benjamins Publishing Company.

Hutchinson, E.W. 1935. "The Lawa in Northern Siam." *JSS* 27(2):153-182.

Katz, A.W., Jr. Caciari, R.W. Gibbs, Jr., & M. Turner. 1998. *Figurative Thought and Language***.** New York: Oxford University Press.

Katz, J.J. 1972. *Semantic Theory***.** New York: Harper and Row.

Komonkitiskun, Jiranan. (1985). "Some general characteristics of Lawa grammar." (La-up Dialect), M.A. thesis, Institute of Language and Culture for Rural Development, Mahidol University

Kunstadter, Peter. 1964. *Research on the Lua and Skaw Karen Hill Peoples of Northern Thai* Duplicated.

Kunstadter, Peter. 1967b. *The Lua' (Lawa) and Skaw Karen of Maehongsorn Province, Northwestern Thailand*, In Kunstadter (ed.) 1967: 639-676.

Lakoff, George. 1993. "The contemporary theory of metaphor." In Andrew Ortony, ed., *Metaphor and Thought*. (2nd edition) Cambridge: Cambridge University Press.

Lakoff, George and Mark Tumer. 1989. *More than Cool Reason*. Vol. l: Theoretical Prereguisites. Standford University Press.

Lyons, J. 1995. *Semantics.* 2 vols. Cambridge: Cambridge University Press.

Premsrirat, Suwilai. 2007. "Endangered languages of Thailand." In Morita, ed., *IJSL* vol. 186, New York: Mouton de Gruyter.

Ratanakul, Suriya. 1979. **"**The Lawa Ləsɔm ʔlɛ Poetry**"**. *Paper Presented to the International Symposium on Austro-Asiatic Languages*, Helsingor, Denmark, October 24-26, 1979, Published in Journal of the Siam Society, 1985.

Ratanakul, Suriya. 1985a. "The phonology of Lawa." Published in *Southeast Asian Linguistic Studies* presented to Andre′ Haudricourt, Bangkok, (1985).

Ratanakul, Suriya. 1996. "Lawa Ləsɔm ʔlɛ Poetry Revisited." *MKS* XXVI: 387-410.

Quinn, N & C. Strauss. 1997. *A Cognitive Theory of Cultural Meaning*. Cambridge: Cambridge University Press.

Reinhart, T. 1976. "On Understanding Poetic metaphor." *Poetics* 5:383-402.
Renkema, Jan. 1993. *Discourse Studies*. Amsterdam: John Benjamins Publishing Company.
Smalley, William A. 1994. *Linguistic Diversity and National Unity in Thailand***.** Chicago: The University of Chicago Press.
Steen, G. 1994. *Understanding Metaphor in Literature: An Empirical Approach.* London: Longman.
Tribal Research Institute, 1991. **"**Hill Tribes: Ways of Life and Development.**"** *Paper Presented in the Exhibition on Golden Anniversary of the Department of Public Welfare at Alliance Française*, Bangkok. August 15-12, 1991. Published in a book.
Ullmann, Stephan. 1962. *Semantics. An Introduction to the Science of Meaning*, Oxford: Basil Blackwell.
Ungerer, Friedrich and Hans-Jorg Schmid. 1996. *An Introduction to Cognitive Linguistics.* London: Longman.

Received: 8 October 2008

Uttaradit Rajabhat University
27 Injaimee Road, Amphoe Muang,
Uttaradit Province, 53000
THAILAND
<niramons@yahoo.com, niramons@hotmail.com>

The origin of the peculiarities of the Vietnamese alphabet*

Translated by Alexis MICHAUD,
LACITO-CNRS, France

André-Georges HAUDRICOURT
LACITO-CNRS, France

Abstract

The missionaries who adapted the Latin alphabet to Vietnamese were Portuguese, Italian and French. The resulting spelling inherited some peculiarities from the spelling systems of Romance languages.

Aspirated consonants H, PH, TH, KH [IPA: /h/, /p^h/, /t^h/, /k^h/] are not found in Romance languages; however, in these languages, the combinations of letters PH, TH and KH are present in words of Greek origin, as equivalents for the Greek initials *phi, theta, khi* (φ, θ, χ), which were aspirated consonants in Greek; and thus these combinations were used to transcribe Vietnamese aspirated stop consonants.

Dorsal stops C, G are only used before the vowels /a/, /o/ and /u/. This is because, in Romance languages, it is the only position where these consonants preserved the obstruent pronunciation they had in Latin; GHE, GHI are used with the phonetic value they have in Italian; the combinations KE and KI resort to the letter K, used in Greek (*kappa*, κ) and in Germanic languages.

Labiovelar stops QU and GU are taken from Italian and Latin spellings.

Among prepalatals, the unvoiced stop CH is taken from the Portuguese and the Spanish, which themselves borrowed this notation from Old French, where it had been created to transcribe a new sound, not found in Latin.

The voiced stop D is used as an approximative notation for a sound not found in Europe, where D is the voiced counterpart of T. In Vietnam, a new letter, Đ, was coined [for a preglottalised alveolar stop: /ɗ/], its horizontal bar pointing to a similarity with the letter T.

The voiced spirant[1] was written as GI, as in Portuguese and French (at the time, J was not yet in use in Europe).

The unvoiced spirant X is borrowed from Portuguese and the Northern dialects of Spanish: in these languages, the S is pronounced at the back of the mouth [IPA: retroflex /ʂ/], as in Vietnamese, whereas elsewhere in Europe S is an anterior sibilant as in French [IPA: alveolar /s/].

NH [for the palatal nasal /ɲ/] is borrowed from the Portuguese; TR is an approximative notation for a sound which is not found in European languages [IPA: /ʈ/].

*Originally published as: L'origine des particularités de l'alphabet vietnamien, *Dân Việt Nam* 3:61-68, 1949.

> Ô, Ê are taken from the Portuguese, which is why the pair Ê, E does not have the same phonetic value as in French. [In French, Ê stands for /ɛ/, E stands for /e/, whereas the opposite is found in Vietnamese *quốc ngữ*, namely Ê for /e/, E for /ɛ/.]
>
> Y is used in Vietnamese *quốc ngữ* in a way similar to what is found in Spanish, where it replaces I in-between vowels or at the end of words. The letter Y comes from the Greek alphabet (*upsilon*, υ).
>
> Ơ and Ư were created to represent Vietnamese vowels that do not have equivalents in Western Europe [IPA /ɤ/, /ɯ/].

Translator's foreword

André-Georges Haudricourt's contribution to Southeast Asian studies is internationally acknowledged, witness the Haudricourt *Festschrift* (Suriya, Thomas and Suwilai 1985). However, many of Haudricourt's works are not yet available to the English-reading public. A volume of the most important papers by André-Georges Haudricourt, translated by an international team of specialists, is currently in preparation. Its aim is to share with the English-speaking academic community Haudricourt's seminal publications, many of which address issues in Southeast Asian languages, linguistics and social anthropology.

The article "The origin of the peculiarities of the Vietnamese alphabet" is not one of Haudricourt's most famous articles, and therefore it will not be included in the projected volume of collected papers. However, to this day, it remains an insightful and vivid account of the origin of the modern Vietnamese script. It traces the peculiarities of this spelling system back to the spelling habits of the Romance languages that were familiar to the authors of this system. The article illustrates Haudricourt's passion for reconstructing the historical origin of complex phenomena, and his skill for adducing evidence from an impressive range of sources.

The article was clearly intended for a broader audience than most of Haudricourt's other publications. Its style is colloquial; technical terms are avoided. This work appeared in the third and last issue of the journal *Dân Việt Nam* (The Vietnamese People) published by the *Ecole Française d'Extrême-Orient* in Vietnam in 1948 and 1949. The original is now difficult to find; moreover, many of its potential present-day readers may not have a command of French, whereas back in 1949, Haudricourt could confidently expect that the general public could read French. The present translation aims to make this document available to anyone interested in Vietnamese, and in writing systems in general.

The Vietnamese publisher obviously had a difficult time composing this article, which uses a wealth of symbols. The present translation corrects some typos: in the original edition, C had been substituted for G on page 64; the diacritic in p‘, t‘, k‘ had been rendered approximately, as pc, tc, kc; etc. Transcriptions in the present-day version of the International Phonetic Alphabet were systematically added at translation. Translator's comments are placed in square brackets or in footnotes.

Many thanks to Michel Ferlus for pointing out this article, to Boyd Michailovsky and Martine Mazaudon for corrections to the translation, and to Jean-Michel Roynard for his help with the figures.

Introduction

Alphabetic writing aims to represent the pronunciation of words. In theory, each letter stands for one sound [one phoneme], a given letter always stands for one and the same sound, and a given sound is always represented by one and the same letter. This ideal situation is not found in the case of the alphabetic writing of languages that have a long history and a literary tradition, because as time goes by, their writing systems remain the same whilst their pronunciation changes. Written texts are passed on from one generation of speakers to the next without changes (or only with minor changes, which lag behind the evolution of pronunciation). Thus the correspondence between written signs and spoken sounds changes as the spoken language evolves. What I refer to as the peculiarities of a given alphabet consists of the lack of one-to-one correspondence between letters and sounds: one and the same letter standing for different sounds, or the same sound being represented by different letters. These peculiarities originate in the historical evolution of pronunciation in the language at issue.

When a language is put into writing for the first time, the writing system devised for this language is generally based on the writing system of another language, and the new writing system inherits peculiarities from the writing system which served as a model. Thus the Vietnamese *quốc ngữ* alphabet was based on the orthography of several Romance languages, and it inherited peculiarities which can be explained by the way the pronunciation of Latin evolved in Europe. Recall that Latin, which in the 3rd century BCE was only spoken in the town of Rome, spread all over Europe at the time of the Roman empire (up until the 5th century CE), only leaving a few non-Latin-speaking areas, e.g. a small region of the Western Pyrénées where the Basque language was preserved. Later, the varieties of Latin spoken in different areas evolved in different ways: the Latin of the region of Florence became the Italian language, the Latin of Paris became the French language, that of the region of Burgos became Spanish. However, throughout the Romance area, Classical Latin continued to be used as the language of religion and science, although its accurate original pronunciation was not known anymore. (To this day, Classical Latin remains the liturgical language of the Roman Catholic church, and scholars doing research in natural history refer to plants and animals by Latin names.) The present-day French, Italian, Spanish and Portuguese languages are called “Latin languages”, “Neo-Latin languages”, or “Romance languages”. The Latin alphabet was adapted to other languages of Europe (Celtic, Germanic and Slavic) by scholars who knew the Latin transcriptions of Greek words, Greek being the language of culture of Antiquity.

In this article, I set out the peculiarities that the Vietnamese alphabet [today known as *quốc ngữ*] inherited from the Romance languages; why it differs from the French alphabet despite the fact that French is a Romance language; and how the Vietnamese alphabet and the French alphabet both differ from the scientific alphabets used by the scientists who study languages.

The order in which the letters of the alphabet are arranged goes back a very long way. It is known that the alphabet was invented in Phoenicia (North-West of Palestine) and spread in Greek-speaking countries; in Eastern Greece it gave birth to the Greek alphabet as we know it; in Western Greece it took on a somewhat different shape and spread into Italy, where it gave birth to the Latin (Roman) alphabet. The present-day order of letters in the Roman alphabet—which we use in dictionaries, for instance—is the same as that of the Greek alphabet and of the Hebrew alphabet (that of the inhabitants of ancient Palestine). It is a purely traditional order, which we are not able to explain or justify. This is quite different from the alphabets of Hindu origin, which are ordered in a scientific way, separating consonants and vowels, and classifying consonants according to their place of articulation, starting out at the back of the mouth and progressing forward up to the lips.

[Consonants are presented first (section 1 of Part I), then vowels (section 2); Part II consists of a short note on tones.]

Part I: Consonants and vowels

The Latin and Greek alphabets were very well adapted to the pronunciation of the Greek and Latin languages. Each letter stood for one given sound.

1. Consonants

Latin had two series of stop consonants (i.e. consonants produced with a full closure at some point in the mouth): a voiced series (i.e. with vibration at the larynx), /b/, /d/, /g/, written B, D, G, and an unvoiced series (without vibration at the larynx): /p/, /t/, /k/, written P, T, C, Q. This system of two series is still found in French. Ancient Greek, on the other hand, had three series of stop consonants: a voiced series (/b/, /d/, /g/, written B, Δ, Γ), an unvoiced unaspirated series (/p/, /t/, /k/, written Π, T, K), and an unvoiced aspirated series (/p^h/, /t^h/, /k^h/, written Φ, Θ, X). Latin literati reproduced this last series by adding an H after an unvoiced consonant, yielding the combinations PH, TH, KH. But as early as the end of Antiquity, the pronunciation of Greek evolved: all the stops became spirants [fricatives], i.e. instead of a full obstruction of the vocal tract during the consonant, there was only a narrowing at one point in the tract. Thus, when the Roman alphabet was employed to transcribe Germanic languages at the beginning of the Middle Ages, the combinations PH, TH, KH were still in use but the words in which they appeared now had spirant consonants and not obstruents, and as a consequence PH, TH, KH were used by scholars with their new phonetic value, which was different from the one they had in Classical Latin: they were

used to transcribe spirants. This is why, to this day, TH is employed for spirants in English, and CH for spirants in German. In Latin, PH (/p^h/) got to be pronounced like F (as /f/), and TH got to be pronounced like T (/t/), hence the French pronunciation: /t/ for TH, /f/ for PH. In Vietnam, the situation is analogous to that found in Greek: TH remains an aspirated stop like that of Ancient Greek; PH is a spirant, as in Modern Greek. In scientific transcription, spirants are to be carefully distinguished from aspirated stops. Spirants are usually represented by means of the Greek letters ϕ, Θ, χ for the unvoiced set (bilabial: /ϕ/, dental: /Θ/, velar: /χ/) and β, δ, γ for the voiced set (bilabial: /β/, dental: /δ/, velar: /γ/). The aspirated stops are ph, th, kh [IPA: /p^h/, /t^h/, /k^h/]; some authors transcribe them as p‘, t‘, k‘, using a small apostrophe with its concave part turned to the right: the *spiritus asper* of Greek writing, created by the Greeks of Alexandria to note aspiration, which existed in their variety of Greek but not in Athens Greek, where the letter H was used for a vowel [written H as a capital letter, η as a small letter]. Indianists use the notations *ph, th, bh, dh*, whereas Sinologists use p‘, t‘, b‘, d‘ instead, to note the same sounds.

Latin had four types of stops: labials B, P (/b/, /p/), formed by joining the lips; apicals (dentals) T, D (/t/, /d/), formed by joining the tongue tip (by its Latin name: the *apex*) with the base of the teeth; dorsals (also called palatals) C, G (/c/, /ɟ/), formed by joining the back of the tongue with the palate. The letter Q represented a velar stop, formed by joining the back of the tongue with the velum; in Latin, this sound was always labialised [IPA /k^w/], i.e. accompanied by lip rounding (as found in the vowel U [IPA: /u/] of Vietnamese spelling). This is why, in the Latin writing system, Q is followed by U. QU [IPA /k^w/] is the unvoiced stop corresponding to the voiced stop GU [IPA: /g^w/]. This opposition is preserved in Rumanian, where Latin AQUA “water” became *apa* whereas Latin LINGUA “tongue” became *limba* [i.e. the earlier opposition of /k^w/ and /g^w/ in Latin remains a voicing opposition in Rumanian, between /p/ and /b/]. In Latin, QUI is a single syllable [IPA: /k^wi/], whereas CUI represents two syllables [IPA: /ku.i/].

The old (Latin) pronunciation of the groups of letters QU and GU is preserved only in Italian, and it is from Italian that this reading was borrowed for use in the Vietnamese script [where QU stands for /k^w/]. In French and in Spanish, labiovelars became simple dorsals (except before A). Thus, the Latin word QUAMQUAM was pronounced CANCAN [IPA: /kɑ̃kɑ̃/, with initial /k/, and not /k^w/] in the schools of Paris in the 16th century.

C, G

The history of C is very complex. To begin with, by comparing the Latin alphabet with the Greek alphabet one observes that the third letter of the Greek alphabet is a voiced dorsal, Γ (small letter: γ), not an unvoiced dorsal as in Latin (C). This is because the Latin alphabet was influenced by a neighbouring people, the Etruscans, who distinguished between aspirated and unaspirated stops [e.g., /t^h/ and /t/] but not between voiced and unvoiced stops [e.g., /d/ and /t/]. (The same holds true of the Chinese of Beijing and Guangdong.) The Etruscans thus used the third letter of the alphabet to note an

unvoiced dorsal. The Latins, on the other hand, did have a distinction between voiced and unvoiced stops; using the same letter for both was not satisfactory. They therefore added a horizontal bar to distinguish a G from a C. C remained in third position in the alphabet, whereas G went into the slot formerly occupied by the Z [small letter: ξ] of the Greek alphabet, which at that time was not employed in Latin. (The two Greek letters Y and Z are found right at the end of the Latin alphabet, because, as we shall see below, they were not borrowed until much later.)

In Latin, C had the same pronunciation in the syllable CE as in the syllable CO. This is still how things stand in the Celtic languages of the British isles (Irish and Welsh), which were the first non-Romance languages to adopt the Latin alphabet. But in Latin, from the 4th century onwards, when the dorsal consonant was followed by a vowel E, I (/e/, /i/), which requires a fronted position of the tongue, it became a prepalatal consonant, during whose articulation the tip and the back of the tongue make simultaneous contact with the front part of the palate. Such consonants have been part of the phonemic inventory of languages of Indochina for a very long time, and they remain stable in these languages; in Europe, on the contrary, these new sounds were difficult to perceive; their perceptual salience was enhanced by shaping the tongue in a channel-like form: if the channel is on the back of the tongue, this yields a non-anterior sibilant stop [in present-day IPA transcription, /tʃ/, a postalveolar affricate], like Italian CI, CE; if the channel is towards the tip of the tongue, this yields an anterior sibilant stop [IPA: /ts/, alveolar affricate]: this is how CI, CE were pronounced in Old French and Old Spanish. This last sound is currently found in Mandarin Chinese and Cantonese Chinese, but not in Indochina.

Meanwhile, the consonants of the Latin syllables TIA, TIO had become anterior sibilant stops [i.e. the affricate /ts/] in Italian as well as in the other Romance languages. The Greek letter Z was borrowed to represent the voiced anterior sibilant stop [i.e. /dz/]. In Italian, Z stands for both the voiced and unvoiced sounds [i.e. /dz/ and /ts/].

K, KH, GH

When the Latin alphabet was adapted to the Germanic languages (in the 7th century), then to the Slavic languages that had these stops, Z and C were employed. One had to borrow the Greek letter K (which in the Latin alphabet had remained at its original place, the one it had in the Greek alphabet) to note the dorsal stop.

This accounts for the Vietnamese spelling, in which one finds CA, CÔ, CU [for /ka/, /ko/, /ku/] but KÊ, KI [for /ke/, /ki/]: in Vietnamese, QU could not be used in the way it is used in French and Spanish [i.e. for /k/] because QU was used to transcribe another Vietnamese sound [IPA /kʷ/]; Vietnamese /k/ could not be transcribed as CH either (the way it is used in Italian), because CH was used for a Vietnamese prepalatal unvoiced stop, as we shall see below. This also precluded the use of CH as a notation for the aspirated stop [IPA: /kʰ/]. This led to the use of KH for the aspirated stop [IPA:

/k^{h}/], before all vowels. The voiced dorsal raised a similar issue: GA, GÔ, GU [for Vietnamese /ɣa/, /ɣo/, /ɣu/], but GHÊ and GHI [for /ge/ and /gi/, present-day /ɣe/ and /ɣi/] and not GÊ and GI; one could not use the combinations GUÊ and GUI, which were used in the case of a labiovelar initial, whereas the Italian notation could be used, hence GHÊ, GHI.

The scientific notation consists in using k for the unvoiced dorsal stop, and g for the sonorant dorsal stop.

CH

In Western Romance languages, a new prepalatal stop emerged and later turned into a non-anterior sibilant stop [IPA: postalveolar affricate, /tʃ/]. In French, it originates in the Latin syllable CA: for instance, *calidum* "hot" became French *chaud*, *caballum* "horse" became French *cheval*. The non-anterior sibilant stop [IPA: postalveolar affricate, /tʃ/] was written as CH, where the function of the H is only to indicate that the C is pronounced neither as in CO, nor as in CE. At the time of Charlemagne (8th century), it was still a prepalatal stop; in Old French (9th century), it had become a non-anterior sibilant stop [IPA: /tʃ/] — a value which is preserved in English. In Portuguese, it originates in initial clusters PL, CL; in Spanish, in the word-medial cluster CT, e.g. Latin *noctem* "night" first changes to *noite*, as preserved in Portuguese, then to *noche*. In Basque the prepalatal stop still exists, e.g. in *éche* "house".

Given this situation in the languages they were familiar with, it was quite natural for the creators of the Vietnamese Roman alphabet to note the Vietnamese prepalatal stop as CH.

The scientific notation of the palatal stop, a rare sound in Europe, is not well established yet [present-day IPA standard: /c/]: it is sometimes noted as k^{ʲ} or t^{ʲ}, where the apostrophe indicates palatalisation, a phenomenon referred to in French as *mouillure*, literally 'wetting', because the horizontal position of the tongue dorsum *wets* the palate, as it were. This notation is borrowed from a Slavic language, Czech; it is familiar to the European public through the famous shoemaker's brand *Bat'a*. English and Chinese phoneticians, objecting to the use of diacritics placed on the side of the letter, add a curl to the bottom of the letter. Indianists note it by C; actually, the prepalatal stop of Sanskrit was only preserved in Indochina, in Sanskrit borrowings into Khmer and Lao: in India, it became a non-anterior sibilant stop [IPA: postalveolar affricate, /tʃ/]. For sibilant stops, the Czech notations are commonly used: C for the anterior sibilant stop [i.e. /ts/], Č for the non-anterior sibilant stop [i.e. /tʃ/] (thus, "Czech" is written *česky*). The sign Č is an abbreviation for CZ, a combination still in use in Polish. Czech does not have a voiced anterior sibilant stop, and Z is used in Czech to note a voiced spirant [i.e. the fricative /z/], as in French.

Old French had such stops [IPA: affricates]; in the 13th century, they weakened into spirants. The anterior sibilant stop became an anterior sibilant spirant, and since then CE, CI have been pronounced in the same way as SE, SI. Meanwhile, the corresponding voiced sound became the voiced

correspondent to /s/, i.e. /z/. The non-anterior sibilant stop [i.e. the affricate /ʧ/] became a non-anterior spirant: in French, CH is not pronounced as the English CH [i.e. /ʧ/] anymore, but as the English SH [i.e. /ʃ/]. A Latin word-initial I followed by a vowel had by that stage become a voiced palatal stop (like G in GE, GI /ge/, /gi/), then a voiced non-anterior sibilant stop [IPA /ʤ/] (word-initial J, originating in Latin I, still has this value in English), and finally, in 13th-century France, a voiced, non-anterior spirant [IPA: /ʒ/].

GI, D, Đ

The notation of the spirant mentioned above (/ʒ/) as J in French only dates back to the 17th century. Earlier on, J was simply the form that the letter I took when at the beginning of a word; thus one wrote jnde (for present-day French *Inde*, "India"), jure (for present-day *ivre*, "drunk"), jean (for the proper noun still written as *Jean* today; English equivalent: "John"), and so on. In Italian, the sequence GI was used in all cases where the sound at issue was a consonant, e.g. *Giovannis* "John". Thus, in the 17th century, when a romanised script was devised for Vietnamese, J did not exist as a distinct letter contrasting with I; one can easily understand why GI was then adopted as a notation for the prepalatal voiced spirant (/ʒ/). When it came to transcribing the voiced prepalatal stop[2], which to the European ear sounded like a palatalised /d/, a simple D was used, without any added diacritic. In Vietnamese there was also a voiced [and preglottalised] alveolar stop [IPA: voiced implosive /ɗ/]; to transcribe it, the creators of the Vietnamese alphabet coined a new letter, Đ, by adding to the letter D a horizontal bar by analogy with the one in the letter T, which is the unvoiced counterpart to Đ.

S, TR

The anterior sibilant stop of Old Spanish, written C and Z, weakened, but did not become confused with S as it did in French. It became a dental spirant, like the English TH [IPA: /θ/]. It is thus more fronted than the former S. In some areas (especially in Soule Basque), Z corresponds to the French S: it is a dental spirant [IPA: /s/], whereas the S moved further back in the mouth and became a cacuminal spirant [IPA: /ʂ/]. 'Cacuminal' means that it is pronounced at the top—Latin *cacumen*—of the palate; it can also be called 'retroflex', meaning that the tongue is curled back. Vietnamese has a cacuminal spirant, which was written as a simple S, by analogy with the sound

[2]Translator's note: since the time when Haudricourt's article was published (1949), it has been shown, on the basis of reconstruction and of Vietnamese borrowings into neighbouring languages, that the phoneme written as D in the Vietnamese orthography was not a voiced prepalatal stop (closest IPA equivalent: /ɟ/) at the time when the orthography was devised, but a dental spirant, written as ð by Ferlus 1982 (see also Ferlus 2001). The closest equivalent of this sound in the International Phonetic Alphabet is /ð/; however, Ferlus makes a difference between spirants and fricatives, following Martinet 1956:24-25, who points out that spirants and fricatives are two different classes of sounds: "It is advisable to distinguish between relaxed articulations which tend towards a vowel-like articulation, for which we reserve the term *spirant*, and firmly articulated consonants, clearly characterised by friction as the air passes through the place of constriction: the latter are *fricatives* properly speaking" [our translation]; see also Martinet 1981, 1985 and Thomas, Bouquiaux and Cloarec-Heiss 1976:29-31. However, Haudricourt does not make a distinction between the two in the present paper: he uses "spirant" throughout.

just mentioned, whereas the corresponding stop [IPA: /ʈ/] was approximated through the combination TR. These consonants are rare in Europe; they are often confused by Europeans with non-anterior sibilant consonants. For instance, Chinese has a consonant transcribed as TCH in the EFEO (Ecole Française d'Extrême-Orient) system of romanisation, and as CH in the Wade-Giles system; these notations appear to suggest that it is a non-anterior sibilant stop consonant [IPA: /tʃ/], pronounced with the back of the tongue and with simultaneous protrusion of the lips, whereas in fact this is a cacuminal sound [IPA: /tʂ/], pronounced with the tip of the tongue curled upwards against the palate. Such consonants [IPA: retroflex consonants] are especially frequent in India; Indianists transcribe them with an added dot: ṣ ṭ ḍ. English and Chinese phoneticians add a forward-pointing hook at the bottom of the letter: ʂ ʈ ɖ [this has become standard in the IPA]. These consonants are rare in Indochina: in the Tonkin delta, they turned into prepalatals.

X

In Latin, the letter X stood for two sounds: it was a graphic abbreviation for CS. In Romance languages, this group of consonants underwent the same evolution as the group CT: it changed to IS (e.g. Latin *coxa*, French *cuisse* [/kɥis/] "thigh"). But in the Middle Ages, there developed a habit of linking the combinations US and IS into a single sign, which looked somewhat like the Latin letter X; moreover, an IS in French words often corresponded to a Latin X; as a consequence, X came to be used instead of S after an U or an I. This is why the French pronounce *soixante* 'sixty' as *soissante* [IPA: /swasɑ̃t/], *Auxerre* (a place name) as *Ausserre* [IPA: /osɛʁ/], and *Bruxelles* (Brussels) as *Brusselles* [IPA: /bʁysɛl/]; this also explains why the plural of *cheval* 'horse' is *chevaux*, that of *genou* 'knee' is *genoux*. It is simply a matter of writing habits. But when words were borrowed from Greek into French in the 16th century, the letter X was again pronounced as KS, GZ [IPA: /ks/, /gz/].

We have seen that in Spanish, the group IT had evolved into a prepalatal stop; the same happened to the group IS, written as X. In Old Spanish, X thus stood for a prepalatal spirant; it was heard by the French as a non-anterior sibilant sound: *Don Quixote* and *Ximena* were written by the French as *Don Quichotte* and *Chimène* [both pronounced with a /ʃ/]. But the name of the famous missionary *Xavier* (a Spanish Basque name, equivalent to French Basque *echeverry*, 'new-house') was treated like a learned word and pronounced *Gzavié* [IPA: /gzavje/].

These considerations explain why the Vietnamese prepalatal spirant was noted X.

The notation of this spirant in the scientific alphabet is ç (the letter c with a small curl below it), but if it is more like a sibilant (anterior, as in Vietnamese, or non-anterior), the diacritic for palatalisation is added to the corresponding sibilant consonant (as used in Polish: s', c'). The anterior

sibilant spirants are always noted S, Z [IPA: /s/, /z/]. (…)[3] According to English and Chinese phoneticians, sibilant stops should be written with two signs, decomposing them into a stop+spirant sequence: /ts/, /dz/ for anterior sibilants, /ʧ/, /ʤ/ for non-anterior sibilants. (…) The cacuminal stops of Vietnam and China are written ʈs̩, ɖz̩ or ʈʂ, ɖʐ. The stops written by phoneticians with two letters are said to be affricated.

N, NH, NG

The nasals (voiced stops with a lowering of the velum, such that air comes out at the nose) N, M have the same value in all languages. But in French, when they come at the end of a syllable, they are not pronounced as such anymore: instead, the vowel is nasalised (the velum is lowered during the pronunciation of the vowel). In Portuguese, the same phenomenon took place, but the N, written on top of the vowel, became ˜. Linguists took up this sign for nasal vowels, e.g. /ɑ̃/. In Spanish, a double N (a sequence of two Ns) was likewise written with a superscript ˜, yielding ñ; later, this double N became a prepalatal nasal; Indianists adopted the sign ñ for this sound. In French and in Italian, the prepalatal nasal originates in the Latin consonant cluster GN; this orthography was preserved in French and Italian, whereas in Portuguese, in Provençal and in Gascon this sound is written NH by analogy with the corresponding stop, CH. The latter spelling was adopted for the Vietnamese alphabet. English and Chinese phoneticians draw out the first stroke of the N into a curl: /ɲ/.

Lastly, in Romance languages the dorsal nasal is only occasionally found, when N is followed by C or G; this sound exists word-finally in English and German, where it is written as NG. It was therefore natural to write it as NG in Vietnamese. English and Chinese phoneticians combine both letters by dragging out the extremity of the n as in a g, thus: ŋ. Indianists write it with a dot on top of the n, thus: ṅ.

V

Let us recall the origin of V. In Latin, V and U were not distinguished: V was the notation found in inscriptions for the labiovelar vowel, and U was the cursive form for the same sound. When U was followed by a vowel, it became a labiovelar consonant [IPA: /w/], then a labiodental consonant (the voiced counterpart to F [IPA: /v/, as opposed to /f/]); this change took place around the 6th century CE, whereas no change took place in the writing until the 17th century, so that two different sounds were written the same way. The letter V was used word-initially, when it referred to a vowel (as in *vn* "one" [written as *un* in present-day French]) as well as when it referred to a consonant (as in *vent* "wind" [still written as *vent* in present-day French]). Word-medially, U was used, thus *chanure* "hemp" [present-day French: *chanvre*] and *mur* [present-day French: *mur*]. But in other languages, a new notation had been introduced much earlier than the 17th century: since the

[3]Translator's note: Two sentences were deleted at translation; they are indicated as (…).

labiovelar consonant had already become labiodental in Romance in the 6th century, the Romance-speaking scholars who devised spelling systems for Germanic languages had to create a new notation for the labiovelar consonant that they encountered in these languages: a double U (still called by that name in English), which led to present-day W (called *double V* in French). English and Northern French language varieties (Picard, Wallon) have retained a labiovelar, hence the pronunciation of the word *Wallon* [/walɔ̃/, and not /valɔ̃/], whereas in German, Dutch and Polish, W transcribes a labiodental (hence *Wagram*, *Würtenberg* are pronounced with an initial labiodental [/v/]). Linguists distinguish labiodental V and labiovelar W.

H

As a last note concerning consonants, let us recall that French has a consonant which is written but not pronounced: the so-called "aspirated H" prevents the elision of the article [thus *la hache*, *le hibou* "the axe", "the owl" and not *l'hache, l'hibou*] but is not pronounced anymore, except in a few areas (such as Lorraine and Normandie); conversely, Vietnamese has a consonant which is pronounced but not written: the glottal stop (sudden opening of the larynx) found at the beginning of vowel-initial words. In scientific writing, it can be noted by means of the *spiritus lenis* of the Greek writing system, a small apostrophe with its concave part oriented backward: ʿ

English and Chinese phoneticians use the interrogation point instead: ? [Present-day IPA notation: ʔ]

2. Vowels

A, E, I, O, U

Latin had five vowels:

(i) A, an open vowel (i.e. with maximum distance between the tongue and the palate);

(ii) and (iii): two maximally closed vowels. I, a prepalatal vowel, better described as an *anterior* vowel in terms of tongue position; I is not rounded, i.e. the lips are spread. U, a labiovelar vowel, better described as a posterior vowel (in terms of tongue position) which is rounded: the lips are protruded and rounded;

(iv) and (v): two intermediate vowels: E in-between I and A; O in-between U and A.

The U changed its phonetic value in Old French around the 10th century: it became an anterior, rounded vowel [IPA: /y/]. The scientific notation of this French sound is ü (and ö is used for the vowel in-between ü and A); this notation comes from the German, where the trema, *Umlaut* in German, stands for an abridged E: ü stands for ue, ö for oe. English phoneticians favour the Scandinavian notation Y.

Y

In Ancient Greek, the rounded posterior closed vowel [IPA: /u/], which was written Y, became an anterior rounded consonant. Latin scholars borrowed the Greek letter Y to transcribe this vowel, which was not present in Latin. In Greek, this vowel later got confused with I, and the Latin too came to pronounce it as I, hence the name "Greek I" which the letter Y retains to this day in French. Concerning its role in the Vietnamese alphabet: in French, it is commonly used to note the voiced prepalatal spirant, which stands in the same relationship to I as W does to U; this is the notation used by Indianists [and Africanists]. But the English phoneticians prefer to note this sound as j. [Notation as /j/ has become standard in IPA.]

Latin had a distinction between long and short vowels. The notations were: a for the long vowel, ă for the short one; the latter sign is used in Vietnamese. Indianists transcribe the long vowel, whereas they omit the short vowel altogether. English and Chinese phoneticians indicate length by means of a colon following the vowel, or by doubling the vowel: a: or aa. [IPA: /aː/]

Ê, Ô, Â

In Romance languages, long vowels appeared when two vowels came together; this was abbreviated by writing the vowel only once, and adding a circumflex accent ˆ, e.g. in French *aage* was rewritten as *âge*, *meur* became *mûr*. (In French, ˆ often stands for a former s, e.g. *fête* "celebration" comes from an earlier *feste*, *pâte* "paste" from *paste*.) The same happened in Portuguese, where *oo* became *ô*, *ee* became *ê*. In Portuguese, the new vowels had a more closed pronunciation: ô was in an intermediate position in-between *o* and *u*, and *ê* was intermediate between *e* and *i*. This is where the notation for Vietnamese vowels was taken up from. Neither Spanish, nor Italian, nor Provençal offered appropriate means for noting two different *e* vowels and two different *o* vowels.

In French, ô [IPA: /o/] is more closed than o [IPA: /ɔ/], and ê [IPA: /e/] is more closed than e [IPA: /ɛ/]; in view of this fact, some linguists use the grave and acute accents: ó, é for the more closed vowels, vs. ò, è for the more open vowels. Another possibility is to use a subscripted dot for the more closed vowels and a subscripted hook for the more open vowels. Lastly, English phoneticians use e, o for the more closed vowels, and ɔ, ɛ (an inverted C, and Greek *epsilon*) for the more open ones. The sign æ is used for a vowel in-between a and ɛ, like the vowel of the English word *cat*. An italicised a (i.e., ɑ) is used for the vowel of the French word *pâte*.

Vietnamese has unrounded back vowels [IPA: /ɯ/, /ɤ/]. They were noted as ư and ơ, perhaps drawing inspiration from the notation ü, ö mentioned above, though in fact ư is quite the opposite of ü in terms of lip position and tongue position: ư [IPA: /ɯ/] is back and unrounded, ü [IPA: /y/] is front and rounded. These vowels are not found in Western Romance languages; however, they are found in a Romance language of Eastern Europe:

in Romanian, the word "dog" could actually be written in Vietnamese spelling as *cưn*. These vowels are sometimes written with a trema: ï, ë, or with a subscript circle. English phoneticians write them as /ɯ/, /ə/ [IPA /ə/ would now be reserved for a truly central vowel, and /ɤ/ be used for a back, unrounded, close-mid vowel], and /ʌ/, i.e. an inverted m, e and v, respectively. The latter (/ʌ/) corresponds to Vietnamese *â*.

Second part: Tones

As a last note, let us indicate that the notation of tones was borrowed from Ancient Greek. Ancient Greek had two tones, written as ´ and ˜, and a third accent was used to note the tone of unaccented words: `. These accents were supplemented by punctuation signs. Indeed, in Romance languages, tone cannot be used to distinguish between words, but it can distinguish between two sentences: for instance, in French, interrogative *C'est vrai?* "Is that true?" and affirmative *C'est vrai.* "That is true." Thus, a simple point (analogous to a full stop) placed under the vowel, and an interrogation mark placed on top of it, supplemented the notation of tones [here are, as an illustration, the six tonal categories for the vowel A: a à á ạ ả ã].

English and Chinese phoneticians use a vertical bar placed to the right of the word, against which a small horizontal bar is added, indicating by its shape and position the height and modulation of the tone [e.g. ˥ for a high tone, ˧ for a mid tone, etc].

Conclusion

The Vietnamese alphabet is thus the product of well-identified historical facts.

The first Europeans who reached the Far East by sea came from the Iberic peninsula: they were Portuguese, Spanish, and Basque. Among them, the scholars who adapted the Latin alphabet to the Vietnamese language belonged to the Catholic clergy, and therefore knew Latin, Italian and Greek. All the peculiarities of the Vietnamese alphabet can be explained in light of these facts.

Looking at the Roman-based scientific alphabets, which were briefly presented in the course of the discussion, it appears that they fall into two groups: (i) those employed for Oriental languages that already had their own alphabets, in which case the Latin transcription is essentially a transliteration, replacing a letter of the indigenous alphabet by a letter of the Latin alphabet with a view to facilitate printing by European printers; (ii) those employed by phoneticians, who aim to transcribe all the nuances of pronunciation. The latter kind of alphabet is used in the countries where there is no alphabetic writing (e.g. China) or where spelling only has a distant link with pronunciation (e.g. England: it is no easy matter to guess the actual vowel sound in an English

word, despite the fact that English uses the Latin alphabet). But the phonetic alphabet is designed for handwriting, and difficult to print.

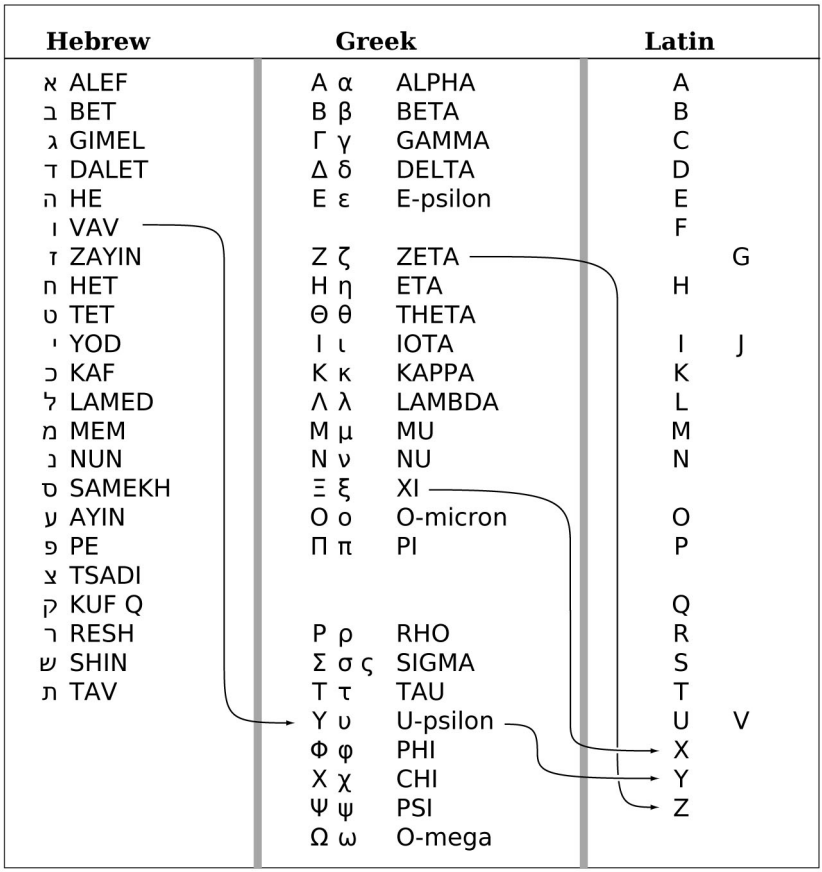


Figure 1. From the Hebrew and Greek alphabets to the Latin alphabet.

Vietnamese	French	Scientific notations : indianists and ethnologists	Scientific notations : phoneticians and sinologists
th		th	t‘
ph			ɸ
	f	f	f
	ch	š	s
	tch	č *or* tš	ʦ
ch		c *or* t´	t *or* ʨ
d		j *or* d´	d *or* ɖ
x		ç *or* ś *or* s´ *or* š´	ɕ *or* ʂ
gi	z´ *or* ž´		ʑ
s		ʂ	ʂ
tr		ʈʂ	ʦ
nh	gn	ñ	ɲ
ng		ṅ	ŋ
vocalic initial		’	ʔ
	eu	ö œ	ø
	u	ü	y
ư			
u	ou	u	u
ơ		ə	ə
	ill	y	j
ê	é	ẹ	e
e	è	e	ɛ
ô	o	ọ	o
	o		ɔ

Figure 2. Equivalences across alphabets: Vietnamese, French, and scientific notations.

REFERENCES

Ferlus, Michel. 1982. “Spirantisation des obstruantes médiales et formation du système consonantique du vietnamien.” *CLAO* 11(1):83-106.

Ferlus, Michel. 2001. “Les hypercorrections dans le thổ de Làng Lỡ (Nghệ An, Vietnam) ou les pièges du comparatisme.” *Paper Presented at the Quinzièmes Journées de Linguistique de l’Asie Orientale,* Ecole des Hautes Etudes en Sciences Sociales, Paris.

Martinet, André. 1956. *La Description Phonologique Avec Application au Parler Franco-Provençal d’Hauteville (Savoie).* Genève: Droz.

Martinet, André. 1981. “Fricatives and spirants.” In B. P. Mallik, ed., *Suniti Kumar Chatterji Commemoration Volume* 145-151. Burdwan, West Bengal, India: Burdwan University Press.

Martinet, André. 1985. “Two proposals.” *The Study of Sounds (Onsei no Kenkyuu), Commemorative Volume for the 50th Anniversary of the Phonetic Society of Japan* XXI:67-72.

Suriya, Ratanakul, David Thomas and Premsrirat Suwilai, Eds. (1985). *Southeast Asian Linguistic Studies Presented to André-G. Haudricourt*. Bangkok: Mahidol University.

Thomas, Jacqueline, Luc Bouquiaux and France Cloarec-Heiss. 1976. *Initiation à la Phonétique: Phonétique Articulatoire et Phonétique Distinctive*. Paris: Presses Universitaires de France.

Received: 24 January 2008

7 rue Guy Môquet, 94800 Villejuif,
FRANCE
<alexis.michaud@vjf.cnrs.fr>

Cua (Kor) historical phonology and classification*

Paul SIDWELL
Australian National University (Canberra), Australia
Centre for Research in Computational Linguistics (Bangkok), Thailand

Abstract

The Cua language of the Central Vietnam Highlands has been little studied, and is barley known beyond a few unpublished wordlists. Recently the author acquired an extensive manuscript lexicon (Maier & Burton 1981), and also made some brief field recordings. With the help of this data the phonological history of Cua is reconstructed, with reference to a preliminary reconstruction of proto-Bahnaric. Additionally, the question of classifiaction is discussed; it was treated as Eastern North Bahnaric by Smith (1973) and Central Bahanric by Sidwell (2002). The present study argues that Cua reflects a distinct Eastern branch which has been influenced by contact with North Bahnaric (in addition to Vietnamese and Chamic).

Introduction

Cua, also known as Kol, Kor or Traw, is something of an enigma among the Bahnaric languages of the Vietnam Central Highlands. It embodies a rather idiosyncratic combination of phonological developments that have served to obscure its history and classification. The language is not well studied (not being easily accessible), yet the data available does permit some useful analyses, especially the manuscript rhyming dictionary of Maier & Burton (1981). In this paper I present the results of internal and comparative reconstruction of Cua historical phonology, framed in terms of its apparent development from Proto-Bahnaric (PB)[1] (the latter modelled here more or less as it was discussed in Sidwell 2002 and 2003).

*Research that made this paper possible was supported by generous assistance of the National Endowment for the Humanities (Washington DC). Any views, findings, conclusions, or recommendations expressed in this publication do not necessarily reflect those of the National Endowment for the Humanities. I would also like to thank Vurnell Cobbey (SIL, Dallas) for valuable assistance with access to archival materials.

[1]Abbreviations used in this paper: PB: Proto-Bahnaric, NB: North Bahnaric, WB: West Bahnaric, CB: Central Bahnaric, MK: Mon-Khmer, SIL: Summer Institute of Linguistics.

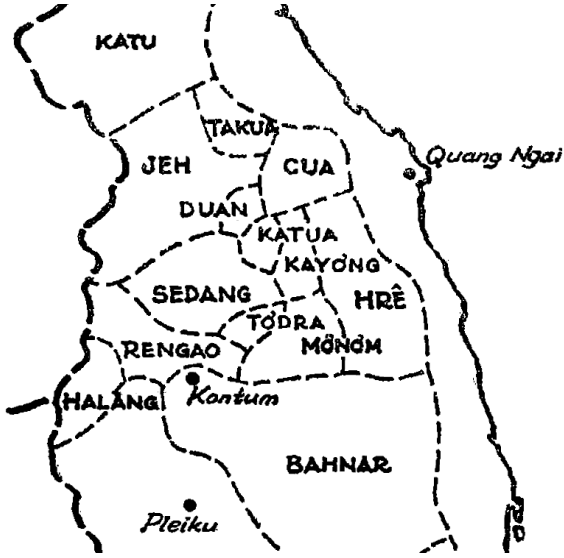


Figure 1. Map showing location of languages of the Vietnam Central Highlands
(fragment of SIL map dated March 1966)

According to Maier & Burton (1981: 2):

> Cua is a member of the Bahnaric branch of the Mon-Khmer family of languages. The majority of the Cua people, who number 10,000-15,000, live in the mountain area of Tra Bong district in Quang Ngai province, central Vietnam. The Cua people call themselves "Kool," their term for 'montagnard' in general, this name being modified to "Cua" by the Vietnamese. The Cua people who live in the Tra Bong valley are called Kool Dong, "Valley Cua," and the Cua who live in the mountains are called Kool Doot or Kool Taal "High Cua."

The language is not extensively documented by Western scholars, and one is largely restricted to using manuscript materials. Sources readily available in English (via the Summer Institute of Linguistics [SIL] in Dallas and Bangkok) are:

- Phillips (1959), a manuscript lexicon of 187 words in four dialects
- Burton (1969), a discussion of clause structure,
- Maier (1969), an analysis of phonemes,
- Maier & Đình Van Cau (1976) a Cua-Vietnamese-English thesaurus in the form of a computer printout,
- Maier & Burton (1971), a manuscript lexicon of 281 words of a lowland Cua dialect
- Smith (1973), a discussion of classification,
- Maier & Burton (1981), a manuscript rhyming dictionary of somewhat more than 3000 items, annotated to the effect that it is based on a 1966 manuscript

In addition there are some minor literacy and bible translation notes at the David Thomas Library (SIL) in Bangkok, and this writer has collected some short wordlists in the field.

Cua Phonology

Maier (1969) offers a phonemic description of Cua. The analysis is of the straightforwardly structuralist type typical of the day, and is very useful, although perhaps not quite exhaustive. The set of phonemes, tabled according to their place in the word-structure, are found to be as follows (transcribed into IPA from Maier's typewriter based notation):

Phonemes of Cua (based on Maier 1969):

Initials

p	**t**	**c**	**k**	**ʔ**	
ph	**th**		**kh**		
b	**d**	**ɟ**	**g**		
ɓ	**ɗ**				
m	**n**	**ɲ**	**ŋ**		
w	**r, l, hl**	**j**			
	s			**h**	

Finals

p	**t**	**c**	**k**	**ʔ**
wʔ		**jʔ**		
m	**n**		**ŋ**	
w	**l**	**j**		
	lh	**jh**		**h**

Vowels:

iː	**ɨː**	**uː**	**i**	**ɪ**	**ɨ**	**u**	**ia**	**ua**
eː	**əː**	**oː**	**e**		**ə**	**o**	**ea**	**oa**
ɛː	**aː**	**ɔː**	**ɛ**		**a**	**ʌ**		

The template for the phonological word is rather strict; it consists minimally of a main syllable $C_1(C_2)V(C_3)$ and an optional presyllable C_4v. Concerning the main syllable onset, C_1 can be any of the "initials" listed above, while the optional medial C_2, is restricted to ***r, l, j*** and ***w***, plus nasals

after initial ***h*** and ***ʔ***, unless we accept sequences ***hm, hn, hɲ, hŋ, ʔm, ʔn, ʔɲ, ʔŋ*** as unitary phonemes. Maier entertains either analysis, tending to the latter (although not indicating this in her table of phonemes). Main syllable codas (C_3) show the same places and manners of articulation but without a general voicing or aspiration contrast. Final postglottalized glides occur, as do an aspirated lateral and yod ("complex finals"). In fact the dictionary of Maier & Burton (1981) also lists a number of forms with postglottalized final nasals, although these appear largely to reflect Vietnamese loans with săc tone. There are no tones or registers mentioned in any descriptions, and I have not heard any.[2] The main vowel (V) is any of the table above, although there are various distributional restrictions that give rise to remarkable asymmetries, which will be discussed in detail below. The presyllable vowel (v) appears to be entirely prosodic, having no distinctive timbre, and is consistently written ***a*** in the dictionary of Maier & Burton.

The Cua phonemes can be compared with the PB phoneme inventory as follows:

Phonemes of Proto-Bahnaric (based on Sidwell 2002, 2003):

Initials

p	**t**	**c**	**k**	**ʔ**
b	**d**	**ɟ**	**g**	
ɓ	**ɗ**			
m	**n**	**ɲ**	**ŋ**	
w	**l, r**	**j**		
	s			**h**

Finals

p	**t**	**c**	**k**	**ʔ**
m	**n**	**ɲ**	**ŋ**	
w	**l, r**	**j**		
	s			**h**

Vowels:

iː	**ɨː**	**uː**	**i**	**ɨ**	**u**		**ia**	**ua**
eː	**əː**	**oː**	**e**	**ə**	**o**			
ɛː	**aː**	**ɔː**	**ɛ**	**a**	**ɔ**			

Historical Development of Cua Phonemes

Initials

At first glance the only difference between the Cua and PB initials is in the aspirated stop series and voiceless lateral. It is often a vexing question whether to treat aspirates in MK languages as sequences of stop + /h/ or unitary phonemes. In languages such as Khmer, Koho and others it is obvious that

[2] I recorded the elicitation of some basic word lists (with the assistance of a young Bahnar fellow who prefers not to be named), and these recordings provide no indication of voice registers.

prevocalic aspirates (as opposed to preconsonantal aspirates)[3] can be split by infixes, confirming their analytical structure, and this is the approach I have taken with PB. But in the available Cua data examples of productive morphology of this type are lacking. Also, as Maier (1969) points out, examples such as ***khwal*** 'curly', ***khwaːl*** 'hoe', although infrequent, suggest aspirates occupying a single C_1 slot before a medial C_2. It appears that a general reduction in morphological productivity, plus increasing borrowing from Vietnamese (whose initial sequences such as ***xw-*** are borrowed as ***khw-***) may have conditioned a reanalysis of these initials in Cua.

The other initial which is missing in PB is the voiceless lateral /hl/. Effectively Cua has two voiceless laminal fricatives (/hl, s/), differing only in their manner of articulation. Comparative data demonstrates that /hl/ is the direct reflex of PB ***s** and ***sl**, part of a larger chain shift in which PB ***c** shifted to /s/, and the resultant empty /c/ slot was filled mainly by loans from Vietnamese and Chamic. Examples:

Gloss	Cua	NB[4]	CB	WB	PB
'hair'	hloːk	*sək (Jeh suk)	*sɔk (Bahnar sɔk)	*sɔk (Jru' sok)	*s-
'divide/share'	ʔahlɔːk	*sɔːŋ (Halang sɔaŋ)	*ʔəsɔːŋ (Bahnar ʔəsɔːŋ)	*sɔːŋ (Brao sɔːŋ)[4]	*s-
'honey(bee)'	hluːt	*sut (Halang sut)	*su(ː)t (Bahnar suːt)	*sut (Jru' suːt)	*s-
'leaf'	hlaː	*hlaː (Jeh laː)	*hlaː (Bahnar hlaː)	*slaː (Jru' hlaː)	*sl-
'dog'	sɔː	*cɔː (Jeh cɔː)	*cɔː (Bahnar sɔː)	*cɔː (Jru' cɔː)	*c-
'bird'	seːp	*cem (Jeh cim)	*ceːm (Bahnar sɛːm)	*ceːm (Jru' ceːm)	
'to eat'	saː	*caː (Jeh caː)	*caː (Bahnar saː)	*caː (Jru' caː)	*c-

[3]As is well known, Khmer preconsonantal stops are phonetically aspirated before nasals and liquids.

[4]NB reconstructions & data are from Sidwell (ms.), a manuscript monograph widely circulated for comments in 2002 and still available upon request from the author; WB reconstructions & data are from Sidwell & Jacq (2003), and CB reconstructions & data are from a ms. of Sidwell, which is still in preparation. Much of this same data can be viewed freely online at the author's project website: www.sealang.net/monkhmer.

[4]Brao *sɔːŋ* 'to pay' may be borrowed from/influenced by Khmer *sɔːŋ* 'give back, payback, restore, compensate'.

Examples of /c/ entering Cua via loans:

Cua ***cih*** 'write'	Proto-Chamic[5] ****cih*** 'write, draw', Wr.Cham ***ciḥ***
Cua ***cowah*** 'sand'	Proto-Chamic ****cuah*** 'write, draw', Wr.Cham ***cụaḥ***
Cua ***cɛː*** 'tea'	Vietnamese *chè*
Cua ***cʌm*** 'squat'	Vietnamese *chồm*

These shifts among the laminals resemble, but are quite independent of, the merger of ****c-*** > /s-/ which occurred in Bahnar, Tampuon and South Bahnaric (together "South Central Bahnaric"). While the latter was a simple merger of ****c-*** and ****s-***, Cua has kept the reflexes of these proto-phonemes distinct.

Both Cua and Proto-Bahnaric show imploded stops /ɓ, ɗ/ in their inventories, but these series are not related. In fact like most Bahnaric language, Cua has merged the imploded stops with the plain voiced series, and the presently imploded stops of Cua have other sources: principally borrowing, and also by assimilation of glottal + nasal sequences. Etymological imploded stops appear to continue unchanged only in Bahnar. Examples:

Cua	**Remarks**
baːl 'two'	< *ɓaːr PB (cf. Bahnar ɓaːr)
tabak 'sprouts'	< *təɓaŋ " (cf. Bahnar ɓaːr 'bamboo shoots')
daːk 'water'	< *ɗaːk " (cf. Bahnar ɗaːk)
doːp 'ripe'	< *ɗuːm " (cf. Bahnar ɗuːm)
taɓak 'hang, suspend'	< taɓak Proto-Chamic
ɓẹŋʔ 'candy, sugar'	< *bánh* Vietnamese (with sac tone realized as post-glottalization)
ɗɔŋ 'hammer'	< *đóng* Vietnamese
ɗwat 'protective hat'	< ɗuan Proto-Chamic (via Bahnar ɗwan ?)[6]
ɗəw 'just, recent'	< ʔnaːw PB? (cf. Bahnar ʔnaːw 'new, recent')
kaɗiap ~ kaʔniap 'close (eyes)'	doublet

Other than the above-mentioned changes, Cua initial consonants continue their PB values essentially unchanged, showing no signs of devoicing or other general restructurings often found in MK languages.

Finals

The history of the word final consonants is characterized by several types of changes that, taken together, have profoundly affected the phonetic

[5]Chamic data and reconstructions used here are from Thurgood (1999).

[6]A cognate is also attested in Vietic, but lacks the diphthonged vowel, e.g. Vietnamese nón 'conical hat'.

character of Cua vis-à-vis other Bahnaric languages.[7] These can be summarised as follows:

- lateralization of *-s
- hardening of nasals to oral stops
- emergence of post-glottalized finals
- loss of laminal versus velar oppostion

Perhaps the single most striking aspect of Cua phonology is the hardening of final nasals. Smith (1973) identified the phonological environment for this process as any syllable not having an initial nasal or laryngeal, such that the great majority of final nasals hardened. He formulated the rule in the following figure:[8]

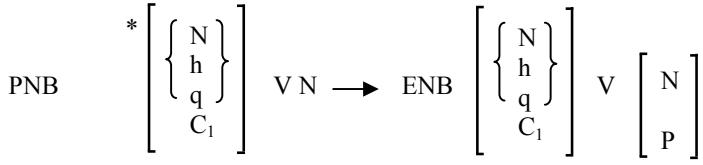


Figure 2. Rule for hardening of nasal finals in Cua and Katua by Smith (1973:115)

The above rule does not appear to remain active, and probably belongs to a much older stage of the languages; an examination of the corpus finds numerous exceptions among borrowings, expressive formations and other more recent lexical innovations/acquisitions. Indicative examples of the operation of the rule, and some exceptions, follow:

Type	Cua	Remarks
*-m > -p	klɔːp ‘liver’	< *klɔːm PB
*-n > -t	pɔːt ‘four’	< *puan PB
*-ɲ > -c	pɛc ‘to shoot’	< *paɲ PB
*-ŋ > -k	hrɛːk ‘100’	< *hriaŋ PB
no change (prevocalic /n/)	kanɨm ‘urinate’	< *kʔnoːm PB
no change (prevocalic /ŋ/)	ŋɨːm ‘sweet’	< *ʔŋaːm PB

[7]All but the first of these changes also occur viously in NB languages, but it is Cua that shows the whole suite.

[8]The rule also applies (apparently) to the language Kotua, and on the basis of this shared feature Smith proposed the sub-grouping East-North-Bahnaric. However, Kotua, although lacking registers, shows the characteristic vowel restructuring of NB languages, indicating that it should more probably be sub-grouped with Hrê-Sedang.

no change (prevocalic /h/)	thəːm ‘eight’	< *thaːm PB
exception (loan)	wɔŋ ‘hammock’	< *võng* Vietnamese
exception (loan)	siːn ‘request permission’	< *xin* Vietnamese
exception (loan)	kuŋ ‘also’	< *cũng* Vietnamese
exception (loan)	phanaːn ‘starving’	< *pha ngan* Vietnamese

Not surprisingly, given similar parallels in Vietnamese and various other languages of the area (especially Jeh and Halang), the final laminals ***-c** and ***-ɲ** underwent various neutralizations. Historical ***-ɲ** and ***-ŋ** have merged, with */ŋ/ the unmarked reflex, and /ɲ/ commonly recorded after front vowels (although most are simply noted with /ŋ/ in Maier & Burton’s 1981 dictionary). Historical ***-c** tends to merge with ***-t** after back/central vowels, and with ***-k** after front vowels (by dissimilation) and there are numerous examples of alternates attesting this in the data. In various cases it is evident that laminal nasals shifted articulation before hardening to stops. Some examples:

Type	Cua	Remarks
*-ŋ > -c ~ -k	ɟeːc ~ ɟeːk ‘to become’	cf. Bahnar ɟiːŋ
*-ɲ > -ŋ > -k	pleːk ‘sky’	cf. Bahnar plɛɲ
*-ɲ > -c ~ -t	taːc ~ taːt ‘to weave’	< *taːɲ PB
*-c > -t	ramuːt ‘ant’	< *smoːc PB
-c ~ -k	paleːc ~ paleːk ‘to pledge’	

Interestingly, Cua has a contrast between two final laminals, /lh/ and /jh/, quite unlike other Bahnaric languages, which usually only have one (typically a sound that has a wide allophonic range: [s ~ ʃ ~ ç ~ jh]). In this case finals /lh/ and /jh/ appear to reflect a split, either a phonologically conditioned split of PB final ***-s,** or a general shift of ***-s** to /-lh/, with /-jh/ subsequently introduced by borrowings. Examples:

Cua	Remarks
pɔːlh ‘calf of leg’	*puas PB
roːlh ‘elephant’	*ruas "
bʌlh ‘snake’	*ɓəs "
wʌlh ‘measure’	*was "
barajh ‘type of mountain rice’	*braːs Proto-Chamic
taːjh ‘to call’	*tas ‘loud noise’, cf. Wr.Khmer *kantas* ‘sneeze’
kawəjh ‘beckon with hand’	Wr.Khmer *vas* ‘gesticulate’, Vietnamese *vẫy* ‘to wave’

The voiced final lateral /-l/ reflects a merger of ***-r** and ***-l** to /-l/ (realized as /-r/ in the Kool Taal dialect). Examples:

Cua	Remarks
ɟamiːl ‘ribs’	cf. ɟəmiːr Bahnar
pʌl ‘to fly’	cf. par Bahnar
gʌl ‘drum’	cf. səgər Bahnar
phɔɔl ‘soul, spirit’	cf. pəhŋɔɔl Bahnar
kʌl ‘chop wood’	cf. kal Bahnar ‘chop large trees’

The post-glottalized finals - somewhat infrequent in the corpus at around 1% only - appear to reflect various complex origins, including: vowel breaking, metathesis, and borrowing. The following examples are indicative:

Cua	Remarks
baraw**ʔ** ‘work’	cf. *bruã**ʔ** Proto-Chamic (with metathesis)
kwaj**ʔ** ‘gather, amass’	cf. kuai**ʔ** Chru, N.Roglai (Highland Chamic only)
luj**ʔ** ‘last born (child)’	cf. *taluc Proto-Chamic, e.g. N.Roglai talui**ʔ**
ɟaːw**ʔ** ~ **ʔ**jaːw**ʔ** ‘to count’	cf. ɟo**ʔ** Bahnar (with vowel breaking)
kaː gaj**ʔ** ‘kind of fish’	< *cá gay* Vietnamese
səw**ʔ** ‘bad, ugly’	< *xấu* Vietnamese

However, there are also a few examples of /-jʔ/ that are puzzling. For example, the first two examples below, **sɔːj****ʔ** and **ʔ**nɔːj**ʔ**, have apparent cognates with finals that do not look like obvious sources of [jʔ]. Such examples will require further investigation. In two further examples below, ‘one’ and ‘before’, the finals may have a straightforward explanation. Various Bahanric languages have a special form for ‘one’ used with classifiers, e.g. the regular term in Bahnar for counting is /moːŋ/, but when followed by a classifier /ʔməjʔ/ is used. The latter is the regular MK etymon for ‘one’ in which the main vowel is reduced, and the final is postglottalized, perhaps as a prosodic juncture. The ‘before, first, ahead’ form may be contaminated by analogy with ‘one’.

Cua	Remarks
sɔːjʔ ‘till, cultivate, prune’	cf. Wr.Khmer bhjuər furrow, Stieng cuər ‘to plow’
ʔnɔːjʔ ‘more, another’	cf. Bahnar ʔnaaw ‘new, recent’, Tampuon naawʔ ‘again, further’
mujʔ ~ muj ‘one’	cf. Bahnar ʔməjʔ ‘one used with classifiers’
ʔadrɔːjʔ ‘before, ahead’	cf. Bahnar hədrɔːj ‘before, first’

The general problem of accounting for the /-jʔ / final in Bahnaric languages, especially in cases where it contrasts with /-c/ (since we would anticipate decomposition of /-c/ as the first source of /-jʔ/), remains unresolved. It may be that a ***-jʔ*** : ***-c*** distinction must be reconstructed for Proto-Bahnaric,

although presently the lack of regular correspondences makes the problem difficult to investigate.

Vocalism

Cua vocalism is exceedingly interesting from the perspective of reconstructing PB vowels. On the surface the Cua vowel inventory looks like what one finds in a more or less typical conservative or "unrestructured" (to use the framework of Huffman 1985) Mon-Khmer language. However, the specific distributions of various vowels tell a very different story, one of a web of conditioned shifts, splits and mergers that is unique to Cua.

There are two main stories in the history of Cua vowels. One is the story of the central vowels raising in timbre, and other is the diphthongs merging and re-emerging from the front and back monophthongs. Both of these complicated changes correlate fairly neatly with the manners and places of articulation of the immediately adjacent consonants.

Raising of Central Vowels

Maier (1969:19) writes:

> Glancing at the occurrences of *ɯ* and *ʊ*, one could wonder whether they are allophones of the same phoneme. The higher *ɯ* may be preceded by nasal consonants but *ʊ* never is; however, with several other consonants they do contrast in minimal environments.

Maier is apparently referring to both the long and short pairs /ɨː, ɨ/ and /əː, ə/. Out of the complete corpus of more than 3000 entries, there are about 60 examples of /ɨː/, 150 of /ɨ/, 120 of /əː/ and 120 of /ə/. Among these we find the following distributions in the etymologically Bahnaric vocabulary:

- /əː, ə/ occur after all initials except nasals (specifically in the prevocalic position),
- /ɨː/ occurs only after nasal and a couple of examples after /r/,
- /ɨ/ occurs after all consonants, approximately half of these after nasals.

Beginning with /ɨː/, the distribution is quite strongly restricted. The couple of examples of /rɨː/ are the adverbial **rɨːt rɨːt** 'carefully' and derivative **hrɨːt** 'careful, small', plus the temporal **tamrɨː** 'day after tomorrow'. The latter is transparently derived from **baːr** 'two' (cf. Nyaheun **mbraː** 'day after tomorrow'), which immediately gives a clue to the source of /ɨː/ in Cua. But first, some background on the question of /ɨː/ in Bahnaric generally.

According to sub-grouping, we generally find that:

- WB languages have a phoneme /ɨː/, often realised as a diphthong [ɨə], which is well distributed and reasonably frequent;
- NB languages lack a long high central vowel, due to restructuring associated with the emergence of registers;
- CB languages have infrequent high central vowels, usually in loans or transparently allophones of other vowels.

In Sidwell (2002) I reconstructed a PB ***ɨː**, on the basis of the correspondence of WB ***ɨː** to CB and NB ***iː**, contrasting with PB ***iː**. Examples:

	PB	**WB**	**CB**	**NB**
'banana'	*prɨːt	*prɨːt (Jru' prɨət)	*priːt (Bahnar priːt)	*priːt (Jeh prịat)
'weep/cry'	*ɲɨːm	*ɲɨːm (Jru' ɲɨəm)	*ɲiːm (Alak ɲiːm)	*ɲiːm (Kayong ɲiem)
'rain'	*ʔmɨː	*ʔmɨː (Jru' ʔmɨə)	*ʔmiː (Alak ɓiː)	*ʔmiː (Halang ʔmịa)
'bone'	*ktsɨːŋ	*k[r]ʔtɨːŋ (Jru' ktɨəŋ)	*ktiːŋ (Bahnar kətiːŋ)	*ksiːŋ (Jeh kəsịaŋ)
'frog'	*kiːt	*kiːt (Jru' ŋkiət)	*kiːt (Bahnar kiːt)	*kiːt (Jeh kịat)
'sick'	*ɟiːʔ	*ɟiʔ (Jru' ɟiʔ)	*ɟiː (Stieng ɟiː)	*ɟiʔ (Jeh ɟịʔ)
'dig'	*ciːr	--	*siːr (Bahnar siːr)	*ciːl (Jeh cịːl)

External comparisons, such as Mon **prạt** 'banana' and Khmu' **kmaʔ** 'rain' indicate that this PB ***ɨː** goes back to PMK ***aː**. It is also clear that the raising of /aː/ to /ɨː/ was an ongoing tendency within Bahnaric, for example, within WB examples of ***ɨː** derive from PB ***aː**, and there is a further tendency to raise /aː/ to /əː/ specifically within Jru' (before labials). Examples:

	PB	**WB**	**CB**	**NB**
'wind'	*kəjaːl	*kəjɨːl (Jru' kəjɨəl)	*kəjaːl (Bahnar kjaːl)	*kəjaːl (Jeh kəjaːl)
'maggot'	*sraːj	*srɨːj (Jru' srɨəj)	*hraːj (Bahnar hraːj)	*hraːj (Sedang hrẹ)
'tiger'	*klaː	*klɨː (Jru' klɨə)	*klaː (Bahnar klaː)	*klaː (Sedang klạ)
'sweet'	*ʔŋaːm	*ʔŋaːm (Jru' ʔŋəːm)	*ʔŋaːm (Bahnar ʔŋaːm)	*ʔŋaːm (Jeh ʔŋaːm)
'blood'	*phaːm	*phaːm (Jru' phəːm)	*phaːm (Bahnar phaːm)	*pəhaːm (Jeh pəhaːm)
'crab'	*ktaːm	*kʔtaːm (Jru' ktəːm)	*ktaːm (Bahnar ktaːm)	*ktaːm (Jeh kətaːm)

As Maier noted above, Cua /ɨː/ is effectively in complementary distribution with /əː/, given their distribution after nasals. However, it is also apparent that Cua /aː/ is similarly restricted; out of some 420 words with /aː/ in the corpus I count only 15 with prevocalic nasals, and among them are various obvious loans, such as **phanaːn** 'starve to death' < Vietnamese *pha ngan*, **ɲam ɲaːm** 'fuzzy, rough' < Vietnamese *nhám* 'rough'. Thus the most straightforward hypothesis is that historical ***aː** and ***əː** both raised and merged to /ɨː/, filling the vowel space that was emptied by the earlier merger of ***ɨː** and ***əː** to ***iː** - apparently generally - beyond WB. Examples:

Cua	**PB**	**WB**	**CB**	**NB**
saŋɨːj 'far'	*cŋaːj	*cŋaːj (Jru' hŋaːj)	*cŋaːj (Alak caŋaːj)	*sʔŋaːj (Jeh ʔiʔŋaːj)
ʔamɨː 'younger aunt'	*maː	--	*maː (Bahnar maː)	*maː (Jeh maː)
kaʔŋɨːl 'blinded'	*k[ʔ]ŋaːl	--	--	*kŋaːl (Halang kəŋaːl)
ŋɨːt 'to let cool off'	(*ŋaːc)	--	Bahnar ŋaːc 'to cool'	--
ʔamɨːj 'daughter-in-law'	*[ ]maːj	*kmaːj (Jru' kmaːj) 'widow'	--	*maːj (Jeh maːj) 'd-in-law'
hmɨːt 'miss, remember'	(*məːt)	*məːt (Jru' mʌːt) 'to love'	--	--
kaʔŋɨːt sʌk 'put up above'	(*[ ]ʔŋəːŋ)		Bahnar ʔŋəəŋ 'pull oneself up'	Rengao həŋe̤ːn 'lean upright (against) '
ʔaŋɨːk 'to look upwards'	(*ŋəːk)	--	Bahnar ŋəːk 'to bend head back'	--

Unlike /ɨː/, Cua /ɨ/ occurs after all types of initials. None-the-less, we still find that, parallel to /əː/, the short /ə/ also does not occur after initial nasals. Additionally, of 430 words with a short /a/, I count only 17 with prevocalic nasals. The working hypothesis therefore, foreshadowed to some extent in Sidwell (2002), is that while both PB ***a** and ***ə** raised and merged to /ɨ/, PB ***ɨ** was not previously lost from the pre-Cua system by merger(s).[9] Consequently

[9]This is not to say that all examples of PB ***ɨ** were transmitted to Cua without change, since some other conditioned shifts/mergers have occurred; e.g. Cua **suʔ** 'to return' < PB ***cɨʔ** (with backing of vowel by dissimilation of initial).

this is why we count as many as 150 examples of words with /ɨ/ in the corpus, since the directly inherited words with /ɨ/ will have been joined by those resulting from the raising of ***a** and ***ə** after nasals.

Elsewhere in Bahnaric, especially in South Central and NB, it is evident that PB ***ɨ** was most often lowered, and in some cases backed or fronted under the influence of final consonants. This had the effect that, beyond WB, only Alak and Taliang/Kasseng show evidence of the PB ***ɨ/*ə** contrast.[10] Unfortunately, the extent of available data for either of these languages is such that I have not found any relevant examples involving PB ***ə** after a nasal, and we have a conundrum for the present: did PB ***ə** raise after nasals in Cua, paralleling the development of ***əː**, or were there simply no examples of such a sequence in the proto-language? More research is required. Examples:

Cua	**PB**	**WB**	**CB**	**NB**
hŋɨj 'day'	*tŋaj	*tŋaj (Jru' tŋaj)	*taŋaj (Alak taŋaj)	--
sanɨm 'year'	*cnam	--	*cnam (Bahnar sənam)	*hnam (Jeh hnam)
panɨl 'wing'	*pnar	*pnar (Jru' pnar)	*pnar (Bahnar pənar)	*pnar (Jeh mənal)
panɨh 'portion, half'	*p[ʔ]nah	--	*pʔnah (Bahnar məʔnah)	*ʔnah (Halang ʔnah)
kʌʔ gɨl 'head'	*gɨl	--	(Bahnar kəl, Alak gɯɯ)	*gəl (Rengao gl̥)
dɨk 'stand up'	*dɨk	*dɨk (Jru' dɨk 'to climb')	*dɨk (Alak dɯk 'rise, stand')	*dək (Jeh dạk 'go up')
rabɨʔ 'at night'	(*[ ]bɨʔ)	*trbɨʔ ~ *ʔmbɨʔ (Jru' hbɨʔ)	(Alak m̥bɤʔ)	--
hnɨm 'to bang against, knock'	*tɨm	*tɨm 'hit', *trnɨm 'hammer, mallet' (Jru' hnəm)	--	--

[10]One reviewer suggested that the ***ɨ** correspondence may instead reflect PB ***ə**, due its limited distribution.

Diphthongs

The development of Cua diphthongs involves various splits and mergers, conditioned by place of articulation of finals. These changes are no longer productive, and subsequently numerous structural exceptions have been introduced by the intrusion of loans, mainly from Katuic, Chamic and Vietnamese.

Within the history of the Bahnaric languages there is a complicated relationship between the diphthongs ***ia** and ***ua**, and the low monophthongs ***ɛː** and ***ɔː** (discussed to some extent in Sidwell 2003). In my reconstruction of PB, ***ua** was restricted to syllables with apical or laminal finals, while ***ia** could only occur with labial, apical or velar finals. Neither could occur in open syllables (zero coda). By contrast, ***ɛː** and ***ɔː** were effectively unrestricted in their distributions, with the one exception that ***ɛː** did not occur before laminals. It is also noteworthy that ***ɛː** was markedly infrequent in closed syllables, possibly reflecting various diphthongizations in the development of Proto-Bahnaric from PMK. We can represent the distributions of these proto-vowels as follows (where shaded areas indicate licit combinations and the lighter shading of the more suspect distributions):

	-p	-m	-w	-t	-n	-r	-l	-s	-c	-ɲ	-j	-k	-ŋ	-h	-ʔ	-ø
ia																
u a																
ɛː																
ɔː																

Analysing the etymologically Bahnaric vocabulary within Cua, some regular patterns emerge:

1) ***ua** and ***ɔː** merged to /ɔ/ before apicals and laminals, creating a temporary gap in the paradigm (e.g. the few examples of /-uat/, /-uac/ in Maier & Burton's dictionary are predominantly loans). Some new cases of /ua/ emerged from the diphthongization of ***uː** before velars.

2) Most cases of ***ia** continued without change into Cua, although before ***-r** and ***-l**, ***ia** merged with ***iː** to /iː/. ***ia** also variously merged with ***eː** to /eː/, and in an apparent chain shift some examples if ***iː** then diphthonged to /ia/.

Examples:

Sound change	**Cua**	**Proto-Bahnaric**
PB *ua > Cua ɔː /_ C[+ apical, + laminal]	pɔːlh ‘calf of leg’ pɔːt ‘four’ paŋɔːt ‘hungry’ hlɔːj ‘tail’ ɟɔːj ‘deer’ rɔːj ‘fly’ rɔːt ‘buy’ kɔːt ‘child’	*puas *puat *paŋuat *suaj *ɟuaj *ruaj *ruat *kuan
PB *uː > Cua ua /_ C[+ velar]	guak ‘ladder’ suak ‘axe’	*guːŋ *cuːŋ
Sound change	**Cua**	**Proto-Bahnaric**
New uaC[+ velar] < borrowing	kuak ‘to dig’ < Viet. *cuốc*	
PB *ia, *iː > Cua iː /_ r, l	ʔiːl ‘chicken’ pariːl ‘hail’ khiːl ‘wind’ siːl ‘dig’ tiːl ‘seed’	*ʔiar *prial *kjaːl *ciːr *tiːl
PB *ia > Cua eː /_ C[+ velar]	hreːk ‘100’	*hriang
PB *ia > Cua ia	rapiat ‘tongue’ kaniah ‘finger nail’ siap ‘raise animals’ kadiap ‘onion’	*lpiat *krniah *ciam *kdîam
PB *iː > Cua ia /_ C[+ velar]	ɟiak ‘rice field’ cpiak ‘civet’	*ɟiːk ‘to hoe’ *spiːk ‘civet’

The innovative diphthongs /ea/ and /oa/ are infrequent in the corpus; I count only 30 examples of /ea/ and 13 of /oa/. Significantly, six examples of /oa/ are written by Maier & Burton with alternates, e.g.: **ɟoaʔ ~ ɟowaʔ** ‘to tread’, and paralleling this, several examples of /ea/ have Bahnaric etymologies suggesting disyllables:

Cua	**Remarks**
dea ‘thatch’	< PB *[g/d]ajaa (cf. Bahnar gajaː, Halang dajaː)
kea ‘ginger’	< PB *kajaa (cf. Alak kajaː, Stieng caː)

These facts suggest that the notation of diphthongs /ea/ and /oa/ may record forms with medial glides, with or without an epenthetic schwa.

Also, there are several examples involving words with prevocalic /r/ that are unambiguously from an earlier /ia/. This environment is a typical trigger for reanalysis of diphthongs in Bahnaric languages, e.g. Chrau **rəwəj** 'fly' < ***ruaj**, **rəjəŋ** '100' < ***riaŋ**. Cua examples:

Cua	**Remarks**
reah 'root'	< PB *riah (cf. Srê rias)
ʔareak 'crab'	< PChamic *ʔariaŋ 'crab'

Other Vowel Changes

There is one further aspect of the vocalism to discuss; Cua has the contrasting pair of /ɪ, i/ among the short high front vowels. The contrast is real, but restricted to before glottal stops only. Tentatively we can explain this as the reanalysis of a PB length contrast which has been otherwise lost from the system, such that /ɪ/ likely reflects an historically short ***i**, while /i/ reflects a long ***i**ː.

There is a structural correspondence between syllables with a short vowel and glottal final in Cua, Bahnar and various other languages, and an open syllable in South Bahnaric, e.g. Cua **ɟiʔ** 'sick' versus Stieng **ɟiː** 'sick'. For some etyma South Bahnaric have a short closed syllable, and for others they have an open syllable. This suggests a Proto-Bahnaric contrast of length before glottal stop, reflected in Cua as a vowel quality difference. The following examples are suggestive although not conclusive:

Cua	**Bahnar**	**SB**	**Alak**	**PB**
ɟiʔ 'sick'	ɟiʔ	ɟiː (Stieng)	ʤiʔ	*ɟiːʔ
karaʔ 'old'	kraʔ	kraː (Srê)	karaʔ	*kraːʔ
baʔ 'father'	ɓaʔ	baː (Stieng)	--	*ɓaːʔ
suʔ 'to return'	--	seʔ (Stieng)	cɯʔ	*cɨʔ
dɪʔ 'all, finished'	ɗiʔ 'all'	--	--	(*ɗiʔ)

Summary & Conclusions

While the analyses and results presented in this - rather preliminary - paper are far from complete, they have identified many features of the historical-phonological evolution of Cua, and surely provide a solid basis for further work. On the face of it Cua shows a peculiar suit of connected conditioned sound changes that are readily explained as direct developments from Proto-Bahnaric. It would therefore seem reasonable to suggest that Cua represents a (fourth) primary branch of Bahnaric on its own, which (contra Sidwell 2002), I suggest be called "East Bahnaric".

The reconstructed development of PB phonemes discussed here is tabled as follows:

Initials				**Finals**		
*p-	>	p-		*-p	>	-p
*b-, *ɓ-	>	b-		*-m	>	-m/-p
*m-	>	m-		*-w	>	-w
*w-	>	w-		*-t	>	-t
*t-	>	t-		*-n	>	-n/-t
*d-, *ɗ-	>	d-		*-r, -l	>	-l
*n-	>	n-		*-c	>	-n/-t/-k
*r-	>	r-		*-s	>	-lh
*l-	>	l-		*-ɲ-	>	-ŋ/-k
*s-	>	hl-		*-k	>	-k
*c-	>	s-		*h-	>	-h
*ɟ-	>	ɟ-		*ʔ-	>	-ʔ
*ɲ-	>	ɲ-				
*j-	>	j-				
*k-	>	k-				
*g-	>	g-				
*ŋ-	>	ŋ-				
*h-	>	h-				
*ʔ-	>	ʔ-				

Vowels								
*iː	>	iː/eː/ia/i	*ɨː	>	iː/ia/eː	*uː	>	uː/ua
			*əː	>	əː/ɨː			
			*aː	>	aː/ɨː			
*ia	>	ia/ja/iː				*u a	>	ua/ɔː/ wa/uː
*i	>	i/ɪ	*ɨ	>	ɨ/u/ʌ			
			*ə	>	ə/ɨ			
			*a	>	a/ɨ			

REFERENCES

Burton, Eva. 1969. “A brief sketch of Cua clause structure.” *MKS* 3:5-8.

Huffman, Franklin E. 1985. Vowel Permutations in Austroasiatic Languages. *Linguistics of the Sino-Tibetan Area: The State of the Art.* Pacific Linguistics Series C-No.87. Canberra: Australian National University, pp141-45.

Maier & Đình Van Cau. 1976. *Cua Vocabulary: Cua-Vietnamese-English Thesaurus (a computer print-out)*. Huntington Beach California, Summer Institute of Linguistics.

Maier, Jacqueline & E. Burton. 1971. *Vietnam Word List (revised): Cua*. (Copy held by David Thomas Library, Bangkok, folder 495.9).
Maier, Jacqueline. 1969. "Cua phonemes." *MKS* 3:9-19
Maier & Burton. 1981. *Cua-English-Vietnamese Dictionary: 1981 Pre-Publication Manuscript Based on a 1966 Manuscript*. (Held by SIL, Dallas)
Phillips, Richard. 1959. *Vietnam Tribes Survey Word Lists: Cua (=Traw)*. (ms. dated July 1959, copy held by David Thomas Library, Bangkok, folder 495.968)
Sidwell, Paul & Pascale Jacq. 2003. *A Handbook of Comparative Bahnaric: Volume 1, West Bahnaric*. Canberra, Pacific Linguistics 551.
Sidwell, P. 2003. "A note on the reconstruction of proto West Bahnaric and investigation of early West Bahnaric-Katuic contact." *MKS* 33:159-166.
Sidwell, Paul. 2002. "Genetic classification of the Bahnaric languages: a comprehensive review." *MKS* 32:1-24.
Sidwell, Paul. ms. *Proto North Bahnaric: a Reconstruction of Phonology and Lexicon*. (draft issued March 2002)
Smith, Kenneth, D. 1973. "Eastern North Bahnaric: Cua and Kotua." *MKS* 4: 113-118.
Thurgood, Graham. 1999. *From ancient Chamic to modern dialects: Two thousand years of language contact and change.* Oceanic Linguistics Special Publications No. 28. Honolulu: University of Hawai.i Press.

Received: 26 September 2008

Department of Linguistics
Research School of Pacific & Asian Studies
Australian National University, 0200, ACT,
AUSTRALIA
<paulsidwell@yahoo.com>

Tai Loanwords in Mal: A Minority Language of Thailand[1]

CHOMMANAD Intajamornrak
Naresuan University, Thailand

Abstract

Mal is a dialect of Lua/Thin, a Mon-Khmer language spoken in Nan Province of Thailand. Almost all of the speakers of the Mal language are bilingual or multilingual in Mal, Tai Yaun and/or Standard Thai. In a situation of bilingualism or multilingualism, a natural linguistic phenomenon is language contact, which can bring about phonological interference. Mal has borrowed a lot of Tai words for many years. The aim of this paper is to study Tai loanwords in Mal and to classify transformational processes of loanword adaptation in Mal.

The data were taken from a wordlist of 2,452 lexical items complied for investigating the 13 languages of Nan (L-Thongkum et al. 2007). To separate loanwords from native words, loanwords that are easily distinguished by phonological form were separated first. Then ambiguously separated loanwords were checked with the other Tai and Mon-Khmer languages spoken in Nan Province. After that the percentage of loanwords found in the data was calculated.

The results show that about one-third of the lexical items are Tai loanwords. Compound nouns and proper nouns are always composed of Mal native words and loanwords. Transformational processes found in the adaptation of Tai loanwords to Mal can be classified into two main categories; 1) Segmental change and 2) Suprasegmental innovation. Segmental change is the substitution of consonant and vowel phonemes which do not appear in the native language, for example, the phoneme /f/ in the source language is replaced by /ph/; the phoneme /ɨə/ is replaced by /iə/. As for suprasegmental innovation, the low-rising tone, which occurs on about half of the Tai loanwords, must be the result of language contact with a tonal language, especially a language with many words on a rising tone. It seems that the emergence of the low-rising tone in Mal is a borrowed device for marking loanwords. The reason for the lack of the low-rising tone on the other half of the Tai loanwords may be the time-depth of borrowing that caused them to be better integrated with the native phonology.

It can be concluded that transformational processes of loanword adaptation in Mal involve the native Mal phonology as well as the emergence of a new feature that is limited to loanwords. Perhaps the most interesting domain in Tai loanwords discussed above is that of prosody. That is, how does tone function as a symbol of non-nativeness and how can language contact contribute tonogenesis.

[1]Paper presented at the 9th Annual Southeast Asian Studies Graduate Student Conference, March 16-18, 2007 Cornell University, U.S.A.

I. Introduction

Mal is a dialect of Lua/Thin, a Mon-Khmer language spoken in Nan Province of Thailand. It is in the Khmuic sub-branch of the Northern Mon-Khmer branch. According to Filbeck (1978), the Lua/Thin language has two main dialects, Mal and Pray. There are three varieties of Mal that agree in vocabulary and sound change. Mal A is spoken only in one village in Thung Chang District. The other two varieties are spoken in several villages in Pua District. For Pray, there are only two varieties found in Thailand. They share similar vocabulary and sound change with the Mal dialects.

Examples of cognates within Lua/Thin from Filbeck (1978) are given below:

	Mal		**Pray**		
Mal A	Mal B	Mal C	Pray A	Pray B	Gloss
phram	*phjam*	*pham*	*khram*	*khjam*	'person'
prɔːŋ	*pjɔːŋ*	*pɔŋ*	*ŋkrɔː*	*ŋkjɔː*	'morning'

From the examples, we can see that Mal A keeps the proto-sound */r/, but it has been changed to /j/ in Mal B and has disappeared in Mal C. It also occurs in Pray A, but it has been changed to /j/ in Pray B.

Almost all of the speakers of the Mal language are bilingual or multilingual. They can speak Tai Yuan (Kam Muang Nan) fluently as well as their native language, because they usually communicate in it with Nan, Hmong, and Mien people nearby (Smalley 1994: 230). Nowadays, some of them who are educated can also speak Standard Thai very well, especially the young generation. Linguistic research on Mal in Thailand has been done by David Filbeck. One of his works that is of interest to me is about loanwords (Filbeck 1997).

In a situation of bilingualism or multilingualism, a natural linguistic phenomenon is language contact, which brings about phonological interference. It was found that Mal has borrowed a lot of Tai words for many years. The Tai loanwords found in Filbeck's study (Filbeck 1997) caused him to classify the Lua villages into three groups depending on how the words are used and changed.

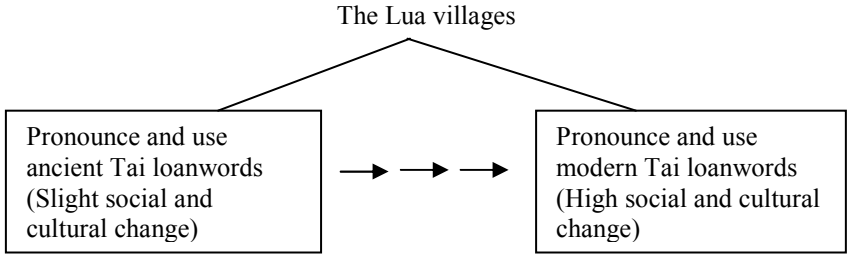


Figure 1. The classification of Pray and Mal villages by their use of Tai loanwords (adapted from Filbeck 1997: 20)

Examples of Tai loanwords having voiceless sonorant as an initial consonant are shown below:

Kwet Village (Mal)	**Wen Village (Pray)**	**Gloss**
hmɔɔ	*mɔɔ*	'doctor'
hlaŋ	*laŋ*	'classifier for house'

It is indeed true that some villages still keep voiceless sonorant sounds which they must have borrowed from ancient Tai words, while the others don't. From my fieldwork at Yot Doi Wattana Village, Bo Kluea Tai Sub-District of Bo Kluea District in Nan Province,[2] I have come to realize that other aspects of the movement of Tai loanwords into Mal may also usefully be studied, in addition to the comparison of the history of Tai and Mal.

Due to the typological differences between the Mon-Khmer and Tai language families, transformational processes of loanword adaptation between two languages of the same family likely differ from those involved in loanwords from another family. Loanword adaptation is a controversial linguistic phenomenon of interest to many phonologists as well as other linguists. Recent studies on loanword adaptation can be found for many languages, such as Thai (Kenstowicz and Suchato 2006), Norwegian (Rice 2006), Mandarin Chinese (Hsieh 2005), Japanese (Dohlus 2005), and Korean (Kenstowicz 2005). The aim of this paper is to study Tai loanwords and to classify transformational processes of loanword adaptation in Mal.

II. Theoretical considerations and previous research

Borrowing is one of the linguistic mechanisms occurring in language-contact situations. It means the incorporation of foreign features into a group's native language by speakers of that language; the native language is

[2]For the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development" directed by Professor Theraphan Luangthongkum.

maintained but is changed by the addition of certain incorporated features (Thomason and Kaufman 1988). In addition, Coetsem (1988) also considers how borrowing begins. Borrowing means features in a source language are transferred to a recipient language by a native speaker of the borrowing language. The component most commonly first transferred is lexicon, that is, lexical borrowing. Then, phonological borrowing, morphological borrowing and syntactic borrowing may also occur when a contact situation is more intense.

Loanword adaptation involves transformations that apply to words when they are borrowed into a foreign language. Two approaches to loanword adaptation have been in dispute as to whether such adaptation is phonetic or phonological (Heffernan 2005, Kenstowicz 2005). Moreover, perception is proposed to have a crucial role in loanword adaptation as well (Peperkamp & Dupoux 2003, Miao 2005).

Kenstowicz and Suchato (2006) study the adaptation of loanwords from English into Thai. Their interpretation of the data is that the adaptation of consonants lacking counterparts in the Thai phonemic inventory, such as the adaptation of [v] to [w], can be explained in terms of auditory similarity, while the change from interdentals to dentals is rather articulatory-based. These authors also discuss the mapping of English voiceless and voiced stops to the Thai three-way, voiceless aspirated, voiceless unaspirated and voiced series. It is observed that Thai speakers assign English word-initial voiceless stops to the Thai aspirated stops. This means that the adaptation is based on surface phonetics. Then the adaptation of prosodic structure and tone is investigated. The repair strategies of vowel lengthening and truncating of final clusters are applied to loanwords.

There has not been as much research on loanwords in Mon-Khmer languages as in some other language families. Berenhult (2001) has studied loanword phonology in Jahai, a language of the Northern Aslian Subgroup of the Mon-Khmer language family, spoken in Malaysia. Jahai speakers have long been in contact with Tamiar, a Central Aslian language, and also with Malay, the Austronesian majority language. He found four main phonological changes in loanword phonology in Jahai: phonetic adaptation, phonemic replacement, reorganization of syllabic structure, and relocation of stress. The crucial point is that the several patterns are the result of indigenous innovation. He also concludes that Jahai loanword phonology is like other Aslian languages with respect to the phonological treatment of Malay loanwords.

In addition to the question of whether the inputs are phonetic or phonemic, various transformational processes are also to be distinguished, namely, segmental and suprasegmental change, deletion, epenthesis, and so on.

Loanwords from English to Thai as examples of segmental and suprasegmental change are the following (Kenstowicz and Suchato 2006).

	English	**Thai**
Segmental	***g****oal*	***k****oo*[32]

	visa	*wiʃ*^{32} *saa*^{42}
Suprasegmental	*camp*	*khɛm*^{34}
	brand	*brɛɛn*^{32}

And loanwords from Malay to Jahai provide examples of deletion and epenthesis (Burenhult 2001).

Deletion	**Malay**	**Jahai**	
	lembu	*l.mu?*	‘cattle’
	rendah	*r.nah*	‘low’
Epenthesis	*ajam*	*hajam*	‘poultry’
	nasi	nasi?	‘cooked rice’

III. Method

This paper will focus on transformational processes of Tai loanwords found in Mal. This means not only Thai, but also other Tai languages spoken in Nan Province. While Thai is the standard language of Thailand, in this setting of language contact, Tai Yuan is the majority language of Nan Province. Lexical borrowing in this situation may be from both Thai and Tai Yuan.

The data was taken from a wordlist of 2,452 lexical items devised for investigating the 13 languages of Nan[3]. This is an output of the research project on “Linguistic Diversity in Nan Province: A Foundation for Tourism Development” sponsored by the Thailand Research Fund. To separate loanwords from native words, loanwords which are easily distinguished by phonological form will be separated first. Then loanwords which are ambiguously separated will be rechecked with the other Tai and Mon-Khmer languages spoken in Nan Province. After that the percentage of loanwords found in the data will be calculated.

Loan-translations (calques) will not be treated in this paper, for example:

Tai Yuan	**Mal**	**Gloss**
kop^{35} *kiŋ*^{23} *diən*^{35}	*teey pɔŋ thuə?*	‘lunar eclipse’
dii^{35} *caj*^{23}	*lɔ?ɲiim*	‘delighted, greatly pleased’

IV. Analysis

A. Tai Phonological systems

Although, Standard Thai and Tai Yuan are in the same language family, there are some phonological differences between them. The first category is the tonal system. Standard Thai has five tones, while Tai Yuan has six tones.

[3]The data were collected in the field by Professor Theraphan Luangthongkum. I am grateful to her for letting me use them for my research.

Word	Gloss	Standard Thai Tone	Tai Yuan (Nan) Tone
khaa	'remain embedded'	32	35
khaa	'galangal'	21	33
khaa	'value'	42	31
khaa	'to trade'	34	43
khaa	'leg'	323	23
khaa	'servant'	42	44'

The second category is the consonants and vowels which are very similar in both languages; however, there are some systematic differences, especially in the consonants. The aspirated sounds in Standard Thai correspond to two sounds in Tai Yuan, aspirated and unaspirated. The correspondences are shown here (adapted from Smalley 1994: 75):

Standard Thai	Tai Yuan	Gloss
phaa^{42}	*phaa*$^{44'}$	'cloth'
thaam^{323}	*thaam*^{23}	'to ask'
chom^{32}	*chom*^{35}	'to admire, to praise'
khaa^{323}	*khaa*^{23}	'leg'
phan^{32}	*pan*^{35}	'thousand'
thaaŋ^{32}	*taaŋ*^{35}	'road'
chi̶̶^{42}	*ci̶̶*^{31}	'name'
kham^{32}	*kam*^{35}	'word'

Also, the phoneme /j/ in Standard Thai corresponds to two consonants, [j] and [ɲ], in Tai Yuan; while the latter does not occur in Standard Thai, for example:

Standard Thai	Tai Yuan	Gloss
juu^{21}	*juu*^{33}	'be at'
jaaw^{32}	*ɲaaw*^{35}	'long'

The other consonantal difference between these two languages is the correspondence between Standard Thai /r/ and consonants of Tai Yuan. To begin with, the stability of the contrast between the traditionally accepted consonants /l/ and /r/ in Standard Thai is rather dubious. In colloquial Standard Thai, there is much vacillation between the two. Here, however, with this situation in mind, we can note examples of words in Tai Yuan that are cognates of words conventionally viewed as beginning with /r/ in Standard Thai.

Standard Thai	Tai Yuan	Gloss
riə^{32} ~ *liə*^{32}	*hiə*^{35}	'boat'
rɔɔj^{34} ~ *lɔɔj*^{34}	*lɔɔj*^{43}	'hundred'

Moreover, Tai Yuan doesn't have consonant clusters with /l/ and /r/ as in Standard Thai.

Standard Thai	Tai Yuan	Gloss
khrɔɔp^{42}-*khruə*^{32}	*khɔɔp*^{31}-*khuə*^{35}	'family'
plaa^{32}	*paa*^{23}	'fish'

B. Mal Phonological system

The phonological system of Mal, the Yot Doi Wattana dialect, is not too complicated compared to other Mon-Khmer languages, especially in final consonants and vowels. There are 28 consonant phonemes of which only 11 can be final consonants. The consonant phoneme inventory is shown in Table 1.

Table 1. The consonant system of Mal (Initials)

	Bilabial	**Alveolar**	**Alveolo -Palatal**	**Palatal**	**Velar**	**Glottal**
Stops	p ph	t th			k kh	ʔ
	b	d			g	
Nasals	m m̥	n n̥		ɲ ɲ̥	ŋ ŋ̥	
Fricatives		s				h
Affricates			c ɟ			
Laterals		l l̥				
Approximants	w w̥			j j̥		

Final consonants: -p, -t, -k, -ʔ, -h, -m, -n, -ŋ, -l, -w, -j

The vowel system of Mal distinguishes only three degrees of height in front, central, and back position. Besides, length has phonemic significance. Table 2 shows the vowel phoneme inventory in Mal.

Table 2. The vowel system of Mal

	Short vowels			**Long vowels**		
	Front	**Central**	**Back**	**Front**	**Central**	**Back**
High	i	ɨ	u	ii	ɨɨ	uu
Mid	e	ə	o	ee	əə	oo
Low	ɛ	a	ɔ	ɛɛ	aa	ɔɔ

Diphthongs with a centering offglide iə uə

Mal Yot Doi Wattana dialect has two tones, high tone and low tone. The low tone is mainly limited to Tai loanwords (LW)[4]. For example:

kɔ̆ɔt (LW)	'to embrace'
khăam (LW)	'to cross over'
bĕɛk (LW)	'to carry on the shoulder'
păan (LW)	'birthmark'
lĕɛk (LW)	'to exchange'

[4]In some Mal non-tonal dialects which are non-tonal language, lexical pitch can vary according to syllable structures: high-falling in non-checked syllable and high-level in checked syllable. More details have been discussed in Intajamornrak (2009).

C. Transformational processes

Since Mal has been spoken in this area for a long time and has been in contact with Tai Yuan, the majority language, Standard Thai, as well as other minority languages, it is not surprising that many loanwords can be found.

Out of my corpus of 2,452 words, 813 or about one-third are loanwords. Some words, especially compound nouns, are composed of Mal native words (M) and loanwords (LW). This construction commonly occurs in Southeast Asian languages, for example:

siəŋ (M)	*khăaŋ* (LW)	'molar teeth'
phŭn (LW)	*puəh* (M)	'ashes'
khɔ̆ɔŋ (LW)	*pɔŋ* (M)	'condiments eaten with rice'
thɔɔŋ (LW)	*soo* (M)	'copper'

Some compound nouns contain a native Mal word to designate a category followed by a loanword to specify the member of the category. Some categories are /thuu/ for vegetable, /phlɛʔ/ for fruit, /jaaŋ/ for flower, /lam/ for tree, and /phjam/ for person. For example:

thuu (M) *ka-phaw* (LW)	'sweet basil (kind of)'
phlɛʔ (M) *mŭəŋ* (LW)	'mango'
jaaŋ (M) *ʔiəŋ* (LW)	'orchid'
lam (M) *təəj* (LW)	'Pandanus amaryllifolius Roxb (kind of plant)'
phjam (M) *cĭin* (LW)	'Chinese person'

Some Tai loanwords can be described in terms of semantic shift. It means that the meanings of the loanwords are related to those of the words of the source language but are somewhat different in Mal. For example:

Standard Thai	**Gloss**	**Mal**	**Gloss**
hoo^{21}	'to cheer'	*hŏo*	'to shout'
leew^{32}, *waj*^{32}	'fast, quick'	*leew, waj*	'to hurry'
phun^{21}	'dust'	*phŭn*	'fine particles'
liəŋ^{323}	'yellow'	*l̥iəŋ*	'pale'
kwaj^{32}	'to swing'	*kwăj*	'to wave'

Words borrowed from Tai become adapted to the native phonological system of Mal. The transformational processes fall into two main categories: segmental change and suprasegmental innovation.

1. Segmental change

A. Tai phonemes with no Mal equivalents

In Tai loanwords the consonant /f/, which is not found in Mal, is replaced by /ph/ or /phw/. For example:

Standard Thai	Tai Yuan	Mal	Gloss
fa-raŋ^{21}	*khoŋ*^{33} *fa-laŋ*^{33}	*phjam pha-lăŋ*	‘westerner’
fuŋ^{21}	*khii*$^{44'}$ *fuŋ*^{33}	*phŭn*	‘dust’
fiŋ^{21}	*fiŋ*^{33}	*phĭn*	‘opium’
faaj^{42}	*faaj*$^{44'}$	*phwaaj*	‘cotton’
faak^{21}	*faak*^{33}	*phwăak*	‘to entrust’

As for the vowel phonemes, the diphthong with a centering offglide /iə/ does not occur in Mal and in Tai loans takes a Mal /iə/ as in the following examples:

Standard Thai	Tai Yuan	Mal	Gloss
liək^{42}	*liək*^{31}	*lĭək*	‘to choose’
liəŋ^{323}	*liəŋ*^{23}	*l̥iəŋ*	‘yellow’
ʔiəŋ^{42}	*ʔiəŋ*^{43}	*jaaŋ ʔiəŋ*	‘wild orchid’
chiə^{42}	*ciə*^{31}	*ciə*	‘to believe’
thiən^{21}	*thiən*^{33}	*thĭən*	‘illegal, unauthorized’

B. The same distinction in Tai and Mal

The most interesting case is that the phonological differences between the source languages, Standard Thai and Tai Yuan, make the adaptation process of loanwords more complicated. The stop consonant phonemes take into account this process. Both the source languages and the native language have three-way distinctions of stop consonants, but the phonemes /ph/, /th/, /kh/ in Standard Thai correspond to /p/ - /ph/, /t/ - /th/, and /k/ - /kh/ respectively in Tai Yuan cognate words.

The first case is that in which the phonemes of the Tai source languages are the same: /p/, /t/ and /k/. In this case, it is interesting that the Mal versions can be sometimes adapted to the sounds [ph], [th], and [kh] respectively. For example:

Standard Thai	Tai Yuan	Mal	Gloss
pet^{21}	*pet*^{35}	*pet*	‘duck’
piŋ^{42}	*piŋ*$^{44'}$	*phiiŋ*	‘to roast, to barbecue’
tɔɔn^{32}	*tɔɔn*^{23}	*tŏɔn*	‘to castrate’
tuəŋ^{32}	*tuəŋ*^{35}	*thŭəŋ*	‘to measure’
kut^{21}	*kut*^{35}	*kut*	‘cut off’
kooŋ^{32}	*kooŋ*^{33}	*khooŋ*	‘humpbacked’

On the other hand, if the phonemes in the source languages are different, varied substitutions can also be found, especially for the phoneme /kh/. For example:

Standard Thai	Tai Yuan	Mal	Gloss
phoŋ^{34}	*poŋ*^{43}	*pŏn*	'free from (danger, misfortune)'
phaŋ^{32}	*paŋ*^{35}	*phan*	'thousand'
phaaj^{32}	*phaaj*^{35}	*phaaj*	'to paddle'
khap^{34}	*kap*^{44}	*kap*	'tight'
khaŋ^{42}	*kaŋ*$^{44'}$	*khăn*	'to separate'
khuə^{42}	*khuə*$^{44'}$	*khuə*	'to parch, to roast'
khɔɔ^{323}	*khɔɔ*^{23}	*khjɔɔ*	'to ask for, to beg'

If the Tai Yuan words begin with /l/, there is no adaptation; the Mal word also begins with /l/. Tai Yuan initial /h/ corresponds to Mal /l/ or /j/.

Standard Thai	Tai Yuan	Mal	Gloss
rɔɔj^{34}	*lɔɔj*^{43}	*lɔɔj*	'hundred'
rəəm^{42}	*ləəm*^{31}	*lăəm*	'to start'
rak^{34}	*hak*^{44}	*jak*	'to love'
ruup^{42}	*huup*^{31}	*jŭup*	'picture, photography'
riən^{32}	*hiəŋ*^{35}	*liən*	'to study'
riə^{32}	*hiə*^{35}	*liə*	'boat'

The same process occurs with the phoneme /j/ in Standard Thai which corresponds to [j] and [ɲ] in Tai Yuan. This means that loanwords having [j] in both Standard Thai and Tai Yuan still keep the same consonant without adaptation, while loanwords having /j/ in Standard Thai but /ɲ/ in Tai Yuan keep the same sound as in Tai Yuan.

Standard Thai	Tai Yuan	Mal	Gloss
jɨt^{34}	*jɨt*^{44}	*jɨt*	'to hold back'
jɨɨm^{32}	*jɨɨm*^{35}	*jɨɨm*	'to borrow'
jaak^{42}	*ɲaak*^{31}	*ɲăak*	'difficult'
jɔɔm^{32}	*ɲɔɔm*^{35}	*ɲɔɔm*	'consent'
jok^{34}	*ɲok*^{44}	*ɲɔk*	'to lift, to rise'

The next segmental change deals with the vowel phonemes. The monophthongal vowels in Mal are exactly the same as in Standard Thai and Tai Yuan, but they are also adapted inconsistently, for example, the substitution of [ɛ] for [a], [o] for [ɔ], [ə] for [o], and [e] for [ɛ].

Tai Yuan	Mal	Gloss
miit^{33} *ɲap*^{44}	*ma-ɲɛp*	'scissor'
naŋ^{35} *wɔŋ*^{43}	*jaaŋ woŋ*	'rubber band'
ʔok^{35}	*ʔək*	'chest'
cɛɛ^{31}	*cĕe*	'to soak'

However, the outstanding vowel change is vowel lengthening, especially when the words are borrowed from Tai Yuan.

Tai Yuan	Mal	Gloss
wɛn^{31}	*wĕɛn*	'mirror'
ma-nun^{23}	*phlɛʔ nŭun*	'jack fruit'
kiw^{43}	*kiiw mat*	'eyebrow'
ŋɔj^{31}	*ŋɔɔj*	'lame'
nɛŋ^{31}	*nĕɛŋ*	'net for catching fish'

C. A more complex case

For the affricate phoneme /ch/, it was shown above that Standard Thai /ch/ also corresponds with two consonants in Tai Yuan, /ch/ and /c/. If the words of Standard Thai and Tai Yuan both have /ch/, it will be replaced in Mal by /s/, while the words with /ch/ in Standard Thai but /c/ in Tai Yuan have in Mal the same consonant as Tai Yuan.

Standard Thai	Tai Yuan	Mal	Gloss
chom^{32}	*chom*^{35}	*som*	'to admire'
chaam^{32} *ʔaaŋ*^{21}	*chaam*^{35}	*saam sək*	'basin'
chaŋ^{42}	*caŋ*^{31}	*căŋ*	'to weigh'
chuu^{34}	*cuu*^{43}	*cuu*	'lover'
chaw^{42}	*caw*^{31}	*căw*	'to rent'

2. Suprasegmental innovation

The low tone is an innovation found mostly in Tai loanwords. It is surely the result of contact with a tonal language, especially a language with many words on a rising tone. It seems likely that the emergence of the low tone in Mal is a borrowed device for marking loanwords.

Out of 813 Tai loanwords, 346 (42.5%) have the low tone. It was found that 654 words of Tai Yuan origin (80.4%) carry the rising tone, and 95 words of Standard Thai origin (11.7%) carry the rising tone. As a result, when Mal speakers hear Tai words, the most striking feature is the rising pitch on so many words, especially in Tai Yuan, which has two rising tones. Apparently, it causes them to pronounce Tai words with the low-rising tone regardless of what tone the original words have. For example:

Standard Thai	Tai Yuan	Mal	Gloss
ta-khɔɔ^{323}	*khɔɔ*^{23}	*khŏɔ*	'hook'
klɔɔŋ^{32}	*kɔɔŋ*^{23}	*kŏɔŋ*	'drum'
khaap^{42}	*kaap*^{31}	*kăap*	'to seize (in the teeth or beak)'
taŋ^{42} *tɛɛ*^{21}	*taŋ*$^{44'}$ *tɛɛ*^{33}	*tăŋ*	'from'
phon^{34}	*pon*^{43}	*pŏn*	'free from'

Most Mon-Khmer languages are classified typologically as non-tonal. The one prominent exception is Vietnamese, which is clearly tonal. Proto-Thin has been reconstructed as a non-tonal language (Filbeck 1978). It means that Mal developed tone later on. Filbeck (1972) suggested that contact with Tai

might be a cause of this linguistic phenomenon. However, he did not explicitly state why it had to have rising pitch no matter what tone that loanword had in the source language.

In trying to find the causes of this tonal evolution, Filbeck looked for internal factors and considered the possibility of a mechanism for marking loanwords; however, he feels that the failure of loanwords to be consistent in taking the rising tone in Mal makes the second factor unlikely. In my view, the reason for the lack of the low-rising tone on the other half of the Tai loanwords may be the time-depth of borrowing that caused them to be better integrated with the native phonology. A good explanation is found in the loanwords having voiceless sonorants, which don't carry the low-rising tone. For example:

Standard Thai	**Tai Yuan**	**Mal**	**Gloss**
phɔɔ^{42} *maaj*^{42}	*pɔɔ*^{31} *maaj*$^{44'}$	*ʔaw m̥aaj*	'widower'
mɔɔŋ^{323}	-	*m̥ɔɔŋ*	'gloomy, depressed'
nep^{21}	*nep*^{35}	*hep*	'to stick in'
lop^{21}	*lop*^{35}	*l̥op*	'to avoid'
liə^{323}	*liə*^{23}	*l̥iə*	'remaining, left over'

The adaptation of these words with Mal voiceless sonorants or /h/ strongly suggests contact with very early stages of Tai, which has been reconstructed by Li (1977) as having voiceless sonorants.

V. Discussion and Conclusion

The transformational processes of loanword adaptation found in Mal can be classified into two categories: segmental change and suprasegmental innovation. Although Standard Thai is the national language, Tai Yuan, which is the majority language of Nan Province, seems to have had much more influence on Mal.

It has been accepted that loanwords are adapted by undergoing phonological reanalysis in which the native sounds repair the illegal non-native forms to agree with the native phonology. In some cases, more than one sound can be substituted. It means that there is more than one possibility to replace a non-native segment by a native one by single feature change (Peperkamp and Dupoux 2003), as in the adding of aspiration to unaspirated stop consonants. Moreover, unnecessary adaptations occur during the transformation of Tai loanwords. Peperkamp (to appear) points out that unnecessary adaptations might be called generalizations since they apply to foreign forms that are well-formed in the native (borrowing) language but do not conform to some default pattern. For example, Tai and Mal have the same monophthongs, nevertheless, Mal is inconsistent in its adaptation of Tai loanwords with monophthongs.

We may conclude then that loanword adaptations in Mal involve the native Mal phonology as well as the emergence of a new feature that is predominantly limited to loanwords. Perhaps the most interesting domain in Tai loanwords discussed above is that of prosody. That is, how does tone function as a symbol of non-nativeness and how can language contact

contribute tonogenesis. It remains to be seen whether a full tonal system will come into being over time or whether the present incipient system will die out.

VI. Acknowledgement

I would like to express my gratitude to Naresuan University and The Royal Golden Jubilee Ph.D. Program for sponsoring my research scholarship (Grant No. PHD0270/2546), and The Thailand Research Fund for sponsoring the research project on "Linguistic Diversity in Nan Province: A Foundation for Tourism Development". I also would like to express my sincerest appreciation to Professor Theraphan Luangthongkum, my advisor, and Professor Arthur S. Abramson for their kind support in my research.

REFERENCES

Berenhult, N. 2001. "Loanword phonology in Jahai." Lund University, Department of Linguistics, *Working Papers* 48(2001):5-14.

Coetsem, F. V. 1988. *Loan Phonology and the two Transfer Types in Language Contact.* Dordrecht: Foris.

Dohlus, K. 2005. "Phonetics or Phonology: Asymmetries in Loanword Adaptations - French and German Mid Front Rounded Vowels in Japanese." Zentrum für Allgemeine Sprachwissenschaft. *Papers in Linguistics* 42:117-135.

Filbeck, D. 1972. "Tone in a dialect of T'in." *AL* 14: 111-118.

Filbeck, D. 1978. "T'in: a historical study." *PL* B:49. Canberra: Australian National University.

Filbeck, D. 1997. "Thai loanwords in the Lua language of Muang Nan." Seminar document paper for the conference *Suepsan Adit An Rueangrong Khong Mueang Nan: Khomun Lae Mummong Mai Thang Borankhadi: Prawattisat Lae Chattiphan.* Tai Studies Project, Payap University and Nan Community, December 18-21, 1997.

Heffernan, K. 2005. "Phonetic similarity and phonemic contrast in loanword adaptation." *Toronto Working Papers in Linguistics* 24:117-123.

Hsieh, C. S. 2005. "Taiwanese Loanwords in Mandarin Chinese: Language Interaction in Taiwan." *Taiwan Papers 5*: October 2005. [available at http://www.ruhr-uni-bochum.de/slc/ taiwan/TaiwanPapers5.pdf]

Intajamornrak, C. 2009. "On becoming a tonal language of Mal and Pray spoken in Nan Province." Ph.D. dissertation, Department of Linguistics, Faculty of Arts, Chulalongkorn University. (in Thai)

Kenstowicz, M. 2005. "The phonetics and phonology of Korean loanword adaptation." Paper presented at *First European Conference on Korean Linguistics,* Leiden University, February 2005.

Kenstowicz, M. and A. Suchato. 2006. "Issues in loanword adaptation: A case study from Thai." *Lingua* 116(7):921-949.

L-Thongkum, T. 1984. "Tonal evolution and tonal development in Mon-Khmer." *Warasarn Aksornsart* 16(2):53-72. (in Thai)

L-Thongkum, T. and C. Intajamornrak. 2009. "Tonal evolution induced by language contact: a case study of the T'in (Lua') language of Nan province, Northern Thailand." *MKS* (38):57-68.

L-Thongkum, T. and S. Puengpa. 1980. *Bru-Thai-English Dictionary. Department of Linguistics.* Chulalongkorn University.

L-Thongkum, T., P. Akharawatthanakun, C. Intajamornrak, K. Putthasatien, Y. Huadsiri and S. Phalipat. 2007. *Dictionary of the 15 Languages Spoken in Nan Province.* Unpublished research report submitted to the Thailand Research Fund (TRF), Bangkok.

Li, F-K. 1977. *A Handbook of Comparative Tai.* Honolulu: The University Press of Hawaii.

Miao, R. 2005. "Loanword Adaptation in Mandarin Chinese: Perceptual, phonological and sociolinguistic factors." Ph.D. dissertation, Department of Linguistics. State University of New York at Stony Brook.

Peperkamp, S. (to appear). "A psycholinguistic theory of loanword adaptations." *Proceedings of the 30th Annual Meeting of the Berkeley Linguistic Society.*

Peperkamp, S. and E. Dupoux. 2003. "Reinterpreting loanword adaptations: the role of perception. Barcelona: Universitat Autonoma de Barcelona." *Proceedings of the 15th International Congress of Phonetic Sciences,* 367-370.

Rice, C. 2006. "Norwegian stress and quantity: The implications of loanwords." *Lingua* 116(7):1171-1194.

Smalley, W. A. 1994. *Linguistic Diversity and National Unity.* Chicago and London: The University of Chicago Press.

Singnoi, U. 1988. "A comparative study of Pray and Mal phonology." M.A. thesis, Institute of Language and Culture for Rural Development, Mahidol University.

Thomason, S. G. and T. Kaufman. 1988. *Language contact, Creolization and Genetic Linguistics.* Los Angeles: University of California Press.

Received: 24 November 2008

Department of Language and Folklore,
Faculty of Humanities, Naresuan University,
Phitsanulok 65000
THAILAND
<katechommanad@gmail.com>

Establishing relative chronology of Palaung sound changes using Tai Loanwords*

PITTAYAWAT Pittayaporn
Chulalongkorn University, Thailand

Abstract

The sequence in which different sound changes occur in a language can be established by identifying feeding/bleeding relationship among the changes. In many cases, however, it is not possible to establish the relative chronology among certain changes because they are not in either feeding or bleeding relationships. The chronology of changes from Proto-Palaung to the Red Palaung presents such a problem. While there is clear evidence for voicing flip-flop and vowel shift, the sequence in which the changes occur is not recoverable from Paluang-internal evidence. Fortunately, Red Palaung has a large number of Tai loanwords, some of which reflect earlier stages of the Shan language. Because a set of loanwords may have been incorporated into Red Palaung before one change but after another, conclusions about the relative chronology among those sound changes can be drawn by applying the principle of feeding/bleeding relationships to those Tai loanwords. In this paper, I argue that Tai loanwords indicate two different series of vowel shifts in Red Palaung intervened by the voicing flip-flop. Moreover, I use these Tai loanwords to locate the three sound changes chronologically with reference, and provide tentative dates for the PR sound changes.

1. Introduction

The historical sequence in which different sound changes occur in a language can be established by identifying feeding/bleeding relationship among the changes. In many cases, however, it is not possible to establish the relative chronology among certain changes because they are neither in feeding nor bleeding relationships. The chronology of changes from Proto-Palaung (PP)[1] as reconstructed by Mitani (1977; 1979)[2] and refined by Diffloth (1988) to the Red Palaung dialect of Pang Daeng Nai (Pittayaporn 2002), henceforth Red Palaung (RP), presents such a problem. While there is clear evidence for vowel changes

*This paper is homage to my teacher Prof. Theraphan L-Thongkum, who inspired me to do linguistic fieldwork, and from whom I first learned linguistic field methods. I would like to thank her for her encouragement and comments. I would like to express my appreciation for members of the Palaung field trip for the wonderful memory: Por, Ple, Koy, and Sophia. Last but not least, I also would like to thank Prof. John Whitman, two anonymous reviewers, and the audience at Chulalongkorn Linguistics Forum for their comments on earlier versions of this paper.

[1]The label "Proto-Palaung" refers to the ancestor language of modern Palaung dialects while "Proto-Palaungic" refers to the parent of a deeper-level subgroup which consists of Palaung dialects and their closest kins Riang, Lamet, Angku, Wa, Blang, etc.

[2]PP forms in Mitani (1977) are re-transcribed with standard IPA symbols.

(see Table 1), the exact changes that brought about the observed reflexes are not recoverable from RP-internal evidence.

Table 1. RP and KP reflexes of PP front and back vowels in open syllables

	PP	**RP**	**KP**	
*-i	*bri:	prɛ:	praj[3]	'forest'
*-e	*ple:	ble:	ble	'fruit'
*-ɛ	*klɛ:	gla:j	glaj	'rain'
*-aj	*rawaj	rəwa:j	rəwaj	'tiger'
*-u	*blu	plɔ:	plaw	'thigh'
*-o	*mo	ma:w	maw	'stone'
*-ɔ	*^{h}ŋɔ	^{h}ŋo:	^{h}ŋo	'paddy'

This problem becomes even more apparent when RP is compared with Kengtung Palaung (KP)[4], its closest kin among the Palaung dialects studied in Mitani (1977). As shown in Table 1, the two varieties both show a voicing flip-flop as evidenced in the cases of 'fruit', 'rain', and 'thigh'. In addition, they both show almost identical vowel reflexes as illustrated 'fruits', 'rain' and 'stone'. However, there are discrepancies between the two dialects with respect to the development of vowels in open syllables as substantiated by 'forest', and 'thigh.' The reflexes of PP */-i/, and */-u/ in RP are /-ɛ/, and /-ɔ/ respectively. In contrast, they are reflected in KP as /aj/ and /-aw/[5]. These discrepancies raise the question of what sequence of changes led to the observed situation.

Fortunately, there are a large number of Tai loanwords in Palaung due to its extensive contact with various Tai groups throughout its attested history. Some of these forms reflect earlier stages of Shan as exemplified in the upper half Table 2. The relevant Proto-Southwestern Tai (PSWT) forms are also provided for comparison. Also see Appendix for a list of Tai loanwords in the RP dialect of Pang Daeng Nai.

Table 2. Examples of Tai loanwords in RP

PSWT	**Shan**	**RP**	
*^{h}laj^{A}	laj^{1}	^{h}lɛ:	'to flow'
*k^{h}uə^{B}	k^{h}o^{3}	k^{h}o	'to fry'
*rɯə^{A}	hɤ^{2}	rɤ:	'boat'
*caw^{C}	caw^{3}	ɟɔ:	'prince'
*tɛ:m^{C}	tɛm^{3}	dɛ:m	'to write'
*^{h}rɔ:k^{D}	hɔk^{2}	hɔ:k	'lance'

[3]It is possible that length is contrastive in KP but it is not indicated in the data.

[4]Mitani (1977) refers to this variety as Darang, obviously a rough transcription of /ˈda:rəˈʔa:ŋ/. His data are from Scott and Hardiman (1900).

[5]Transcribed as <au> or <ao> in the original source.

Because some loanwords may have been incorporated into RP before one change but after another, conclusions about the relative chronology among Palaung sound changes can also be drawn by applying the principle of feeding/bleeding relationships to those Tai loanwords. RP /do:/ 'figure, animal', which goes back to Shan /to^{A1}/ (< */tuə^{A}/) is an excellent example. The form must have been borrowed before the "voicing flip-flop" which transformed as final results the original voiceless unaspirated stops into voiced stops and voiced stops into their voiceless unaspirated counterparts[6]. This is because the initial consonant */t-/ became voiced just like native words. In contrast, it must have not entered RP in time for the shift of PP */-o/ to */-ɔ/ because the RP reflex would be */da:w/ from an intermediate stage */dɔ:/ otherwise.

In this paper, I focus on changes that Palaung vowels in open syllables have gone through. I identify changes that occurred during the development from PP to RP and propose a relative chronology of these changes. By considering Palaung-internal evidence as well as Tai loanwords in RP, I hypothesize that there were four changes in the history of RP that affected PP open rimes: 1) diphthongization of high vowels, 2) raising of low vowels, 3) diphthongization of mid vowels, and 4) diphthong warping, presented in chronological order.

2. Introduction to Palaung and its speakers

According to reports on the Palaung (Chun 1989; Howard and Wattanapun 2001; Sila n.d.; Yangderm 1995), Palaung dialects are spoken in Yunnan Province in China and the Southern and Eastern States of Burma, mainly Shan State. The Burmese call the speakers of this language "Palaung" while the Shan call them "Kunloi" or "mountain people". In China, they are officially recognized as a distinct nationality under the label "De'ang" (德昂). Palaung speakers, however, refer to themselves variously as /daˈʔa:ŋ/, /diˈʔa:ŋ/, /raˈʔa:ŋ/ or /ˈda: rəˈʔa:ŋ/. In Shan State, they live in compact villages, hill tops, or ridges while their Tai neighbors are plain settlers. They are well known for their profession as ancient tea growers although they are also engaged with other forms of agriculture and trade. They have established a centuries-long relationship with the Shan,[7] from whom they acquire cloth and jewelry as well as other products. The Palaung are devout Buddhists even though animism is still practiced to a considerable degree.

The RP dialect used in this study is that spoken in Pang Daeng Nai Village, Chiang Dao District, Chiang Mai Province, Thailand. The Palaung of Pang Daeng Nai migrated from Burma in the 1980's because of civil wars in their home country. Before settling in Northern Thailand, they lived near a hill

[6]The RP flip-flop is best characterized as a series of changes in phonation-types, PP */t-/ > */ɗ-/ > */d-/ and PP */d-/ > */t-/ (Diffloth 1988). Although the crucial steps within this apparent process of flip-flop are the implosivization of voiceless stops and the devoicing of their unaspirated counterparts, the term "voicing flip-flop" is used in this paper to refer collectively to these two related changes.

[7]Shan is used here as a cover term for Tai speaking groups in Burma.

near Kengtung called /lɔ:j la:j/ (corresponding to /dɔ:j1 la:j1/ in Thai). This group of Palaung is sometimes called Pale or Silver Palaung in the literature (e.g. Howard and Wattanapun 2001) but they, however, refer to themselves as Red Palaung (/ˈda: rəˈʔa:ŋ reɲ/).

The data analyzed in this study is collected mainly from Mr. Chang Kana, 80, and Mr. Namsaeng Changmueang, 55, with help from Mr. No Lungsoi who worked as a language assistant. Both informants speak fluent Shan (or Tai Long, henceforth TL) and a little Northern Thai (or Tai Yuan) in addition to their native Red Palaung dialect (henceforth RP). This dialect, spoken in a different village, has been described by Kasisopa (2003) but the phonological analysis differs markedly from this analysis with respect to the vocalisms and the final consonants.

A prosodic word in RP is either monosyllabic or sesquisyllabic[8]. There is relatively little difficulty in establishing the system of initials while the vowels and final inventories are more problematic. With respect to RP vowels, the major problem is whether vowel length is contrastive for vowel pairs other than /a/ and /a:/. In RP, vowel length seems to be predictable from the syllable structure. Vowels in open syllable are always pronounced long and those followed by stops are pronounced short. However, the presence of two separate sets of final nasals seems to be related to the question of vowel length.

In RP, there are two series of final nasal consonants— plain nasals and stops with nasal plosion. Kasisopa (2003) views the latter as realizations of final voiced consonants contrasting with the plain nasals. I differ from her in considering the two series as allophones of the same sounds conditioned by contrastive length of the preceding vowels. In this analysis (Pittayaporn n.d.), final nasals preceded by a long vowel are realized as plain while ones following a short vowel are pronounced as voiceless stop with nasal plosion, cf. [kʰu:n] for /kʰu:n/ ‘wind’ and [kʰutⁿ] for /kʰun/ ‘lord.’ The RP phoneme inventory is given in Table 3. Also note that all RP vowels are always long in open syllables. They show off-glides when not followed by final consonants, except for /a:/.

[8]A sesquisyllable is regarded as consisting of one and a half syllable. For discussions on sesquisyllabicity, see Matisoff (1973).

Table 3. RP phoneme inventory

	labial	**alveolar**	**palatal**	**velar**	**glottal**
stops	p pʰ b	t tʰ d	c cʰ ɟ	k kʰ g	ʔ
nasals	ʰm m	ʰn n	ʰɲ ɲ	ʰŋ ŋ	
fricatives	f	s			h
liquids		ʰr r ʰl l			
glides	w		ʰj j		

	front	**back**	
	unrounded	**unrounded**	**rounded**
high	i, iː	ɯ, ɯː	u, uː
mid	e, eː	ɤ, ɤː	o, oː
low	ɛ, ɛː	a, aː	ɔ, ɔː

Palaung varieties have been in contact mainly with four modern Tai varieties, 1) Tai Luang, 2) Tai Khuen, 3) Tai Yuan, and 4) Siamese Thai. However, the most likely donor is Tai Luang or Southern Shan (TL), which is the lingua franca in Shan States[9]. All these varieties belong to the Southwestern branch of the Tai family according to Li (1977)'s classification. Proto-Southwestern Tai (PSWT) has been reconstructed by various authors (e.g. Jonsson 1991; Li 1977) but the reconstruction used in this study is that proposed by Pittayaporn (2008). A number of Tai loans in RP show features reconstructed for PSWT but not attested in any modern Tai languages.

3. Historical background of RP

According to Diffloth (1988), the RP dialect belongs to the Palaungic branch of the Austroasiatic family as schematized in

Figure *1*. Within the Palaung group itself, Mitani (1977) gives a tentative classification of Palaung, in which modern dialects are grouped into four branches: Central Palaung, Northern Palaung, Southern Palaung, and Omachawn[10]. In this schema, KP belongs to the Southern branch of the Palaung group. Therefore, it is safe to assume that this RP dialect is also a Southern Palaung variety.

[9]TL data used in this paper are from Hudak (1994).
[10]He also notes two unclassified dialects separately from these four groups.

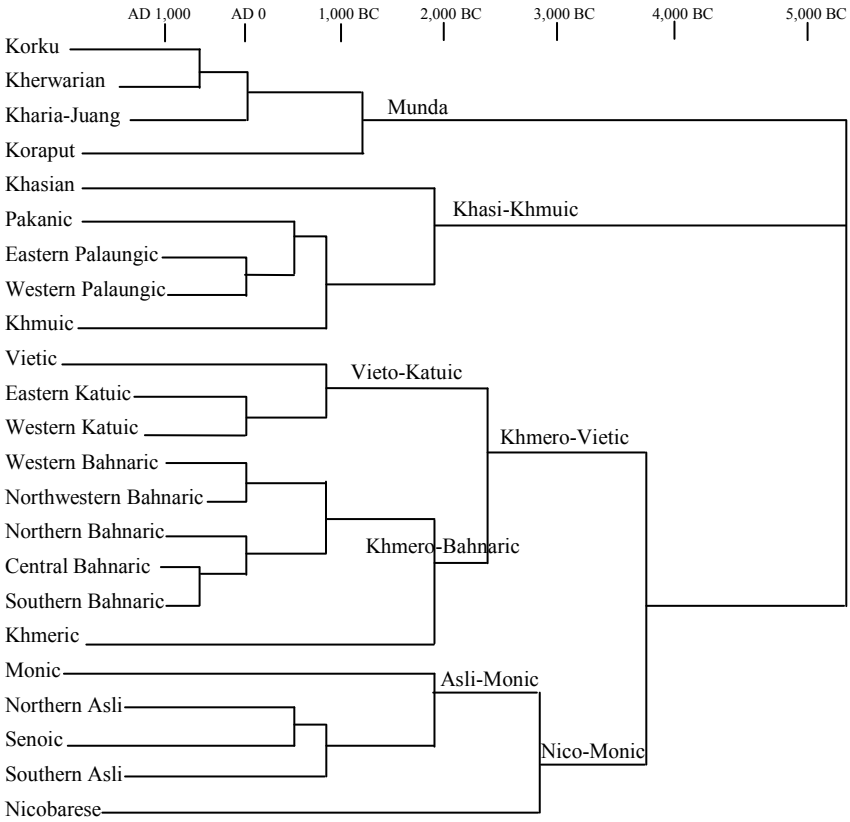


Figure 1. Austroasiatic family tree (from Diffloth 2005)

Important to studies of loanwords are earlier stages of the languages in question. Mitani (1977) reconstructs Proto-Palaung using data from varieties spoken in Myanmar. Diffloth (1988) later refined Mitani's reconstruction by proposing that RP non-high vowels before */-n/ were always long. As far as vowels in open syllables are concerned, I propose that Mitani's */-e/ be reconstructed as */-ɛ/ and his */-ɛ/ as */-e/ as shown in Table 4.

Table 4. Reflexes of Proto-Palaung (PP) in Red Palaung (RP) and Kengtung Palaung (KP)

		Mitani's PP	**RP**	**KP**	**Riang**
*-ɛ	'wood' 'fruit' 'witch'	*he *ple *bre	he: ble: -	he - -	plɛʔ[11] $k^hɛʔ$ prɛʔ
*-e	'earth' 'new' 'rain'	*kətɛ *kənmɛ *klɛ	kəda:j kəma:j gla:j	kədaj - glaj	kəteʔ tənmeʔ (klɛʔ)

Mitani reconstructs */-e/ in such etyma as 'wood', 'fruit', and 'witch' and */-ɛ/ in 'earth', 'new' and 'rain' possibly because Ta-ang, his representative dialect for the Northern Palaung branch, shows /-e/ for the first set and /-ɛ/ for the second. However, in RP and KP Mitani's */-e/ patterns diachronically with his /*ɔ/ but his */-ɛ/ patterns with his /*-o/. This suggests that in fact Ta-ang may have been the dialect that innovated by changing PP */-e/ into /-ɛ/ and */-ɛ/ into /-e/. Extra-Palaung evidence seems to show that this alternative reconstruction is to be preferred. Riang varieties included in Mitani (1977; 1979) show /-ɛ/ in 'wood', 'fruit', and 'witch' and /-e/ in 'earth', 'new' and 'rain,' suggesting that Proto-Palaung Riang */-e/ and */-ɛ/ are retained both in Riang as well as RP and KP. Mitani (1977) notes that his PP */-ɛ/ became Ta-ang /-e/ in many cases but remains agnostic about the conditioning environment. In the alternative proposed here, /-ɛ/ would be the normal reflex of PP */-ɛ/. Etyma which shows /-e/ in Ta-ang for PP */-ɛ/ would be the innovative cases.

RP and KP differ markedly from PP due to phonological changes that occurred through their history. The reflexes of PP vowels in open syllables are summarized in Table 5. Since length was not contrastive in PP open syllables, it will not be indicated in reconstructed PP forms. Note that the voicing flip-flop must have already occurred by the Pre-RP stage since both RP and KP agree in this respect.

Table 5. Reflexes of Proto-Palaung vowels in open syllables

PP	**RP**	**KP**
*-a	-a:	-a:
*-ɤ	-ɤ:	-ɤ
*-i	-ɛ:	-aj
*-e	-a:j	-aj
*-ɛ	-e:	-e
*-u	-ɔ:	-aw
*-o	-a:w	-aw
*-ɔ	-o:	-o

[11]Tone marks are omitted.

Considering the PP reflexes in RP[12] and KP as summarized above, the most important step toward understanding the development of RP vocalism is to figure out the reflexes of PP vowels in the immediate common ancestor of PR and KP, labeled Pre-Red Palaung in this paper. In the following section, I will propose a scenario of sound changes from PP to Pre-RP and to RP.

4. Tai loanwords and the history of Red Palaung

Borrowed words incorporated into a language at different times show different outcomes in the modern language. In this section, I propose the reflexes of PP vowels in open syllables in Pre-Red Palaung (Pre-RP), the hypothetical common ancestor of RP and KP, drawing evidence from Tai loanwords. In addition, I use data from Tai borrowings to establish a relative chronology of the sound changes that RP has undergone.

4.1 Reconstructing Pre-Red Palaung

To reconstruct Pre-Red Palaung, Palaung-internal evidence is fundamental. As shown in Table 5, RP and KP agree in having /a:/, /ɤ/, /a(:)j/, /a(:)w/ and /o(:)/ for PP */a:/, */ɤ/, */e/, */ɛ/, */o/ and */ɔ/ respectively. For these cases, the modern reflexes are projected back to the Pre-RP stage as shown in Table 6.

Table 6. Reconstruction of Pre-Red Palaung vowels in open syllables

PP	Pre-RP	RP	KP
*-a	*-a:	-a:	-a:
*-ɤ	*-ɤ	-ɤ	-ɤ
*-i	?	-ɛ:	-aj
*-e	*-a:j	-a:j	-aj
*-ɛ	*-e:	-e:	-e
*-u	?	-ɔ:	-aw
*-o	*-a:w	-a:w	-aw
*-ɔ	*-o:	-o:	-o

As seen above, data from RP and KP together are not enough to establish the quality of the reflexes of PP */-i/ and */-u/ in Pre-RP. Owing to the contact between Palaung and its Tai neighbors, Tai loanwords in Palaung can shed light on the reconstruction of these Pre-RP rimes. Consider the set of Tai loanwords given in Table 7.

[12]Refer to section 2 for a brief description of the present-day RP vowel system.

Table 7. Tai loanwords in PR going back to TL /-aj/ and /-aw/.

RP	PSWT	TL
$t^{h}ɛ$: 'to plough'	*$t^{h}aj^{A}$	$t^{h}aj^{A1}$
$t^{h}ɛ$: 'Thai'[13]	*daj^{A}	taj^{A2}
$^{h}lɛ$: 'to flow'	*$^{h}laj^{A}$	laj^{A1}
ɟɔ: 'prince'	*caw^{C}	saw^{C1}

These borrowings all have either /-ɛ:/ or /-ɔ:/ in RP but they go back to forms with /-aj/ or /-aw/ in all the potential donor languages. This indicates that there was a process in RP that transformed /-aj/ into /-ɛ:/. It is safe to assume that this process is exactly the one that causes RP and KP to diverge with respect to vowels in open syllables. That is, this set of Tai loanwords had /-aj/ or /-aw/ when they were incorporated into Palaung. Consequently, there was a change in RP in which PP */-aj/ and */-aw/ became /-ɛ:/ and /-ɔ:/ respectively. Therefore, this set of Tai loanwords in RP indicate that PP */-i/ and */-u/ were reflected as */-aj/ and */-aw/ in Pre-RP. The changes of vowels in open syllables from PP to RP are schematized in Table 8 below.

Table 8. Development of PP vowels in open syllables

	PP		Pre-RP		RP
1	*-i *-u	→	*-aj *-aw	→	-ɛ: -ɔ:
2	*-e *-o	→	*-a:j *-a:w	→	-a:j -a:w
3	*-ɛ *-ɔ	→	*-e: *-o:	→	-e: -o:
4	*-a:	→	*-a:	→	-a:
5	*-ɤ	→	*-ɤ	→	-ɤ:

Since PP */-a:/ and */-ɤ/ have remained stable since the PP stage, they will be ignored in the rest of this paper. The following section situates Tai loanwords in the contexts of sound changes in Palaung.

4.2 Tai loanwords and sound changes in RP

Reconstruction of an earlier stage of a language depends largely on the availability of comparative data as well as the interpretation of the data by the researcher. In this process, loanwords play a major part in such interpretations. Not only do they provide hints for the interpretation, they also attest or disprove particular choices of reconstructions. Tai loanwords in RP both confirm and refine the reconstruction of Pre-RP proposed above. Here each Palaung sound changes are discussed in relation to relevant Tai loanwords.

[13]These RP forms refer specifically to the Thai of Thailand.

4.2.1 Raising of RP low vowels

As discussed above, PP */-ɛ/ and */-ɔ/ became */-e/ and */-o/ in the Pre-RP stage. This is a case of raising in which low vowels became mid vowels. A number of Tai loanwords in Palaung confirm that such raising occurred. Consider the set of data given in Table 9.

Table 9. Tai loans that went through the low-vowel raising

	RP	**PSWT**	**TL**
'palace'	ho:	*hɔ:^{A}	hɔ^{A1}
'to cast'	^{h}lo:	*^{h}lɔ:^{B}	lɔ^{B1}

The two loanwords above are confirmations that PP low vowels *ɛ/ and */-ɔ/ did change to */-e/ and */-o/ in Pre-RP. They must have been incorporated into Palaung before the raising occurred. These mid vowels are still preserved in modern RP. Note that the phonemicization of RP /-o:/ and /-e:/ obscures the fact that these modern vowels are diphthongal, i.e., they are pronounced as [-ow] and [-ej] respectively. It is then possible that there was an intermediate stage when PP */-ɛ/ and */-ɔ/ were realized as *[-ɑj] and *[-ɑw] respectively.

4.2.2 Diphthong warping

In Table 7 above, Tai loanwords were used to inform the reconstruction of Pre-RP reflexes of PP */-i/ and */-u/. Needless to say, that set of Tai loanwords serve as evidence for positing a change from */-aj/ and */-aw/ to /-ɛ:/ and /-ɔ:/ respectively. The data are repeated in Table 10.

Table 10. Tai loanwords showing the diphthong warping

	RP	**PSWT**	**TL**	**Note**
'to plough'	t^{h}ɛ:	*thaj^{A}	t^{h}aj^{A1}	
'Thai'	t^{h}ɛ:	*daj^{A}	taj^{A2}	TH t^{h}aj^{A2}
'to flow'	^{h}lɛ:	*^{h}laj^{A}	laj^{A1}	
'prince'	ɟɔ:	*caw^{C}	saw^{C1}	

This change must have taken place quite late in the history of Palaung because KP did not go through this change, suggesting that the change occurred at the time or after Pre-RP split into KP and RP. Note that the phonemicization of RP /-ɔ:/ and /-ɛ:/ obscures the fact that these modern vowels are phonetically diphthongs, i.e. they are pronounced as [-ɔw] and [-ɛj] respectively. Therefore, the change of vowel warping can be understood as assimilation of the main vocalic elements to their following glides.

4.2.3 Diphthongization of mid vowels

By the Pre-RP stage, PP */-e/ and */-o/ had become */-a:j/ and */-a:w/ as discussed above. The phonetic difference between mid vowels and diphthongs with long /a:/ as first element is quite great, suggesting some

intermediate step. One Tai loanword sheds light on the question as shown in Table 11.

Table 11. Tai loanwords showing the diphthongization of mid vowels

	RP	**PSWT**	**TL**	**Note**
'to weave'	da:w	*dɔ:	tɔ:	

The form 'to weave' seems to have been incorporated into RP after */-o/ had changed into something approximating *[-ɔ:] but before that intermediate stage diphthongized into */-a:w/. Positing */-ɔ:/ and */-ɛ:/ for the intermediate step is appealing but problematic because such scenario entails a flip-flop between PP */-e/ and */-o/ on one hand and PP */-ɛ/ and */-ɔ/ on the other. The solution to this problem lies in two facts about the history of RP. First, PP */-i/ and */-u/ diphthongized sometime between PP and Pre-RP stage, which is also the time when PP */-e/ and */-o/ developed into the intermediate stage in question. This suggests that PP mid vowels may have also diphthongized during this time. Second, all modern RP open-syllable rimes except for /-a:/ are diphthongal, including those rimes that are analyzed as monophthongs phonemically. This suggests that these rimes must have been diphthongs at some point in their history. I speculate that at the intermediate stage PP */-e/ and */-o/ was reflected as *[-ʌj] and *[-ʌw] respectively. These two diphthongs then lengthened to Pre-RP *[a:j] and *[a:w]. The proposed development of PP */-e/ and */-o/ can be schematized in relation to other vowels as in Table 12.

Table 12. Intermediate steps in the development of PP vowels in open syllables

PP				**Pre-RP**		**RP**
*-i *-u	→	(*-ɤj) (*-ɤw)	→	*-aj *-aw	→	-ɛ: [-ɛj] -ɔ: [-ɔw]
*-e *-o	→	(*-ʌj) (*-ʌw)	→	*-a:j *-a:w	→	-a:j -a:w
*-ɛ *-ɔ	→	(*-ɑj) (*-ɑw)	→	*-e: [-ej] *-o: [-ow]	→	-e: [-ej] -o: [-ow]

Not only does the proposed account of the development from PP */-e/ and */-o/ to RP /-a:j/ and /-a:w/ avoid the flip-flop problem but it also allows us to view the intermediate step as an open-syllable diphthongization in which the first half of the vowel systematically lowered one level. The resulting intermediate system should then be characterized as still keeping the PP height contrast among simple vowels but having developed an allophonic variation between monophthongal realizations in closed syllables and diphthongal variants in open syllables. In the proposed schema, the TL form /tɔ^{A1}/ was incorporated into Palaung as *[tʌw]. Also, notice that the first part of */-ɛ/ and */-ɔ/ did not lower.

4.2.4 Diphthongization of high vowels

In addition to the above three changes, the diphthongization of PP */-i/ and */-u/ to Pre-RP /-aj/ and /-aw/ is another very important change. This gliding phenomenon is very similar to the case of the Great Vowel Shift in English, in which /i:/ became /aj/. Unfortunately, no Tai loanwords in RP recorded so far attests this pattern of change. It is very likely that PP */-i/ and */-u/ went through an intermediate step where they were reflected as *[-ɤj] and *[-ɤw] respectively, before they finally became /-aj/ and /-aw/ as observed in Pre-RP. This is also illustrated in Table 12. This change must have occurred relatively early in the history of Palaung because it is one that both KP and RP went through.

4.3 Chronology of the RP sound changes

The four changes discussed in the preceding section can be ordered chronologically using bleeding/feeding relationships among themselves as well as information provided by Tai loanwords. Consider the chronological ordering of sound changes in open syllables presented in Table 13.

Table 13. Proposed relative chronology

1) Diphthongization of high vowels: *-u > (*-ɤw)> *-aw
2) Raising of low vowels: *-ɔ > (*-ɑw)> -o: [-ow]
3) Diphthongization of mid vowels: *-o > (*-ʌw) > -a:w
4) Diphthong warping: *-aw > -ɔ: [-ɔw]

Because the chronology of the diphthongization of high vowels is largely speculative, the ordering among the other three changes can be considered first. The raising of PP low vowels must have been completed the earliest, followed by the diphthongization of mid vowels, and lastly by the diphthong warping.

The evidence for ordering the raising of low vowels before the diphthongization of mid vowel crucially comes from the fact that Tai low vowels are reflected either as */-o:/ or */-ɔ:/. Recall the data in Table 9 and Table 11. Crucially, there must have been a stage in which PP */-ɔ/ was already reflected as *[-ow] but PP */-o/ was still *[-ʌw]. If the diphthongization had been completed before the raising, such stage would not have existed. TL /tɔ:^{A2}/ would have been incorporated with the reflex of PP */-ɔ/, which would be more similar to Tai */-ɔ/ than the reflex of */-o/. Compare the proposed scenario in Table 14 to the ordering ruled out by the existence of RP /da:w/ in Table 15.

Table 14. Schematization of the ordering between the low-vowel raising and the diphthongization of mid-vowels

TL hɔ^{A1} ↓				**RP**
*hɔ				ho: [how] ‘palace’
PP *-ɔ	> (*-ɑw)	> [-ow]		
PP *-o		> (*-ʌw)	> -a:w	
		*tʌw		da:w ‘to weave’
		↑ TL tɔ^{A2}		

Table 15. Schematization of the incorrect ordering between the raising and the diphthongization

TL hɔ^{A1} ↓		**TL TL tɔ^{A2}** ↓		**RP**
*hɔ		tɑw		ho: [how] ‘palace’
PP *-ɔ		>(*-ɑw)	>[-ow]	do: [dow] ‘to weave’
PP *-o	> (*-ʌw)	> -a:w		

The ordering of the diphthong warping is not established on the basis of loanwords but by examining its feeding/bleeding relationships to other changes. It is apparent that the change from */-u/ > */-aw/ must have fed into the change */-aw/ > /-ɔ:/ so that PP */-u/ is reflected as */-aw/ in RP. This ordering also holds for PP front vowels. Many Tai loanwords in RP followed the proposed path of development exemplified in Table 16.

Table 16. Schematization of the ordering between the diphthongization of high vowels and the diphthong warping

		PSWT caw^{C} ↓		**RP**
	*caw			
Diphthongization	*-u	> -aw		
Warping		-aw	> ɔ: [ɔw]	ɟɔ: [ɟɔw] ‘prince’

More importantly, diphthong warping must have occurred after the PP low vowels had raised to /-o:/. Otherwise, the low vowels that resulted from the warping would have fed into the raising as shown in Table 17.

Table 17. Schematization of the ordering between the low-vowel raising and the diphthong warping

	PSWT caw^{C} ↓			**RP**
	*caw			
Warping	*-aw	> -ɔ:		
Raising		*-ɔ	> -o:	ɟo: [ɟow] ‘prince’

In addition, diphthong warping seems to have occurred after the diphthongization of mid vowels was completed. The evidence is from the comparison between RP and KP. While both RP and KP went through the diphthongization, only RP underwent the warping. This suggests that the warping occurred after RP and KP split, but the diphthongization took place at the latest in the Pre-RP period. This places diphthong warping very late in the history of Palaung.

Having established the ordering of the raising of low vowels, the diphthongization of mid vowels, and the vowel warping, the relative chronology of the diphthongization of PP high vowels can be examined. As discussed above, the change must have preceded the diphthong warping. The question then is how it is ordered relative to the low-vowel raising and the diphthongization of PP mid vowels. Unfortunately, this change does not interact with other changes except for the vowel warping. Therefore, no feeding/bleeding relationship can be deduced. Moreover, no Tai loanwords provide any additional information. However, the total absence of Tai loanwords that went through this particular change allows us to speculate that it must have taken place even before Palaung started to borrow from Tai. This speculation places this diphthongization as the first change in the history of RP.

5. Conclusion

This paper has drawn from both Palaung-internal evidence and evidence from Tai loanwords to establish a chronology of changes that Proto-Palaung vowels in open syllables underwent. I have proposed that the four changes occurred in the following order: 1) diphthongization of high vowels, 2) raising of low vowels, 3) diphthongization of mid vowels, and 4) diphthong warping. The proposed relative chronology is another step toward an understanding of the history of Palaung dialects.

REFERENCES

Chun, Shizeng, ed. 1989. *Chinese Nationalities.* Photography and Art Press of China Nationalities.

Diffloth, Gérard. 1988. "Palaungic vowels in Mon-Khmer perspective." In H. C. S. Davidson, Jeremy, ed., *Austroasiatic Languages: Essays in Honour of H.L. Shorto*, pp. 13-27. London: School of Oriental and African Studies, University of London.

Diffloth, Gérard. 2005. "The contribution of linguistic paleontology to the homeland of Austroasiatics." In L. Sagart, R. Blench and A. Sanchez-Mazas, eds., *The Peopling of East Asia: Putting Together Archeology, Linguistics, and Genetics*, pp. 79-82. London: RoutledgeCurzon.

Howard, Michael C. and Wattana Wattanapun. 2001. *The Palaung in Thailand.* Chiangmai, Thailand: Silkworrm Book.

Hudak, Thomas J., ed. 1994. William J. *Gedney's Southwestern Tai Dialects: Glossaries, Texts and Translation.* Ann Arbor: Center for South and Southeast Asian Studies, University of Michigan.

Jonsson, Nanna. 1991. "Proto Southwestern Tai." Ph.D. dissertation. Department of Linguistics, State University of New York at Albany. Albany, NY

Kasisopa, Benjawan. 2003. "Phonological study and genetic classification of Dara-ang (Palaung) spoken at Nawlae Village, Fang District, Chiangmai Province." M.A. thesis. Department of Linguistics, Chulalongkorn University. Bangkok.

Li, Fang-Kuei. 1977. *Handbook of Comparative Tai.* Honolulu: University of Hawai'i Press.

Matisoff, James A. 1973. "Tonogenesis in Southeast Asia." In L. M. Hyman, ed., *Consonant Types & Tones*, pp. 71-95. Los Angeles: The Linguistic Program, University of Southern California.

Mitani, Yasuyuki. 1977. "Palaung dialects: a preliminary comparison." *Southeast Asian Studies* 15:193-212.

Mitani, Yasuyuki. 1979. "Vowel correspondences between Riang and Palaung." In T. L. Thongkum, ed., *Studies in Thai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J. A. Henderson*, pp. 142-150. Bangkok: Chulalongkorn University Press.

Pittayaporn, Pittayawat. 2002. *Red Palaung.* Fieldnote.

Pittayaporn, Pittayawat. 2008. "Proto-Southwestern Tai revised: A new reconstruction." *Paper presented at the 18th Annual Meeting of the Southeast Asian Linguistics Society,* Universiti Kebangsaan Malaysia, Bangi, Malaysia, May 21-22, 2008.

Pittayaporn, Pittayawat. in preparation. *A Handbook of Proto-Southwestern Tai.*

Pittayaporn, Pittayawat. n.d. "Final Nasals and Vowels in Red Dar'ang." Unpublished ms. 11 pp.

Scott, J. G. and J.P. Hardiman, eds. 1900. "Gazetteer of Upper Burma and the Shan States." Compiled from official papers by J. George Scott, assisted by J. P. Hardiman. Rangoon: Printed by the superintendent, Government printing, Burma, 1900-01.

Sila, Sarapee. n.d. Palaung. Unpublished report. "Center for Minority Research."

Yangderm, Khuen. 1995. "Journey to Salween and Thanon Thongchai: Ways of creeks and legends of valleys." Bangkok: Patchuban Sangsan.

Received: 25 January 2008

Department of Linguistics,
Faculty of Arts, Chulalongkorn University
Bangkok, THAILAND
<Pittayawat.P@chula.ac.th>

APPENDIX

Tai loans in Red Palaung

Note: 1) PSWT forms are given only when they are found in the lexicon of PSWT etyma reconstructed in Pittayaporn (in preparation).
2) While many forms show clear Shan affinities, in most cases it is not possible to pinpoint the donor Tai dialects due to lack of evidence that would allow discrimination among various possible Tai sources. This is especially true for older borrowings.

No.	Items	Meaning	PSWT	Thai	Tai Long
A. Natural Objects					
1	geːŋ	‘shin’	*ɣɛŋ^{B}	kʰɛŋ^{B2}	kʰɛŋ^{B2}
2	raːŋ	‘body’	*raːŋ^{B}	raːŋ^{B2}	haːŋ^{B2}
3	naː	‘face’	*ʰnaː^{C}	naː^{C1}	naː^{C1}
4	ʰnoŋ	‘lake’	*ʰnɔŋ^{A}	nɔːŋ^{A1}	nɔŋ^{A1}
5	sɛŋ	‘precious stone’	*sɛːŋ^{A}	sɛːŋ^{A1}	šɛŋ^{A1}
6	glɔːj	‘banana’	*kluəj^{C}	kluəj^{C1}	koj^{C1}
7	guŋ	‘shrimp’	*kuŋ^{C}	kuŋ^{C1}	kuŋ^{C1}
8	ŋoŋ	‘elephant trunk’	*ŋuəŋ^{A}	ŋuəŋ^{A2}	ŋoŋ^{A2}
9	kaː	‘rice seeding’	*klaː^{C}	klaː^{C1}	kaː^{C1}
10	moː	‘lotus’	*ʔbuə^{A}	buə^{A1}	mo^{A1}
11	bliŋ	‘leech’	*pliːŋ^{A}	pliŋ^{A1}	piŋ^{A1}
12	pʰiːn	‘opium’		fin^{B1}	pʰin^{B1}
13	laːŋ	‘jack fruit’		laːŋ^{A2}	laːŋ^{A2}
14	ʰnuːn	‘cotton’	*ʰnun^{C}	nun^{C1}	
15	moːŋ	‘mango’	*maːk^{D} muəŋ^{B}	maːk^{DL1} muəŋ^{B2}	maːk^{DL1} moŋ^{B2}
16	vɔːn	‘Caladium’	*ʔbɔːn^{A}	bɔːn^{A1}	mɔn^{A1}
17	doː	‘figure animal’	*tuə^{A}	tuə^{A1}	to^{A1}
18	saːŋ	‘elephant’	*ɟaːŋ^{C}	cʰaːŋ^{C2}	saːŋ^{C2}
19	hɔːj	‘shellfish’	*ʰrɔːj^{A}	hɔj ^{A1}	hɔːj^{A1}
20	moːŋ	‘gong’		moːŋ^{A2}	moŋ^{A2}
21	bep	‘duck’	*pet^{D}	pet^{DS1}	pit^{DS1}
22	ma ˈkʰɤː ma ˈkʰɤː	‘eggplant’	*maːk^{D} kʰɯə^{A}	ma ˈkʰɯə^{A1}	maːk^{DL1} khɤː^{A1} maːk^{DL1}
23	soːm	‘tomato’			khɤː^{A1} šom^{C1}
24	ma ˈʔoː	‘pomelo’		som^{C1} ʔoː^{A1}	
B. Man-made Objects					
1	hoː	‘palace’		hɔː^{A1}	hɔ^{A1}
2	lɤː	‘saw’	*lɯə^{B}	lɯəj^{B2}	lɤ^{B2}
3	ʰmɔːn	‘pillow’	*ʰmɔːn^{A}	mɔːn^{A1}	mɔn^{A1}

No.	Items	Meaning	PSWT	Thai	Tai Long
4	deːn	‘candle’	*diən^{A}	tʰiən^{A2}	ten^{A2}
5	hɔːk	‘lance’	*ʰrɔk^{D}	hɔːk^{DL1}	hɔk^{DL1}
6	ʰmaːj	‘aim’	*ʰmaːj^{A}	maːj^{A1}	maːj^{A1}
7	duŋ	‘flag’		tʰoŋ^{A2}	tuŋ^{A2}
8	lɔː	‘cart’	*lɔː^{C}	lɔː^{C2}	lɔ^{C2}
9	rɤː	‘boat’	*rɯə^{A}	rɯə^{A2}	hɤ^{A2}
10	kʰaːw	‘news’		kʰaːw^{B1}	kʰaːw^{B1}
11	pɛː	‘raft’	*bɛː^{A}	pʰɛː^{A2}	pɛ^{A2}
12	roːŋ	‘ditch’	*rɔːŋ^{B}	rɔŋ^{B2}	hɔŋ^{B2}
13	kʰroŋ	‘cage’	*kroŋ^{A}	kroŋ^{A1}	kʰoŋ^{A1}
14	kʰlɯp	‘bamboo hat’	*klup^{D}		kup^{DS1}
C. Society					
1	ʰlaːn	‘grandchild’	*ʰlaːn^{A}	laːn^{A1}	laːn^{A1}
2	cɯː	‘name’	*ɟɯː^{B}	cʰɯ^{B2}	sɯ^{B2}
3	joːn	‘Northern Thai’		juən^{A2}	
4	tʰɛː	‘Thai’	*daj^{A}	tʰaj^{A2}	taj^{A2}
5	ɟɔː	‘prince’	*caw^{C}	caːw^{C1}	saw^{C1}
6	sɯk	‘soldier’	*sɯk^{D}	sɯk^{DS1}	sɯk^{DS1}
7	kʰun	‘lord’		kʰun^{A1}	kʰun^{A1}
8	kʰrɯː	‘family lineage’		kʰrɯə^{A2}	
9	su ˈmɤːŋ	‘guardian angel of city’		sɯə^{C1} mɯəŋ^{A2}	sɤ^{C1} mɤŋ^{A2}
D. Verbs					
1	kʰoː	‘to fry’		kʰuə^{C1}	kʰo^{C1}
2	ʰloː	‘to cast’	*ʰlɔː^{B}	lɔː^{B1}	lɔ^{B1}
3	heːn	‘to learn’		riən^{A2}	
4	dɛːm	‘to write’	*tɛːm^{C}	tɛːm^{C1}	tɛm^{C1}
5	ciːm	‘to taste’	*ɟiːm^{A}	cʰim^{A2}	sim^{A2}
6	ʰlɛː	‘to flow’	*ʰlaj^{A}	laj^{A1}	laj^{A1}
7	jɔŋ	‘to hate’	*ɟaŋ^{A}	cʰaŋ^{A2}	saŋ^{A2}
8	cɔːj	‘to help’	*ɟuəj^{B}	cʰuəj^{B2}	sɔj^{B2}
9	piːŋ	‘to roast’	*piːŋ^{C}	piŋ^{C1}	piŋ^{C1}
10	ban	‘to mould’	*pan^{C}	pan^{C1}	pan^{C1}
11	ʰraːj	‘to disappear’	*ʰraːj^{A}	haːj^{A1}	haːj^{A1}
12	khɛŋ	‘to compete’		kʰɛŋ^{B1}	kʰɛŋ^{B1}
13	rɔk	‘to love’	*rak^{D}	rak^{DS2}	hak^{DS2}
14	laːk	‘to drag’	*laːk^{D}	laːk^{DL2}	laːk^{DL2}
15	tʰɛː	‘to plough’	*tʰaj^{A}	tʰaj^{A1}	tʰaj^{A1}
16	daːw	‘to weave’	*dɔː^{A}	tʰɔː^{A2}	tɔ^{A2}
17	kʰrɔŋ	‘to imprison’	*kʰraŋ^{A}	kʰaŋ^{A1}	kʰaŋ^{A1}
E. Others					
1	pʰɯːn	‘classifier for cloth’		pʰɯːn^{A1}	pʰɯn^{A1}
2	lɯk	‘deep’	*lɤk^{D}	lɯk^{DS2}	lɯk^{DS2}
3	hɛŋ	‘strength’	*rɛːŋ^{A}	rɛːŋ^{A2}	hɛŋ^{A2}
4	haːj	‘wild’	*raːj^{C}	raːj^{C2}	haːj^{C2}

No.	Items	Meaning	PSWT	Thai	Tai Long
5	loːŋ	‘downward’	*loŋ^{A}	loŋ^{A2}	loŋ^{A2}
6	^{h}mɯn	‘10,000’		mɯːn^{B1}	
7	sɛːn	‘100,000’	*sɛːn^{A}	sɛːn^{A1}	šɛn^{A1}
8	laːn	‘1,000,000’	*laːn^{C}	laːn^{C2}	laːn^{C2}
9	gɛːn	‘center’		kɛːn^{A1}	
10	diː	‘locative marker’	*diː^{B}	t^{h}i^{B2}	ti^{B2}
11	heŋ	‘dry’	*hɛːŋ^{C}	hɛːŋ^{C1}	hɛŋ^{C1}

On k^{hw}–f alternations in Bangkok Thai and other Tai languages[1]

Daniel HUBER
Université Rennes 2, France

Abstract

The paper presents a historical phonological explanation for the Bangkok Thai k^{hw}–f alternations as in *faa*^{A1} for Standard Thai *khwaa*^{A1} 'right (side)'. While this phenomenon has attracted the attention of scholars in various linguistic fields such as the phonetics of Thai, sociolinguistics of Bangkok Thai and applied linguistics (teaching Thai as a "foreign" language), the potential of Tai dialectology and historical linguistics has not been fully taken into account. The paper argues that such alternations are widely attested among Tai varieties, that they are neither a synchronic alternation, nor an innovation in Bangkok Thai, and that the phenomenon ultimately goes back to a change that turned PSWT *x^w- into /f-/.

1 Introduction

The alternation between a labio-velar stop and a labial fricative in working class Bangkok Thai, as exemplified by *faa*^{A1} for literary (=Standard) Thai *khwaa*^{A1} 'right (side)', has been touched upon in the literature from a number of perspectives. Most descriptions have only concentrated on some particular aspect of this phenomenon, pointing out its sociolinguistic connotations. The Bangkok Thai pronunciation *faa*^{A1} for Standard Thai *khwaa*^{A1} 'right (side)' is generally associated with specifically working class Bangkok Thai. No account, however, has come up with a satisfyingly complete linguistic explanation for this particular change. The areas of study where this alternation has gained prominence include the phonetics of Standard Thai (for instance, Harris 1972, Henderson 1985, 1987), the sociolinguistics of Bangkok Thai (Beebe 2001), the teaching of Standard Thai as a foreign language (Brown 1979/1997, Smyth 2003), while Tai dialectology and historical linguistics also has much to say about such changes from a broader perspective (Jones 1965, Li 1977). Interestingly, the potential of the last field of study, Tai dialectology and historical linguistics, has been hardly put to use to explain this stigmatized Bangkok usage.

It will be shown in this paper that Tai dialectology and historical linguistics can offer a reasonable explanation for the Bangkok Thai sociolinguistically conditioned alternation. It will be shown that this usage is

[1]The research for this paper, carried out at SOAS between September 2005 and January 2006, was made possible by the Hungarian State Eötvös Scholarship.

not an active alternation at all, but the **remnants** rather of a historical change where *x^w- turned into /f-/. Accounts which have considered Standard Thai exclusively cannot offer a satisfying explanation, and sometimes they are simply wrong in their assumptions. While the phenomenon tends to have negative sociolinguistic connotations in Bangkok Thai, the change of a labio-velar to a plain labial is very much attested in other Tai varieties, such as Songkhla, in exactly those few lexical items that happen to be cited in the literature on the Bangkok Thai phenomenon. This observation, while clarifying the historical phonological reasons for the change, also opens up the possibility of a more accurate sociolinguistic explanation, which partly contradicts Beebe's (2001) position. While the present paper is phonological and historical in its scope, it may appeal to scholars in other fields of linguistics, theoretical or applied, and scholars of Thai social history.

2 Earlier accounts of the phenomenon

Let us sum up earlier accounts briefly. J.G. Harris (1972) lists [f-] as a possible allophone of /k^{hw}-/ in his description of Standard Thai consonants, and notes that the usage is linked to the speech of lower class speakers. Henderson (1985), citing strikingly similar data from the Songkhla dialect, argues for a feature shuffling whereby the (abstract/underlying) labiality associated with the initial segment either gets integrated completely into the initial consonantal slot of a word to give a plain labial, [f-], or it has a time lag over the underlying velarity, resulting in [$k(^h)^w$-]. In another article, Henderson (1987) points to what she terms the **inherent velarity** of Thai /t s f/, which had also been noted by Harris (1972) as a possible phonetic cause for this alternation. Beebe (2001) analyzes the Bangkok phenomenon from a sociolinguistic perspective, arguing that changes of labio-velars are instances of cluster reduction, and, together with other cluster reductions in Bangkok Thai such as /kl-/>/k-/ or /p^hr-/>/p^h-/, it is an innovation by speakers of lower social class. Brown (1979/1997:51, and notes 48 and 50) also describes such "compromise consonants", but he attaches this usage to natives of the vicinity rather than a social class of Bangkok. Finally, this particular change features prominently in Thai language teaching as well, where specific attention is drawn to this typical Bangkok working class pronunciation (Smyth 2002, 2003). Let us now turn to a more detailed analysis of these earlier approaches.

While all these accounts highlight important aspects of this phenomenon, they leave something to desire. From a methodological point of view, there is the problem that most of the above sources tend not to mention many concrete examples of the change in Bangkok Thai so that one could see a range of lexical items that show this phenomenon. Harris (1972:10-11), writing from a purely phonetic point of view, cites *khwaa*^{A1} 'right (side)' as an example of the change, and mentions (1972:11) that [faa] for this word is "considered low class pronunciation by educated speakers". Beebe (2001), in her turn, gives no examples at all, although alluding to the simplification of labio-velars. The lack of data is particularly painful in this case since she

explicitly aims at analyzing precisely such reductions.[2] Finally, Smyth (2002:7) mentions this "more radical transformation, associated with Bangkok working-class speech", and (1) includes the items he cites:

(1)	**Standard Thai**		**working-class**	**gloss**
	khwaa^{A1}	>	*faa*^{A1}	right
	khwaam^{A2} *sùk*	>	*faam*^{A2} *sùk*	happiness

Notice that Smyth (2002:7) makes it explicit that it is aspirated /k^{hw}/ that gets transformed into /f/, not an unaspirated one. The same solitary example, *khwaa*^{A1}, is also found in Smyth (2003:xiii). Recall that Harris (1972) also gives this single example, one with an aspirated stop, not plain /k^{w}/. In fact, Beebe's (2001:31) statistics also show that unaspirated /kw/ is less prone to reduction than its aspirated counterpart /k^{hw}/: 82-82% of unaspirated /k^{w}/ is retained (in the two age groups examined) as opposed to a slightly lower proportion, 73-79%, of /k^{hw}/ (see (9) below). Nevertheless, Brown (1979/1997:51) cites *khwaai*^{A2} 'water buffalo', with apirated /k^{hw}/, **as well as** *kwaat*^{D1L} 'to sweep', with plain /k^{w}/ as examples having /f/ in Bangkok Thai. The latter item is problematic – for the time being, it seems safest to consider it a divergent development.

Studies on Tai dialects present strikingly similar changes, but, again, there are no long lists of words affected. For instance, (2) includes all the data Henderson (1985:10) gives from the Songkhla dialect for free variation between a labio-velar and a plain labial initial consonant:

(2)	**Songkhla free**		**variation**	**Standard Thai**		**gloss**
	fai	-	*kwai*	*fai*	A2	fire
	fon	-	*khon*	*fon*	A1	rain

Although in (2) different items are concerned than in (1) above and the Standard Thai forms are invariably with initial /f/, the alternation still involves (labio-)velars and plain labials. Jones (1965:208) also confirms the alternation in Songkhla *kwai*^{A2} 'fire', but indicates (1965:210) no alternation in Songkhla *kh*w*on*^{A1} 'rain'.[3] Both cases in (2) then seem to involve **velar stops in a labial environment** alternating with plain labial fricatives. What is important is that in Songkhla both aspirated and unaspirated labio-velars alternate with plain labial /f/. And recall that Brown also cites an example with plain velar: *kwaat*^{D1L} 'to sweep'.

It will have been noted that, except for Beebe (2001), none of these works were written specifically on the Bangkok Thai phenomenon. To sum up the argumentation so far: while the stigmatized status of this usage is emphasised in many cases, there is simply no exhaustive list of the lexical items involved, only a handful of scattered words, listed in (1) and (2) above

[2]To do justice to Beebe, I am aware of work by her which might contain some relevant data, still it is a pity that Beebe (2001) cites none.

[3]One reviewer has kindly noted that 'rain' is *khon*, not *khwon* in Songkhla as affirmed by an elderly native speaker. Jones might then either be wrong here, or the difference may be due to the considerable difference in the date of data collection.

(and Brown's two further examples), are related to the phenomenon in some way.

From a theoretical phonological point of view, it is more problematic that no author, to the best of my knowledge, has ever commented on the fact that in this change sounds of different manners of articulation, namely a stop and a fricative, alternate. In other words, beyond the reconfiguration of the place of articulation, there is a concomitant change in the manner of articulation as well: a stop becomes a fricative. This observation seems to have passed unnoticed. It is true that changes of manners of articulation are not uncommon in languages (see, for instance, Grimm's Law), but in labial–velar interactions, the following possibilities are found to obtain:

(3a)	k^w	>	p	for instance: Latin to Romanian
(3b)	x^w	>	f	for instance: Late Middle Chinese to Cantonese
(3c)	k^w	>	*f	**unattested**
(3d)	x^w	>	*p	**can be excluded on other grounds**

While alternations between identical manners of articulation, like /k^w/ and /p/, both stops, are just as attested across languages as are /x^w/ > /f/ changes, why does a **stop** alternate with a **fricative** in Bangkok Thai since it is not even in a syllabic position where such lenitions are expected?[4] This observation holds true even for Henderson (1985), who gives an explanation for this phenomenon in terms of **feature shuffling**, an account which is strikingly autosegmental in spirit. She argues (1985:4) that such shufflings of features ([labiality], [velarity], etc.) are like "hands of playing cards" dealt out to whole syllables rather than to individual segments. While this analysis is essentially correct as far as the switch in the place of articulation is concerned, she does not consider the additional problem that not only places of articulation, but manners of articulation too are affected in this case. This problem also has to be reconciled in order to get a fuller picture of this phenomenon. And indeed, Tai dialect studies provide some important insights.

3 What Tai dialect studies and historical linguistics can say about this variation

Below it will be shown that the change **predates** modern Standard Thai, and is not an active phonological alternation at all. It is in fact a remnant of an earlier change from fricative /x^w-/ to fricative /f/. In other words, the change did originally involve **identical manners of articulation**, and the alternation was a simple case of labio-velars becoming plain labials. This observation has two implications for the linguistic situation of today. First, under this view, native speakers who have the [f] form, which had developed from *x^w-, have a different lexical item, evidently stored with /f-/, in their mental lexicon rather than synchronically manipulate initial /k^{hw}-/ to surface as [f]. The [f-] is, thus, not the result of a putative synchronic phonological process /k^{hw}-/ > [f-]. Second, /f/ for /k^{hw}-/ is clearly not an innovation on the part of modern speakers, as opposed to

[4]One could argue, following Harris (1972:11), that the phoneme /k^{hw}/ can be phonetically [k^{xhw}], but this is immaterial to the argumentation since such an affricate and [f] would still be of different manners of articulation.

Beebe's (2001) view, since it must predate the emergence of modern Standard Thai /k^{hw}-/ from earlier *x^{w}-. (It is quite a different matter whether /f/-forms are actually spreading from lower class speakers; more on this in section 4 below.) The ultimate source for this stigmatized usage in Bangkok Thai finds natural explanation in essentially identical variations attested across Tai dialects, such as Songkhla.

Jones (1965) has a most detailed comparative list of 81 lexical items in 39 Tai varieties for a reconstruction of Tai dialects. Although the main focus there is on the variation in tone systems, the items show segmental variations, too. They include /r/~/h/, /ŋ/~/h/ alternations, which will not be considered in this paper. As for the phenomenon at hand, Jones's list has only Standard Thai *khwan*^{A2} 'smoke', *khwaai*^{A2} 'water buffalo' and *khwaa*^{A1} 'right' as words having a variant in /f/ in at least some Southern Tai dialects of the 39 Tai varieties examined. It is important to point out that the parallel between the Bangkok Thai sociolinguistically conditioned variation and the variation in Tai dialects is based on the observation that (a) the data cited for Bangkok Thai also feature in dialectal variations in Jones, and (b) the Songkhla data show a phenomenon essentially identical to the variation in Bangkok Thai. This parallelism has also passed unnoticed, although it can be effectively put to explain the Bangkok Thai variation, as shown below.

Most interestingly, Jones also has three words that are with initial /f/ in Standard Thai, but show alternation in some Southern varieties. These are: *fai*^{A2} 'fire', *faa*^{C2} 'sky' and *fon*^{A1} 'rain'. Recall that Henderson's two examples cited in (2) above happen to belong to this latter set of lexical items: *fai*^{A2} and *fon*^{A1}. It is important in Jones's list that some other words that could potentially show this alternation, namely *kwaa*^{B1} 'exceed', *kwaang*^{C1} 'wide', are not indicated to show any segmental change even in Southern varieties (nor are the cognates of *kwaat*^{DL1}). This observation is crucial, as will be discussed later.[5]

The fullest list of words that show alternation, either dialectally or socially, between a voiceless labial fricative /f/ and a voiceless labio-velar cluster /k(^{h})^{w}/ then comprises the lexical items in (4a-b). The list below has been assembled mainly on the basis of Jones (1965).

(4) Fullest list of words with alternating /f/~/k^{hw}/

	alternating social/ or dialectal forms		**Standard Thai**		**gloss**
(a)	k(h)wai	fai	fai	A1	fire
	k(h)waa	faa	faa	C2	sky
	k(h)(w)on	fon	fon	A1	rain
(b)	k(h)waa	faa	khwaa	A1	right
	k(h)waam	faam	khwaam	A2	matter
	k(h)wan	fan	khwan	A2	smoke
	k(h)waai	faai	khwaai	A2	buffalo
(c)	no alternation:				

[5]One reviewer has kindly pointed out that some Bangkok Thai speakers of lower social class pronounce *kwaat* as *faat* 'to sweep' as well as *kwaa* as *faa* 'to exceed'. So an alternation between unaspirated labio-velar *kw* and *f* may also have to be recognized in some varieties.

	kwaa		kwaa	B1	exceed
	kwaaŋ		kwaaŋ	C1	wide
(d)	kwaat	faat	kwaat	D1L	to sweep
	kwaa	faa	kwaa	B1	exceed[6]

No doubt there exist other examples for (4b): since this phenomenon is one of the major shibboleths in Bangkok Thai pronunciation, it is not very likely that only a handful of words can attract such popular disdain. However, others do not appear in discussions of the phenomenon.

As can be seen in (4), the synchronic aspiration in the modern social and dialectal forms seems to play a role in the variation. Importantly, it is only those items that have a non-aspirated labio-velar /k^w/ in Standard Thai, (4c), that do not show alternation in the dialects, according to Jones. As for (4b), even though dialects might have lost aspiration (*kh^w > k^w), the alternation is only attested in items that show aspiration elsewhere, for instance in Standard Thai. This means that the emergence of the alternation in such dialects is earlier than the loss of aspiration. However, examples in (4d) are peculiar in that there is no sign of aspiration in any dialect. One could propose, for the time being, that these words came to begin with /x^w/ through *k^w > x^w in such varieties. This hypothesis needs further confirmation.

Li's *Handbook of Comparative Tai* (1977) has something to say about all the above words. In (5) are included all reconstructed initials for the items in (4). The words are cited in their Standard Thai forms, and they are assigned their Proto-Tai (PT) initials as reconstructed in Li (1977; tone classes are relevant for reconstructing voice distinctions):

(5)	**Thai**	**tone Class**	**gloss**		**PT initials**
(a)	fai	A2	fire	<	*v (p.79)
	faa	C2	sky	<	*v (p.79)
	fon	A1	rain	<	*f (p.78)
(b)	khwaa	A1	right	<	*khw (p.238)
	khwaam	A2	matter, spirit	<	*γw (p.242)
	khwan	A2	smoke	<	*γw (p.242)
	khwaai	A2	buffalo	<	*γw (p.242)
(c)	kwaa	B1	exceed	<	*kw (p.236)
	kwaaŋ	C1	wide	<	*kw (p.236)
(d)	kwaat	D1L	to sweep	<	*kw (p.236)

One finds that all the items with /k^{hw}/ in Standard Thai (5b=4b) go back either to a PT labio-velar fricative *γ^w- or a labio-velar aspirated stop *k^{hw}-, while those in /f/ in Thai are reconstructed (5a=4a) with PT *f- and *v- by Li. In (4c) the items that show no variation according to Jones are reconstructed by Li with a PT plain stop *k^w-, (5c). These reconstructions confirm the observation made above that plain *k^w- does not lead to alternation

[6]This item shows no alternation in Jones, but see preceding note.

while a labio-velar fricative or an aspirated labio-velar can. As for (5d) and *kwaa*^{B1} (see footnote 5), the *kw reconstructed by Li is also confirmed by another speech variety, Chiang Mai.

A few observations about Li's reconstructions in (5) are in order here. First, his reconstructions for PT *k^{hw}- are slightly problematic and might in fact be wrong. Li (1977:238) lists only four words (*khwai*^{B1} 'disorder', *khwai* ^{C1} 'twisted', *khwaaŋ*^{A1} 'broad' and *khwaa*^{A1} 'right') under that heading, the first two of which he claims to be related, and one is likely to be unaspirated *k^{w}- instead, he claims, hence it does not belong to this group. In connection with the fourth one, *khwaa*^{A1}, Li notes (ibid.) that there is "irregularity in the initial of the dialect forms." Hence it is conceivable that this item has PT *x^{w}-, not *k^{hw}-, as Li suggests. This means then that all the items in (5b) go back to a voiced or voiceless labio-velar **fricative**, PT *γ^{w}- or *x^{w}-.

Second, a word has to be said about PT *f and *v. Li sometimes falls victim in his reconstructions to the practice of treating Standard Thai as a workable basis for PT reconstruction, hence it is obvious to him that whatever is /f/ in Thai should go back to a PT *f or *v. In fact, in connection with *fai*^{A2} and *faa*^{C2}, there are some reservations (for instance, the vocalism of *fai*^{A2} is fairly problematic). Jones's data, however, offer a less biassed approach. Although all 39 varieties seem to point to a PT *f or *v in *fai*^{A2} and *faa*^{C2} (with regular *w*- reflex in Thô for both, Jones 1965:223), the data show interesting reflexes of *fon*^{A1} in various Eastern varieties: *hun*^{A1} in Dioi and *hyn*^{A1} in Po-Ai with *fyn*^{A1} in Thô (Jones 1965:225). This is tabulated in (6) below:

(6) Development of PT *f, *v, *γ^{w}, *γ and *k^{w} in Eastern dialects

	PT	**Standard Thai**		**Thô**	**Po-ai**	**Dioi**	**gloss**
(a)	*f, *v	fai	A2	wei	fii	fi	fire
		faa	C2	wa?	-	-	sky
	!	fon	A1	fyn	hyn	hun	rain
(b)	*γw	khwan	A2	wan	hon	hɔn	smoke
	*γ	khyyn	A2	kyn	hyn	hyn	night
(c)	*kw	kwaa	B1	-	kwaa	kwa	exceed
		kwaaŋ	C1	kwaaŋ	-	kwaŋ	wide

Since Dioi and Po-ai show /h-/ reflexes for PT velars in (6b), it cannot be excluded, definitely not on the basis of Standard Thai, that in Standard Thai *fon*^{A1} 'rain' the initial consonant was a PT velar fricative *x^{w}-, not a plain labial *f-.[7] (There are other objections as well in connection with PT *f- and *v- in general, but they will not be considered here since they would lead the argumentation further afield.) It can be established then that in two cases of (5a) the reconstruction with a PT labial fricative is uncertain (*fai*^{A2} and *faa*^{C2}), while

[7]One reviewer does not agree with this proposition, pointing out, correctly, that the reconstruction of the initial consonant of *fon*^{A1} 'rain' as *x^{w}- would clash with the existing *x^{w}- having the reflex k^{hw}. Indeed, this is predicted under my proposal. It must be the following labial vowel that labialized *x^{w}- to *f-. But since my reviewer added that the reflex of *f- as *h*- in Po-ai and Dioi is a regular sound change, I shall leave the question open here.

in the rest of the cases in (5), **including** *fon*[A1], there is clearly some PT labio-velar involved.

Turning now to the later development of the various PT labio-velars illustrated in (5b-c), according to Li, the fricatives in (5b) all became Proto-Southwestern-Tai *x^w- (through devoicing), of which Siamese developed modern /k^{hw}-/, while *k^w- in (5c) was retained. This development is crucial for the change at hand. Because if it is found that the words that show the alternation, that is, items in (5b), go back to a Proto-Southwestern-Tai *x^w- or *γ^w-, this is evidence that only those words could show the change that were of identical manner of articulation at the point in time when the switch took place. To put it differently, it has to be shown that only those words show variation, either in dialects or in Bangkok Thai, that go back to an earlier *x^w-/*γ^w- initial.

Below is a comparative list of Bangkok Thai and Chiang Mai Thai. Chiang Mai is important because it has retained earlier /x^w-/ where Bangkok Thai has developed /k^{hw}-/. The data in (7) reveal that all the items that show alternation still have a fricative rather than an aspirated stop initial in Chiang Mai (while having /f/ where Siamese also has /f/). The data in (7) come from Gedney (as edited by Hudak 1997), Siamese (=Standard Thai) equivalents have been supplied:

(7)	**Siamese**	**Chiang Mai**		**gloss**	
(a)	fai	fai	A2	fire	(p.738)
	faa	faa	C2	sky	(p.738)
	fon	fon	A1	rain	(p.739)
(b)	khwaa	xwaa	A1	right	(p.787)
	khwaam	<not cited>	A2	matter, spirit	
	khwan	xwɔn	A2	smoke	(p.787)
	khwaai	xwaai	A2	buffalo	(p.787)
(c)	kwaa	<not cited>	B1	exceed	
	kwaaŋ	kwaaŋ	C1	wide	(p.750)
(d)	kwaat	kwaat	D1L	to sweep	(p.750)

What these data reveal is that words that have been found in (4b) to show alternation all have a labio-velar fricative in Chiang Mai while words that do not show such alternation, those in (4c), have a plain velar stop /k/ in this variety. It can be concluded then that the phonological basis of these variations between a synchronic labio-velar and labial is simply an earlier change from *x^w- to modern /f/. Crucially, the development involves identical manners of articulation.

4 Reinterpereting the Bangkok Thai social variation; conclusions

The discussion so far has resulted in the following observations about the variation in Bangkok Thai. First of all, a workable database, in (4), had to be established to give a better view of the range of the phenomenon. A parallel was drawn between the Bangkok Thai variation (illustrated by very few data) and dialectal material (with more data) showing the same phonological variation. Then, drawing on Li's reconstruction of Proto-Tai initials in (5), it

was observed that all alternating words go back to a labio-velar fricative (or an aspirated stop, but recall the objections in the comment on (5b) above), which uniformly developed into a Proto-Southwestern-Tai fricative *x^w-, while, crucially, PT unaspirated labio-velar stops remained unaspirated labio-velar stops, PSWT *k^w-. The different developments were supported by the Chiang Mai data in (7). Based on these observations, it is concluded that the variations go back to a period where a PSWT fricative *x^w- turned into a labial fricative /f/. In other words, it is not the synchronic stop /k^{hw}-/ that alternates with /f/ in working class Bangkok Thai. This observation has not been pointed out in the literature so far.

The development can be represented in (8) below:

(8) Summary: developments from PT to the modern dialects

a) PT *x^w-,*γ^w-	> PSWT *x^w-	>	/f-/ in some Southwestern Tai varieties, including "working-class" Bangkok Thai
		>	/x^w-/ retained in Chiang Mai
		>	/k^{hw}-/ in Standard Thai
b) PT *k^w-, *g^w-	> PSWT *k^w-	>	/k^w-/ retained in all Southwestern Tai varieties
c) PT *f-, *v-	> PSWT *f-	>	/f-/ retained in all Southwestern Tai varieties

After discussing the view that changes essentially identical to the Bangkok Thai phenomenon are wide-spread across Tai dialects and can be shown to predate the modern variation, it remains to be seen what import it has on the interpretation of the Bangkok Thai variation itself. Most notable is the qustion of innovation or "how the variation got into Bangkok Thai". Above, historical evidence was presented that, as opposed to Beebe's (2001) view, the alternation is not an innovation at all. In fact, Beebe's own data refute the author's claim that the cluster reduction is an innovation in this case. In the study, all clusters of Standard Thai are examined: kl-, k^hl-, pl-, p^hl- (L in the table below), kr- k^hr-, pr-, p^hr-, tr- (R below) and k^w-, k^{hw}- (W below). Beebe (2001:30) presents the following statistics for cluster retention, that is where the clusters are not simplified:

(9) Cluster retention in Bangkok Thai speech, in percentages

cluster type	**age group**	
	<35	36<
R asp	9	16
R unasp	21	23
L asp	40	53
L unasp	39	55
W asp	73	79
W unasp	83	82

As can be seen from these percentages, the older generation of speakers, (36<), retains more clusters. This is why Beebe says that cluster simplification is an innovation. It can also be seen, and this was not pointed out by her at all, that /k^{w}-/ and /k^{hw}-/ behave significantly differently from other clusters in that they are much more often and much more consistently preserved than any of the other clusters in **both** age groups. Observe that these labio-velars are preserved in 73-83% of the cases, which does not compare with the 9-23% of kr-, pr, khr-, phr-clusters. In other words, /k^{w}-/ and /k^{hw}-/ clusters are not really affected by cluster simplification, but they are rather stable. Moreover, it is very likely that most of the k^{w}-/k^{hw}- reductions are to /f-/, rather than to /k-/ or /k^{h}-/ (although Beebe does not claim or mention this). These changes are then fundamentally different from any other reductions in the language where, for instance, /k^{h}l-/ reduces to /k^{h}-/. Methodologically, it is a serious objection against Beebe that k^{w}-/k^{hw}- are treated as clusters on par with clusters like pl-, k^{h}r-, etc. They are phonemes (like they were in Indo-European and other languages), not clusters–and in fact, Beebe's statistics support this view since k^{w}-/k^{hw}- do not behave like ordinary clusters: they do not simplify easily. In addition, the k^{w}-/k^{hw}- reductions are also different because they really do not indicate any innovation since there is not too much difference in the speech of younger and older speakers.

Since there is general agreement that the pronunciation of /f/ for /k^{hw}/ is characteristic of lower class speech, there are a number of possibilities for its origins in Bangkok Thai. One is that workers from surrounding areas had settled in Bangkok and they brought with them a speech that happened to be a dialect with variation (or **only** /f/ indeed!). Immigrant workers from other dialect backgrounds or the original working class inhabitants of the Bangkok region could speak non-alternating varieties, and the alternation spread later in lower working class speech. Another is that the original working class Bangkok speech already belonged to the group of dialects showing /f/, and it was upon this dialect that Standard Thai was imposed, which spread but imperfectly among working class speakers. To settle this matter would require an in-depth study of the social development of Bangkok, with special reference to immigration issues and Bangkok's rise to become capital of the Thai Kingdom in the late 18th century. This is beyond the scope of this paper, unfortunately.

It is hoped that the present discussion draws attention to the need for further sociological and sociolinguistic studies. It would be particularly welcome to see the range of the phenomenon in contemporary spoken Bangkok Thai: what other words show this alternation. From a phonological and historical point of view, all that has to be concluded is that the variation itself is far from being new to Tai dialects and that it can only appear in words that had an earlier labio-velar fricative *γ^{w}- or *x^{w}-.[8]

[8]I would like to thank my anonymous reviewers for the corrections and valuable suggestions.

REFERENCES

Beebe, Leslie M. 2001. "Initial consonant cluster reduction as a function of age group in Bangkok Thai speakers." In M R Kalaya Tingsabadh and Arthur S Abramson eds., *Essays in Tai Linguistics*, pp. 11-35. Bangkok: Chulalongkorn University Press.

Brown, J Marvin. 1979. *A.U.A Language Center Thai Course*. Reprinted 1997, 2001 by Southeast Asia Program Publications, Cornell University, Ithaca, New York.

Harris, Jimmy G. 1972. "Phonetic notes on some Siamese consonants." In Jimmy G Harris and Richard Noss eds., *Tai Phonetics and Phonology*, pp. 8-22. Bangkok: Central Institute of English Language, Mahidol University.

Henderson, Eugénie J A. 1985. "Feature shuffling in Southeast Asian languages." In Suriya Ratanakul, David Thomas and Suwilai Premsrirat eds., *SALS*, pp. 1-22. Bangkok: Mahidol University.

Henderson, Eugénie J A. 1987. "A phonetic oddity in Thai." In Jeremy H C S Davidson ed., *Laī Sū' Thai. Essays in honour of E.H.S. Simmonds*, pp. 52-62. London: SOAS, University of London.

Hudak, Thomas John (ed.). 1997. *William J. Gedney's Tai Dialect Studies. Glossaries, Texts, and Translations*. Michigan Papers on South and Southeast Asia, Number 45. The University of Michigan, Ann Arbor.

Jones, Robert B. 1965. "On the reconstruction of Proto-Thai." In G B Milner and Eugénie J A Henderson eds., *Lingua 14. IPLS. Part I. Historical Linguistics*, pp. 194-229. Amsterdam: North-Holland Publishing Company.

Li, Fang-Kuei. 1977. "A Handbook of Comparative Tai." *OL* Special Publication No.15. The University Press of Hawaii.

Smyth, David. 2002. *Thai. An Essential Grammar*. Routledge, London and New York.

Smyth, David. 2003. *Teach Yourself Thai*. Hodder Stoughton Educational, Hodder Headline Ltd., London.

Received: 22 July 2008

4. Affiliation Université Rennes 2, FRANCE
<huberd@freemail.hu>

Religious syncretism in healing non-communicable diseases: The role of folk healer monks

SAOWAPA Pornsiripongse
Mahidol University, Thailand

Abstract

Most societies, developed, developing or under-developed have three types of medical practices: popular, folk and professional. For Thai people, about 85 percent of their ailments are treated through the popular methods, 10 percent by folk healing, leaving 5 percent in the hands of modern health professionals. However, almost the entire health budget is spent on developing modern medicine and the professional sector. Hospitals and health centers with emphasis on modern medical science have been established all over the country. A large number of doctors and health providers continue to be trained to work in those health centers. The government provides not only the personnel, finances, and equipment but is also trying to change the paradigm from one of self sufficiency in health to one of dependence on doctors.

Nowadays, modern medicine is considered to be the norm for health care while popular and folk medicine offer alternative care. By using a case study of folk healer monks in the Central Thailand, this paper analyzes how the folk sector adapts and survives in the health arena where dependence on modern medicine has influenced people's ways of life even though health care coverage has expanded all over the country.

The study found that about 36 percent of patients, who visit the folk healer monks suffer from paralysis. The monks have adapted Lord Buddha's teaching to explain the cause of illness as *"previous Kamma"*. Therefore, the process of healing is an holistic approach, using physical, mental and spiritual methods.

Keywords: religious syncretism, non-communicable disease, folk healer monks.

Introduction

Approaching this study from a medical anthropological perspective, as mentioned above, most societies, developed, developing and under-developed, have pluralistic medical systems (Komart Chingsateinsab, 2006; Kleinman, A., 1980) featuring popular, folk and professional sectors. The popular sector is a self-care pattern whereby people use their indigenous knowledge, or their relatives', or neighbors' experiences to care for themselves, which costs little or nothing. Social networks and families have an influence on this sector. The folk sector is another health care pattern in which treatment is performed by folk healers. And lastly, health problems managed by medical doctors make up the professional sector, of modern or western

medicine. It is reported that in Thailand about 85 percent of sickness is treated by the popular sector, 10 percent of these illnesses which are not successfully treated by the popular level, goes to the folk sector and 5 percent health professionals, with severity, acuteness and accidents, is managed medicine. (Pennapa Subchareaon, et al, 2004)

Western medicine arrived in Thailand and firmly established itself in Thai society more than one hundred years ago. In those days, Thai traditional medicine or folk medicine were the main sources of health care. (Suvit Viboonponprasert and Komart Chingsateinsab, 1987, Saowapa Pornsiripongs, and Pornthip Usuparat, 1984). But nowadays, the professional (modern health) sector has become the mainstream health care with a large proportion of the health budget provided for building hospitals, health centers all over the country and training providers. So the numbers of health providers has increased while the number of folk sector healers has gradually decreased.

The objectives of this article are to find out: 1) Since Universal Health Care Coverage provides modern medicine free of charge, why do patients, especially chronic patients, still visit folk healer monks? 2) How do folk healers adapt their methods to cure non-communicable disease?

The article used a discourse analysis to explain how folk healers adapt their methods to the insufficiencies of modern medicine. Two famous folk healer monks in the central part of Thailand, in Nakonpathom and Ayuthaya province, were case studies. These monks are very popular and well known among paralysed patients. In-depth interviews were used to collect data from the healer monks and their patients (Chaai Potisita, 2004).

Development of Thai health care systems: a confrontation between modern and traditional medicine.

Thai traditional medicine has been the mainstream of the Thai health care system for a long time. Tracing back to the Ayudhaya period, Thai people, both commons and royal families, used herb, massage and magic for curing illness (Lalubere, 1976:280). Their health knowledge was handed down from their ancestors, and from observations and experiences. Skilled folk healers, called Royal doctors, were recruited all over the country to work for the royal court[1]. The remaining folk healers for the rest of the population were called Lay doctors[2]. Modern medicine came along with Portuguese merchants and missionaries in the reign of King Narayana but it was primarily used among foreigners.

[1]People from all over the country who were trained to be doctors since they were young, practiced with expert doctors by helping them to cure patients everywhere until they were skillful or talented healers to be appointed as Royal Doctors and serve for royal families and noble men.

[2]Folk healers who gained health knowledge from their ancestors and their own experiences, to help ordinary patients.

In the Ratanakosin Era, King Rama IV developed Thailand as a modernized country using a European model. Modern medicine, a symbol of modernization, was pushed through the Thai medical system but it was not well-known. In 1887, King Rama V established the first modern hospital[3] in Bangkok. When modern medicine was firstly introduced into Thai society, it was resisted by both the commoners and the nobles. Though the government used many strategies to change people's health care patterns from using folk to modern medicine, such as giving a set of birth appliances[4] to a mother who had delivered a child in the hospital, it took a long time to change Thai attitude and behavior to become familiar with modern treatment (Saowapa Pornsiripongse, and Pornthip Usuparat, 1984).

In 1888, the king supported Thai traditional and modern medicine curricula to be taught in Siriraj hospital. But due to different bodies of medicinal knowledge, concepts and styles of teaching, Thai traditional medicine remained separated from Siriraj hospital but was continued in traditional healers' homes, passed down to apprentices and through self-learning. Supported by the government and the Rockefeller Foundation, modern medical education was gradually strengthened while Thai traditional medicine had to be self supporting and slowly faded from the formal health system.

In 1936, the Health Professional License Act was passed. As a result, folk healers were declared illegal unless they were trained and passed the Thai traditional medicine examination. Only legal folk healers were allowed to treat patients. In 1942, the Ministry of Public Health was established. Since then, many hospitals were established and developed bit by bit all over the country. Then western medicine became the mainstream of the health care system while Thai traditional medicine was relegated to marginal treatment.

According to the 1977 Alma Ata declaration, the World Health Organization (WHO) encouraged the member countries to develop and promote folk and herbal medicine as a part of primary health care. As a result, traditional medicine was taken into the National Health Plan but confined only to herbal medicine, not folk healers. The turning point for traditional medicine happened in 1994 when the Institute of Thai Traditional Medicine was founded by the Ministry of Public Health and in 2002 the Department of Thai Traditional and Alternative Medicine was established. This is the revival of Thai traditional medicine in the formal health system again after having been separated for more than 100 years.

The Ministry of Public Health has tried to change the paradigm from one of self-sufficiency in health to that of dependence on doctors. "Do not cure yourself, consult the doctors whenever you have health problems" was the message sent to Thai people through all kinds of media for more than 30 years.

---

[3]Now Siriraj hospital is the most famous public hospital and medical school in Thailand.

[4]Necessary items for infants such as clothes, small and soft mattresses.

In 2001, the Thai government introduced a new policy on universal coverage of health insurance "Sam Sib Baht Ruksa Tuk Roke" or "Thirty baht covered all diseases" which changed to "Free for all diseases" in 2007. This policy encourages more dependence on health providers so now it is more convenient for patients to visit a medical doctor rather than a folk healer. However, many patients, especially those with chronic diseases, still visit folk healer monks even though they can afford medical doctors.

After it became separated from Siriraj hospital, traditional medicine was learned through self-study. Interested persons had to study with licensed healers, and practice with patients until they met the qualifications[5] for the Thai traditional medicinal examination. As a result, very few folk healers are able to pass the exam because almost all of them have poor education, and gain health knowledge from experience, caring for themselves and learning from neighbors and relatives. Their accumulated knowledge is different from that recognized in the health education system. Consequently there are far fewer folk healers with licenses than unlicensed healers. These healers can only help patients in the villages unofficially.

There are many kinds of folk healers such as herbal healers, bone healers, magic healers and massage healers. A folk healer monk is one who uses herbs and magic in his treatment. In the past, monks had to study folk medicine in order to take care of themselves as the first priority, and help villagers when requested, since modern medicine could be found only in the cities.

The role of folk healers was mentioned again when the AIDS epidemic hit Thailand in 1990 and could not be cured by modern medicine.[6] Folk medicine became an alternative form of care for those patients (Yongsak Tantepidok, 2005). Nowadays Thai people accept medical pluralism, using more herbal medicine and other kinds of alternative medicine.[7]

Economic and social development has been the pattern of health concerns shift from communicable to non-communicable diseases. The number of patients suffering from hypertension, heart disease and diabetes has increased and these diseases are now among the top five causes of death in Thailand (Ministry of Public Health, 2008). One effect of non-communicable

---

[5]Thai traditional medicine has three branches which are Medicine, Pharmacy and Midwifery. To apply for Medicine, persons have to study with the licensed healers for three years, and one year for Pharmacy and Midwifery.

[6]The AIDS epidemic crisis has made a great contribution to folk medicine because it reminds people of the limitations of modern medicine in conquering behavioral diseases. So, they have to share health space for other alternative medicines. At one time many folk healers, at Chiangmai, Lopburi and Kanchanaburi province, claimed that they had herbal medicine to control HIV virus.

[7]In Thai society, the mainstream of health care is modern medicine so other practices base themselves on traditional or pre-scientific understandings of medicine, folk knowledge, spiritual, metaphysical, or religious beliefs, or newly contrived approaches to healing called Alternative medicine. Alternative approaches are often used in conjunction with conventional medicine especially for chronic diseases.

disease, especially hypertension, is paralysis. It takes time and considerable expense to recover from this disease. So, many paralyzed patients seek alternative care and visit folk healers as one of their choices.

Folk healer monks and non-communicable disease healing: A case study

Folk healer monks

Two famous folk healer monks in two provinces in the central part of Thailand were studied. One monk is an abbot, 65 years old, and has been in the monkhood for 45 years. He received the Monk Developer Award from the Religious Department, the Ministry of Education in 1999. Knowledge of folk medicine was passed to him by his uncle. When his uncle passed away in 1972, he became a folk healer monk and eventually gained more experience. Now he is very famous, not only in the province, but also around the country. He attends patients every day in the morning from 10 to 12am and from 3 to 4pm in the afternoon.

Another folk healer monk who is now 73 years old, joined the monkhood as a novice when he was 13 years old. As a novice, he worked as a helper to a folk healer monk, preparing herbal remedies for the patients. After considering his personal characteristics (he was honest, a good man, not greedy and interested in folk medicine), the folk healer monk accepted him as an inheritor. He has been a folk healer monk now for 50 years. He is very famous for chronic diseases and paralysis. He attends patients every day in the morning, except Buddhist holy days.

Characteristics of the patients

The number of the patients seen by the monks is about 30-50 persons a day on weekdays and 70-100 persons a day on weekends. The patients come from every part of the country, Bangkok, central, north, northeast, east, and south, as shown in the table below.

Table 1. Shows where the patients who came from in a one-month period

Homeland	Number of patients (N=829)	Percentage
Bangkok	166	20.02
Central part	525 (from 19 provinces)	63.33
Northern part	68 (from 8 provinces)	8.21
Northeastern part	40 (from 13 provinces)	4.83
Eastern part	15 (from 1 provinces)	1.81
Southern part	15 (from 4 provinces)	1.81

Table 2. Shows sex of the patients who visit the folk healer monks in a one-month period

Homeland	Sex		Total (N=829)	percent	
	male	female		male	female
Bangkok	77	89	166	46.38	53.62
Central part	233	287	525	45.33	54.67
Northern part	23	45	68	33.82	66.18
Northeastern part	18	22	40	45.00	55.00
Eastern part	5	10	15	33.33	66.67
Southern part	6	9	15	40.00	60.00
Total	362	467	829	43.66	56.34

Table 3. Shows the ages of the patients who visit the folk healer monks in a one-month period

Age (years)	Number of the patients (N=829)	Percentage
1-9	7	0.84
10-19	14	1.69
20-29	44	5.31
30-39	83	10.01
40-49	156	18.82
50-59	228	27.50
60-69	177	21.35
70-79	106	12.79
>80	14	1.69

Table 4. Shows diseases and symptoms of the patients who visit the folk healer monks in a one-month period

Diseases and symptoms	Number of the patients (Total cases 829)	Percentage
Paralysis	305	36.79
Leg, knee, neck, and back pain	257	31.00
Diabetes	83	10.01
Menstrual fever	39	4.70
Gout	27	3.26
Lack of energy	18	2.17
Lack of appetite	16	1.93
Headaches	16	1.93
Hemorrhoids	14	1.69
Gastric intestinal problems	14	1.69
Spasms	14	1.69
Sinus	14	1.69
Cirrhosis	12	1.45

The number of patients visiting the folk healer monks during the one-month research period was 829 of which 66.33 percent were from 19 provinces in central Thailand, 20.02 percent from Bangkok, 8.21 percent from the north, 4.83 percent from the northeast and 1.81 percent from the south and the east. The ratio between male and female patients was not very different. That is 43.66 percent were male and 56.34 percent were female. The largest group, 82.15 percent, were aged between 40-80 years, of which 27.50 percent were aged between 50-59 years, 21.35 percent from 60-69 years, 18.82 percent from the 40-49 years and 12.79 percent in the 70-79 years group. That means that folk medicine is rather popular among the elders who are now confronted with non-communicable diseases. The statistics in table 3 confirm this fact. During the period of study, the top three diseases and symptoms were paralysis, muscle pain and diabetes. About 30 percent had visited 4-5 times in 1 year and 70 percent were new patients.

Healing process

Because of the number of patients each day, the folk healer monk organized and applied administrative procedures like those used in a modern hospital such as: issuing a member card for a new registration and patient cards; putting the sequence number and patient card on the monk's desk; providing wheel-chairs for disabled patients. But the difference from modern medicine is that every patient has to pay respect to *Cheewogakomarapat*, Lord Buddha's doctor by presenting an offering set consisting of cigarettes, plastic flowers, candles and joss sticks to the healers. These steps were not confusing because a temple announcer would constantly broadcast what to do.

The announcer also persuaded patients and their relatives to make other kinds of merit such as making merit for the birthday Buddha image. There were seven Buddha images representing each day of the week. People who were born on Sunday, could put money on the Sunday Buddha image. Sticking gold leaf on nine sacred stone balls, *"Loog Nimit"* was also popular. These nine balls would be buried to identify the area for a new Buddhist temple. Other offerings included donating money for the water and electricity used by the temple, and offerings dedicated to Buddhist monks, etc. Patients and their relatives were told that "donating by making merit can redeem karma which was the cause of illness". While donating, the patients had to pray and beg for good health and for their illness to be cured. Many patients took an oath that if they recovered from illness, they would make a lot of merit such as arranging movies, *Likay* (Thai local play) for the temple annual festival and buy *Loog Nimit* for the new temple.

Patients can easily give charity if they want. As the temple was established on a river bank, the abbot created a no-fishing zone where people can feed the fish.

Almost all of the patients who visited the folk healer monk, especially new patients, would make merit, give charity and take an oath to get well. After making merit, patients would wait for the healer who called each patient one by one. About 90 percent of the patients had already visited medical

doctors but did not get better. So, especially the paralyzed patients, tried herbal medicine and prayed to get well.

The healer explained two causes of paralysis, the behavioral cause and the "previous Karma' cause or one's past deeds, in order to reduce the patient's anxiety and help them to accept the fact that it takes a long time to recover. The causes of illness in the Thai world-view, are not only pathogen but also misfortune. Illness caused by pathogen can be cured by modern medicine while misfortune has to be eliminated by making merit and offerings. Some patients saw the effect of non-communicable diseases such as paralysis as caused by misfortune or former Karma. So, using only modern medicine could not help them recover a normal life. As a result, many patients with paralysis used both modern and alternative medicine.

The alternative medicine used by the folk healer monks consists of herbs, magic and mental healing. The process of healing is not only physical healing but also involves relieving the patient's mental condition. Physical ailments are cured with herbal medicine while making merit, charity, and prayers seeking the blessings of Lord Buddha, are the remedies for mind healing. When patients have made merit, they can redeem their Karma. If they intend to take herbal medicine regularly, they can get well soon. Patients believe that monks endowed the medicine with special healing powers, so the herbal medicine here is more than just medicine. After treatment, the healers give patients a small alms bowl as a piggy bank to collect money just one baht a day which they bring back to the healers on the annual temple festival day. The piggy bank is changed every year-sometimes it is a bamboo tube or a small tank. Recovered patients return these piggy banks during the three days of the annual temple festival period where lots of food and drink is available for everybody.

Conclusion and discussion

Why do many patients visit the folk healer monks when they can visit the hospital free charge? The answers are as follows:

Using holistic health care

While modern medicine focuses only on physical health, its efficacy can be measured objectively, so it is suitable for acute symptom or diseases caused by germs. But it cannot manage non-communicable or chronic diseases that take a long time to recover from. The monks realized the limitations of modern medicine so they focus on the strong points of folk medicine that concern physical and mental aspects. For mental support, the monks bring "Karma", a core belief of Buddhism, to explain the causes of paralysis.

It is accepted that chronic diseases take a long time to recover from and some cannot be cured. Almost all of the patients and their relatives realized this fact but never gave up. When the folk healer monk explained that the causes of illness correspond with the patients' belief, they chose folk

medicine as another alternative. After 1-2 visits and recovery, many patients stop using modern medicine.

By visiting the healers, they can both cure their illness and reduce their former *karma* or misfortune that affects recovery from paralysis, in a short time. According to Buddhism, physical health is the normal functioning of the body and organs. When one of the other fails to function effectively, the disease sets in. Disease can be treated with herbal medicine and mental health can be regained by making merit and giving charity.

Complicated administrative steps and length of time

There is one weak point in modern medicine which is not shared by folk medicine. Even though treatment at public hospitals is free, administrative steps are complicated and take a long time because of the large number of patients per day. Paralysed patients can neither help themselves nor wait for a long time. Contrary to this folk medicine, patients have only to see the healer in person the first time. In the event that they cannot visit again, their relatives, who are able to describe the symptoms can visit the healer on their behalf and get 2-3 bundles of herbal medicine for them.

Inexpensive treatment

Though visiting the folk healer monk does incure expenses, while going to hospitals under the universal coverage health insurance scheme[8] is free of charge, some patients, particularly those suffering paralysis decide to visit the folk healer because they know exactly what the cost would be. And they can generally afford treatment as it is cheap only 100 baht for a bundle of herbal medicine lasting for 15-20 days. If they don't have enough money, they pay as much as they can.

Facilities available for paralytic patients

As the folk healer monks are very well known among paralysed and disabled patients, they have prepared many facilities for those patients, for example wheel-chairs, pathways with slopes, toilets, disabled parking, etc.

The folk sector tries to fill the gaps in modern medicine such as its inability to cure non-communicable diseases, and its focus on physical illness with very few minutes of consultation. For these reasons, folk medicine continues to survive in a health arena where medicalization has influenced people's ways of life.

[8]People with no health insurance, such as social security, or government welfare, have to use universal coverage health insurance by choosing government hospitals in their residential area. Whenever they get sick, whether as out- patients or in- patients, they can use medical services free of charge.

REFERENCES

Chaai Potisita. 2004. *Science and Arts of Qualitative Research*. Bangkok: Ammarin Printing and Publishing.

Kleinman, A. 1980. *Patients and healers in the context of culture*. Berkeley: University of California Press.

Komart Chingsateinsab. 2006. *Medical Pluralism and Health in Social and Cultural Perspectives*. Bangkok: Princess Maha Chakri Sirindhon Anthropological Center.

Lalubere, D. 1967. *De Lalubere Memorandum* Vol. 1. Translated by Sun T. Komut. Bangkok: Kao Nha Publishing.

Ministry of Public Health. 2008. *The Survey of the Health Expenses of Thai People*, Nonthaburi: Health System Research Institute.

Pennapa Subchareaon, et al. 2004. "Herbal and Traditional Research" in *Strategic Plan for Health Research*, Revised Edition. Bangkok: Thailand National Research Council.

Saowapa Pornsiripongs, and Pornthip Usuparat. 1984. *Document and Knowledge Transfer of Thai traditional Medicine*. Bangkok: Institute of Thai Traditional Medicine, Ministry of Public health.

Shan States." Compiled from official papers by J. George Scott, assisted by J. P. Hardiman. Rangoon: Printed by the superintendent, Government printing, Burma, 1900-01.

Suvit Viboonponprasert and Komart Chingsateinsab. 1987. *Thai Traditional Medicine: Self Reliance Wisdom*. Bangkok: Doctor Magazine.

Yongsak Tantepidok. 2005. "Social Network of Folk Healers and Social Space Construction in Thai Health System." *JSS* 17(1) pp. 90-111.

Received: 30 October 2008

Research Institute for Languages
and Cultures of Asia, Mahidol University,
Salaya, Nakhon Pathom 73170
THAILAND
<lcsps@mahidol.ac.th>

BOOK REVIEWS

Review of *Korku Language: Grammar, Texts and Vocabulary*, by K.S. Nagaraja, Institute for the Study of Languages of Asia and Africa, Tokyo University of Foreign Studies, 1999, 336 pages.

Reviewed by Norman H. ZIDE
University of Chicago, USA

1. Introduction

Nagaraja has written the first fairly comprehensive grammar of the Korku language since Drake's work of 1903. It includes 172 pages of glossed texts which no previous linguist has made available. Anyone interested in the Korku language will have to consult it.

Korku is a North Munda language spoken in several non-contiguous areas in central India, in south central Madhya Pradesh and the adjacent parts of Maharashtra (Vidarbha). It is the westernmost of the Munda languages, and thus the westernmost language of the vast Austroasiatic phylum. The North Munda family branches into Korku and Kherwarian. Kherwarian includes the languages with by far the largest populations of any of the Munda languages. According to the 1981 Census (quoted by Nagaraja) there are more than 300,000 Korkus, most of them, presumably, (still) speaking Korku. The North Munda languages with sizable populations – more than 200,000 – (besides Korku these are Santali, Ho, and Mundari) have increased in population and in number of speakers. The languages with small populations (e.g. Korwa and Turi in North Munda, and apparently, all the South Munda languages in Orissa and northern Andhra, (big (Sora) and small (Gutob, Gorum (Parenga), etc.) have been losing speakers for some time. Arlene Zide (personal communication) noted that in 1963, speakers of Gorum of middle age (c 40) spoke the language natively, but were equally fluent in the Indo-Aryan Desia (Oriya), and many preferred to use that at home. The language of these speakers had many more lexical replacements than that of the older generation. Gregory Anderson (personal communication) reported that a group of young Gorum adults he met in 2005 knew no Gorum at all.

There are important aspects of Korku that are not treated in the grammar; for example sociolinguistic matters are not discussed. Fifty years ago everyone in the Melghat Korku villages spoke Korku as a first language, and most adults – almost all of the men – were bilingual (in some variety of

Hindi or Marathi – or both). One would like to know whether this is still true – and that the children speak Korku as their first, and in their early years, their only language – in the villages near Chikhaldara where Nagaraja's informants came from. Fifty years ago in Zide's Dharni (Amravati District) villages, Hindi songs had become popular at the expense of Korku songs, particularly among the men, but less so among the women.

Although four of the linguists have worked on Korku – Drake, Girard, Zide, and Nagaraja – it is unfortunate that none have done comparative work on the easternmost dialect of Korku, Mowasi (Mawasi). Based on the little in the Linguistic Survey of India, I anticipate that new and interesting material would come out of any examination of Mowasi. What little we know of Betul, and Hoshangabad Korku, and Mowasi is found in the old (1906) fourth volume of the Linguistic Survey of India. S. Bhattacharya did work briefly on several Korku dialects, including Mowasi (in 1958), but all there is of Mowasi in print that I have seen are a very few lexical attestations in his comparative vocabularies, and very brief mentions in his 'Studies in Comparative Munda Linguistics', but no (contrastive) description of it. Zide's few days of work with a (Lahi) Hoshangabad informant turned up notable differences, e.g. the complete absence of the dual, phonological differences (the predicator is /-ò/, not Melghat /-bà/), and some differences in verb morphology, e.g. the presence of a 'confirmative construction', and a few instances of subject suffix agreement in transitive verbs. (Nagaraja notes a few in his data.) D.S. Dwivedi, at one time teaching at Sagar, published two papers on Korku that Nagaraja lists in his bibliography, but I have been unable to see these papers. It is quite possible he worked on Hoshangabad or Betul Korku. There are interesting dialect chain-like features extending across all of North India, from northern Orissa to eastern Maharashtra (e.g. in the demonstratives).

The book covers segmental phonology, phonotactics, and morphology and basic syntax in an 'old-fashioned' (nineteen sixties) model of grammar, and provides texts for twelve tales totaling more than one hundred and fifty pages of texts, wide-spaced pages with interlinear word by word, and – usually – morpheme by morpheme translation as well, with a translation of each sentence, though the latter is unidiomatic and sometimes uninterpretable. Although both Girard and Zide recorded and transcribed texts, none of these texts was published, so we are grateful to Nagaraja for his extensive text material. Girard had the advantage of recording women speakers as well as men, something the others did not – and could not – do. The only other full treatment of Korku ('Kurku') grammar published is Drake's 1903 volume, which is difficult to find. Drake's grammar uses the traditional Latin and Greek grammatical terms and concepts. He knew the language well. His book is thorough and well organized. His is the least known and appreciated of the three substantial North Munda language grammars of the early twentieth century: Hoffman's for Mundari (1903) and Bodding's for Santali (1923, 1929) are the others; both Hoffmann and Bodding build on their predecessors' work. All three of these men were foreign missionaries, Drake English, Hoffmann German, and Bodding Norwegian, and they represented different Christian denominations.

The other, more recent and more limited, studies of topics in Korku grammar, those of Girard and Zide, were published in more restricted and less accessible forms: Zide's as an American dissertation, Girard's dictionary with an appendix on grammar as a mimeographed book in India. Neither is easily obtainable. Several papers by Zide were published in journals, festschrifts, and – for the numerals – a small monograph. One should also note Dwivedi's two papers – in Indian journals apparently unavailable in American libraries.

Nagaraja's study presents his own work, data collected and analyzed by himself in considerable detail. He has very little to say about the data and analyses of others, and some of what he does say is inaccurate. What he says about Pinnow's views on the passive is misleading, and when he says that Zide finds 'a lack of glottalized stops', this is not just mistaken, but the opposite of what Zide says and discusses in detail in his dissertation (which is listed in Nagaraja's bibliography). He does note that Zide finds tone in Korku, which he – Nagaraja – does not. This is important. See the section on phonology following this one. It is not clear that he has seen all the works mentioned in his bibliography. In his earlier bibliography he listed Korku papers that were announced but never written as if, in fact, they were extant papers.

Any comparisons, different analyses, comments, material supplementing and contrasting with the data and analyses of the other Korku scholars will therefore have to be made by others. (See below.) Nagaraja shows little interest in diachronic linguistics or sociolinguistics, although he includes as an appendix a list of Mundari cognates for about 300 Korku forms. His Korku has borrowed heavily from (mostly) Hindi, but Nagaraja does not, in his glossary, mark loanwords and indicate their sources. There is a great deal of influence of Hindi in morphology and syntax as well, but Nagaraja, although he is aware of this, only rarely mentions such borrowings.

This review discusses various topics under the heads Phonology, Morphology, and Language Change, and there are some miscellaneous remarks at the conclusion of the review.

2. Phonology

This reviewer was particularly interested in what Nagaraja had to say about phonology, his analyses being based on data from the two informants from a Chikhaldara village that he worked with. He – Zide – wrote a dissertation with detailed sections on phonetics and phonology. Nagajara's descriptions and analyses were disappointing. He lists the dissertation in his bibliography but it is not clear that he read it, and if he did, he has nothing to say about the detailed analyses there, many of them very different from his own. Except for one important matter, he did not find a tone distinction. (Zide found two tones, unmarked high, and low.) I was particularly interested in dialect differences in low-toned lexemes, but since Nagaraja looked for low tone and did not find – hear? – it, he can provide no such data. He suggests instrumental phonetic analysis would be a good idea. I agree. Nevertheless, I am fairly certain that Nagaraja is wrong to state that, for instance, verbs with

initial vowel lack overt (phonologically marked) reduplication. Reduplication is a complicated matter: in my dissertation I described it at some length – and for all (several) of my informants, a verb like (intransitive) /iraʔ/, 'to return', has as its reduplicated form (transitivizing it in this case) /iìraʔ/. Similarly /higra/, 'to fear, be afraid', and /hiìgra/.

For anyone interested in reconstructing Proto-Munda (or Proto-North Munda), writing – hearing – /Do(:)/[1] – which is equivalent to /Do/ – for 'put. keep', and not /Doò/ – equivalent to /Doho/ – prevents us from seeing the correspondence – the kinds of correspondences there are – roughly – between Korku /Doho/, Kherwarian /DOhO/ and South Munda /DVko/. The semantics – 'put, keep' – are borrowed from Hindi, and do not represent the older – Proto-Munda – meanings of */DVkO/.

There are many words in Korku, most of them loanwords from Hindi, with long – double – vowels, e.g. /raato/, 'night', where the double – long – (high tone) vowel sequence contrasts with the double high-low tone vowel sequence, as in the verb 'to return' noted above, and also in borrowings from Hindi with /h(V)/ in the second syllable, e.g. Hindi /peh(V)laa/, written *pahlaa*, 'first, previous', Korku /peèla/. So I question his transcription – hearing – of the forms where I find low tone, and of length as well. (He sometimes writes length for examples of both of the above-mentioned double vowel sequences, sometimes not.) And I have doubts about his recording of glottal stop as well. The correspondences between Girard, Zide, and Nagaraja with regard to glottal stop are not systematic, but look erratic. I am less sure of his mishearing here. There is variation in certain words as they occur in different dialects and idiolects, sometimes found with and sometimes without final glottal stop.

His transcription of juncture phenomena is also, I think, questionable. He writes as one unit what the other linguists write as two phonological units – words (e.g. Number Noun, Adjective Noun, Demonstrative Noun) – not following up the implications for the syntax of such analyses. I do not hear, e.g. /ini koro/, 'this man' as Nagaraja's /inikoro/, if word junctures are defined – marked – in the usual ways. I have many other disagreements with Nagaraja at this level, but – enough.

3. Morphology

Nagaraja dedicates his book to the memory of his teacher, H.S. Biligiri, but he lacks the interest and skill in morphological analysis of Biligiri (see, for instance, Biligiri's elegant analysis of the Sora verb). He will come up with dubious morphemes, e.g. the verb stem infixes – one example of each – that he identifies, and makes as much of an unsurprising assimilation, /jophe/ (Z /jophèʔ/ for the expected */jomkhèʔ/, – /jom-/ 'to eat', /-ki/ intensive mode, /-èʔ/ 'transitive past') – as he does for the interesting, unexpected – archaic – examples of /jom-/ taking, in the transitive past, *subject*- and not object-

[1]For dentals, capital letters indicated retroflexion.

marking pronominal suffixes. But he does find and identify such forms, unlikely though they seem. /jom-/ is the only verb that he has found exhibiting such suffixation, and he finds it only with the Mode suffix /-ña/ (from /-ya/?, or is this a different (and not noted elsewhere) Mode suffix which switches pronominal suffixation from object to subject?) and in the three forms [jom-ña-piñ-ba] 'will eat (dual)', presumably second person dual. [jom-ña-lañ-ba] 'will eat (pl.)' – i.e., first person inclusive plural, and [jom-ña-ki-mi-ba], 'will eat, probably first person'. Nagaraja finds /-mi-/, for other linguists only used for second person singular, used, usually, for first person singular as well. The sequence of two Mode suffixes (/-ña/ and /-ki/) in a verb form is rare, and ought to have been noted. Nagaraja avoids questions of morphemic identity and meaning – e.g. of /-ya-/ and /-ki/ in the different tense systems by not analyzing them out in the first place. But he does point out nice semantic distinctions between the morpheme complexes /-yen/ and /-ken/ (Zide [-ya-en] and [-ki-en]), the /-en/ being the intransitive past), the former being a more remote past, the latter a (very) recent past, or even indicating an action not completed (compare /kulyen/ and /kulken/, both 'sent').

The questionable phonetics noted under Phonology leads to wrong analyses in the morphology. Even if his glottal stop-less transcriptions do accurately represent what his informants say, a simpler analysis would show the /g/ (in his /-g-àʔ/ and /-g-èn/) as part of the noun stem, and not providing additional genitive /-àʔ/ or locative /-èn/ allomorphs. The genitive has the basic allomorph /-àʔ/, and the locative /-èn/ – no /g/ allomorphs – but there is vowel sandhi whereby the vowel in these affixes is lost – the previous noun-final vowel being retained; thus there are allomorphs /-ˋʔ/ and /-ˋn/, with, additionally, for the former, the glottal stop being lost in non-final (phrase-final or sentence-final) position, the allomorph /ˆ/ i.e. low tone, that being the only mark of the genitive in these cases, and that, too, is lost in the (comparatively small number of) nouns where the final syllable is already low. Nagaraja states that there is a zero genitive where its noun ends in a final vowel, i.e. the genitive is homonymous with the nominative. If in fact the low tone is lost – which seems possible but unlikely since this would lead to a genitive identical with the nominative in such cases, and other complications, although context would usually indicate the genitivity – then there is homonymy; otherwise – with the low tone not lost – the sandhi retracts the low tone – the indication of genitivity – to the previous noun-final vowel. Thus /dadu-àʔ/ 'Dadu (name)-genitive', 'Dadu's' > /dadùʔ/ which then can become /dadù/, the final glottal stop of the affix being lost nonfinally, i.e. phrase- or sentence-medially.

There are a few examples in Nagaraja (and Girard as well) of the locative for 'in the house' being /uran/ (Zide uràn ?) instead of the usual – older – /uragèn/. Presumably for some speakers 'house' is now /ura/ and no longer /uraʔ/, and the genitive is regularized to accord with the current form.

Missed as well are such allomorphic alternations – presumably there, but not heard – not distinguished – by Nagaraja as (from the verb /koòñ/, infinitive /kokhòñ/, 'to call') //koòñ-/, occurring before suffixes with initial vowel and /koñ-/, occurring before verb suffixes with initial consonants, but

the underlying blocked-in-initial-position low tone lowering the tone of the following verb suffixes. See similar misjudgments noted in the Phonology section concerning the reduplication of verb stems with initial vowels.

I do not mean to suggest that most of the morphology is not recognized, or is obscured by Nagaraja. It is not, and he does dig out e.g. reciprocal verb stems, and other interesting and comparatively obscure verb forms. Much of his material on negation – morphological and syntactic (as noted elsewhere in this review) – is not found elsewhere, and is fully presented and well analyzed. Some of the restricted sets of morphologically contrasting forms are well presented (e.g. the complex demonstratives), and the data are there for the reader wanting to make finer morphological distinctions. Much of the data, e.g. on the verb, is fairly fully presented without benefit – or need? – of much morphological analysis. We can see that /-ki/ and /-ya/ (with a rich variety of allomorphs for the latter not found in the other recorded dialects of Korku) are mode suffixes (M) and he gives examples showing how they are used, but he does not see that /-lì/, which figures in his /-lè/ (from [-lì-èʔ] 'cislocative transitive past', and /-wa, -va/ are also mode suffixes, if less frequent, perhaps, in his material that they are in Girard's and Zide's. He has enough examples of /-lì-/ with cislocative meaning (e.g. [i-lì] 'give (me, us)', [saʔ-lì] 'bring') but he decides /li ~ le/ is a verb occurring in double verb stems. He lists other – rather different – double verb stems with /li ~ le/, but he gives no evidence that /li ~ le/ as verb occurs anywhere but in this sort of complex – modal – verb form where it occurs in mode position. Drake also misses the role of /-lì/.

4. Language Change

4.1. hoy

The form *hoy*, found in Nagaraja's materials but unknown to Drake, Girard, and Zide, is a present tense verb form obviously borrowed from Hindi and used much like Hindi *hai* 'is', except that the Korku form is used for all persons and numbers. There are intransitive verbs in Korku, e.g. /taàkhaʔ/ 'to be located, to exist', that mark subject on the verb with pronominal suffixes (but only for 3rd person), but such person-marking on a borrowed verb form like *hoy* would be highly unlikely.

hoy is found in present tense adjectival and nominal predications in Nagaraja's Korku with some optional variation with 'zero'. If there is something systematic about the variation, I am not aware of it (about locatives see below). It occurs in such sentences as /ini uraʔ khad hoy/ 'this house is big', /dadu masTar hoy/ 'Dadu is a school teacher'. It looks as if *hoy* simply – neatly – replaced 'zero' in the earlier language. But all semantic adjectives in earlier Korku (Zide, unpublished notes – there is no trace of this in Nagaraja's material) were not of the same class. The class of 'simple adjectives' (A) like *khad* 'big', took 'zero' in such sentences, but a second class, 'adjectival verbs' (AV) like *simìl* 'sweet', took the predicator /-bà/, the Present-Future marker, in such sentences as /dii ambe simìlba/ [simìl-bà], 'those mangoes are sweet'.

For Nagaraja's informants all adjectives in these constructions behave identically: they take *hoy*. Thus for Nagaraja 'those mangoes are sweet' is /dii ambe simìl hoy/. Since both classes of Zide's adjectives – like all stative verbs – take the same – usual – verbal suffixes in the other tenses, the only place these were distinguished for Zide was in the present – Present for 'zero', but for /-bà/ 'Present-Future'.

Examples of adjectival verb forms not in Present or Present-Future: /khadjen/ [khad-ya-en], 'became big, was made big' (the past passive of kakhàd, 'to make big', similarly /khadken/, [khad-ki-en], which can be translated the same way, but /khadken/ is a more recent and/or intensive past). /simìlyen/, /simìlken/, translated similarly: 'became sweet, was made sweet', etc.

An interesting question to which we can give only a tentative answer is, which (semantic) adjectives fall in which class in Zide's data? Roughly, adjectives of size and color, and loanwords (e.g., /laNRa/ 'lame') are simple adjectives (A). The rest (/simìl/ 'sweet, tasty', /khamal/, 'heavy', etc.) are adjectival verbs (AV).

Locative predications are formed (for third person subjects) with Noun-èn (locative)-3rd animate subject pronoun, e.g. /di-ku ton-èn-ku/ 'Where are they?', [di-ku] 'they', 'that-animate plural', [ton-èn] 'where', from /ton ~ Ton-/ 'which', /-èn/ locative. Nagaraja sometimes finds *hoy* alternating with zero, e.g. /am je (hoy)/, 'Who are you?', /am/, 'you' singular, /je/, 'who'.

For the past the existing form, /Dan/ (from [Da-en]?) is used for 'was'. *Da-* can appear with intransitive suffixes, 'to be, become' and with transitive suffixes 'to make' (used more or less like Hindi *honaa/karnaa*). *Da-* is used as an auxiliary to form perfectives, etc like Hindi *honaa*, where the cognates of *Taàkha(?)* (see above) 'to be located, to exist' are used similarly in the other North Munda languages. *Taàkha?* in Korku is also used to translate 'have', with the English subject in the locative /Dij-èn uphùn kon-ku Taàkha-ku/, literally 'in him four sons are', 'he has four sons'. Another – related – common intransitive *ThaàR, ThanàR* – which does not take subject suffixes – means 'to remain, dwell'.

4.2. Negatives

Korku has two negative systems: the Present-Future and imperative being formed with the adverb/negative copula *(he)ban* / *(he)ban(n)-è?,* and the Past with the negative clitic copula *-Dùn*. In the Munda languages this system is peculiar to Korku. For a note on the North Munda – Kherwarian – language Ho and its negatives see below. The South Munda negatives are still more different.

The partial breakdown of the two-negatives system that we find in Nagaraja's grammar – the one for the Present-Future, and a different one for the Past, with no great symmetry between Past and Present-Future negatives – might seem to bring it closer to Hindi with its negation marking common to Past, Present, and Future, but the odd 'double negative system' in Nagaraja's

data is nothing like Hindi, or like any other language I am familiar with. The recent – see Nagaraja – apparent attrition of the old system introduces – confusingly – the Present-Future negative *(he)ban* into Past sentences, but these sentences are clearly Past, and (still) marked, primarily, by the Past Negative *-Dùn*. The introduced Present-Future negatives apparently have lost their tense meaning and now function as additional emphasizers – additional to the common emphatic suffix *-ka?*.

A reconsideration of 'double negative' sentences suggests that the double negation in Nagaraja's material (and found in none of the other grammars) is not an innovation but a retention. Earlier *ban* (found in Kherwarian with no tense-restricted meaning) must have had that tense-free meaning in early Korku, but with the later development of *-Dùn* (which has no cognates in Kherwarian) *(he) ban* became reserved for Present-Future meaning only, with the Past negative being marked by *-Dùn*. In most Melghat villages this distinction – contrast – was consistently made, but not in the villages from which Nagaraja's speakers come. There the *(he)ban* in sentences with 'double negative' marking retained the earlier and still familiar meaning of 'negative' with no tense implications, so that there was no 'contradiction' or confusion, which speakers of other dialects might now find in such 'doubly negative' Past sentences.

The past negative is marked by the clitic *-Dùn* following a bare verb stem. Past positive forms can provide information on voice (there are different past markers for transitive and intransitive), person (for object suffixes), and mode (intensive, cislocative, etc). All these are 'neutralized' in the past negative, where nothing but the bare verb stem can precede *-Dùn*, the additional information to be indicated externally – externally to the verb form.

The Present-Future negative is *ban*, which occurs in the adverbial *(he)ban* either preceding or following a verb form, usually the infinitive, and the negative copula *(he)ban(n)-e?*, which occurs sentence-finally. (Korku is a SOV language.) The prohibitive 'don't (V)' is *ba-ki*, from *ban* and the intensive mode suffix *-ki. baki* precedes an infinitive and a Verbal-Object for transitive verbs. It precedes the intransitive infinitive V*-ù?,* the neutralized form of the three *-ù?* forms. The positive Present-Future is richer, morphologically, than the positive Past, and the negative Present-Future is more restricted; it is reduced to the following forms: the (active) infinitive, reduplicated for transitive verbs, and an intransitive infinitive, V*-ù?*, which neutralizes three different forms with *-ù?:* /-ù?/, /-yù?/, [-ya-ù?], and /-khù?/ [-ki-ù?]. (There are problems and dialect differences with regard to /-khù?/.) For some speakers, apparently verb + object occurs with *(he)ban* and *(he)ban(n)-è?*.

The 'double negatives' in Nagaraja's material add *(he)ban* to past sentences with *-Dùn*. These sentences have past meaning. Just what these 'doubly negative' past sentences mean, i.e. how they differ in their emphasizing from ordinary past-plus-emphatic-*ka?* sentences, is not clear.

Examples of negative usage in Korku with a few in other Munda languages.

Positive Present-Future forms: /kukulbà/ [REDUP-kul-bà], 'sends, will send (customary)', [kul-miñ-bà], 'sends, will send you, dual'; [kul-lì-bà], 'sends, I will send here/this way' (/-lì/ cislocative, inanimate subject(s) unmarked); /kulùba/, 'is being / will be sent'. The negatives for the transitive forms are /(he)ban-ku-kul/, /ku-kul (he)ban/, and /ku-kul (he)bannè?/. For the intransitives: /(he)ban kulù?/, /kulù? (he)ban/, and /kulù? (he)bannè?/.

The [-è?] verbalizer is homonymous with past transitive /-è?/, but not to be identified with it. It is probably to be identified with the 'exclamatory' verbalizer found with such demonstrative derivatives as [hu-ku-è?] 'There they are!' ([hu-] 'that, yonder', [-ku], plural animate, [-è?]), and /naàne?/ 'Here it/they (inanimate) is/are!' from [na-in-è?] (/na/ 'emphatic before /in/', /in/ 'this'). Inanimates are not marked for number.

Past positive forms: /kulkhè?/ [kul-ki-è?], 'send-intensive, transitive past', /kullenej/ [kul-lì-è?-ej], 'send, cislocative, him-her', /kulyen/ [kul-ya-en], /kulken/ [kul-ki-en], both 'was sent', the latter 'recently and intensively', the former a more anterior past; /ira?ken/, /ira?/ as an intransitive stem, 'to return', as a transitive stem 'to return something/someone', /ira?ken/ 'returned, was returned'. Negatives of these forms are: /kulDùn/, 'did not send, was not sent (objects, direction, voice neutralized), /ira?Dùn/, 'did not return, did not return something(s)/someone(s), was not returned, etc'.

Examples of double negatives (from Nagaraja), partly retranscribed: /Dij heba giTij-Dun/ (Zide /Dij heban giTij-Dùn/) 'He did not sleep'. Nagaraja writes "when *heba* is used, past marker *Du(n)* can also be used". I assume he means that when *Dùn* is used the sentence is past; if not, not. /Dij gITij-Dùn/, 'He/she did not sleep', occurs in Nagaraja as well as the other grammars.

On Ho negatives: *ban*, cognate with the Korku Present-Future negative, is used (see Deeney), meaning 'to not be', not, as in Korku, as a negative 'adverb' – roughly 'not' – with other verbs. Ho, like the other Kherwarian languages, and unlike Korku, has a positive copula, /mena?/. 'To have' is translated with the copula, positive or negative, the (English) subject taking /ta?-re/, /-re/ being the locative 'in'. (Compare the Korku, also using the locative (/-èn/) with a verb /Taàkha(?)/ meaning 'to be (located)', which has a Ho cognate /taikena/ that is used in have-constructions in the Ho Past.)

For the prohibitive Ho uses /alo/ plus pronoun plus a positive verb form, e.g. /alo-ben seno?wa/ ([ben], 'you, dual') 'don't you two go!'. [ka] – subject pronoun is used – suffixed – to /ka/, /alo/, and other words preceding the main verb form – in negative statements. This sort of morphosyntax, characteristic of and important in Kherwarian, has no parallel in Korku.

North Munda, if it marks subject on the verb, does so with suffixes. Probably Proto-Munda did this with prefixes (we can still see this in the South Munda languages Gorum and Gta?). Non-singular 3rd person markers on the

verb (number suffixes, dual or plural) may be old in Munda to mark 3rd person non-singular, singular being unmarked. As noted in the discussion of *hoy*, Korku lacks – has lost – subject suffixation in all but a few intransitive verbs, locatives and a few odd – irregular – others (transitives).

'Not-yet-ness', in Korku /atikà?/, Nagaraja /aThika/, plus infinitive, is similarly expressed in Ho with /auri/. It is expressed in the (South Munda) Gutob language with a separate auxiliary verb, /oroj/, roughly 'to not yet V', versus the regular 'to not V' with /ura?/.

4.3. Relative clauses and relative clause equivalents

The problems in identifying and describing changes in the repertory and uses of relative clauses and relative clause equivalents are different from those discussed above. Here there are alternative constructions, but to understand which are used when and how we need the sort of data that we lack: on discourse and socio-linguistic contexts and distinctions. For instance, the use of the borrowed relative-correlative structure seems to be correlated with the education – literacy – of the informants, and the sorts of discourse – literary, in some sense – they are commonly used for elsewhere. Our data consist of tales (all of Nagaraja's texts, some of Girard's and Zide's), personal narratives, and sentences elicited for the purpose of finding out their Korku equivalents. All of the linguists are eliciting in standard Hindi, so that it is no surprise that the relative correlative calque on the Hindi is the usual translation – initial translation – of such sentences. But we find them rarely elsewhere. This calque equivalent is not newly invented; it is familiar to most – but not all – speakers. (Miss Girard had the advantage of being able to work with female speakers, which none of the others could do, but their texts are no different – i.e. in sticking to simple sentences – than the texts from her male informants.)

In Nagaraja's texts and Girard's texts – the circa fifty pages I checked – (Drake did not publish texts) there are no relative-correlative clauses. What we find are two simple sentences connected with the usual anaphoric devices to provide the sort of information found in complex relative sentences.

Despite this, and with the knowledge of this structure becoming more widespread, these relative-correlative sentences are much rarer for Nagaraja's informants (to the extent one can judge this from his grammar) than I would have expected from my experience and data of others.

Formal communications were not written in Korku. Has this changed? The carving out of a more or less 'tribal' state, Jharkhand, from what was greater Bihar, and the resulting use of Santali, Mundari, etc., in primary education was a major event, but Santali and Mundari had been used for 'literary' purposes earlier, while Korku presumably was not, certainly not to that degree. As of the nineteen-fifties the Korku I knew were unaware of speakers of related languages in Bihar, or even of Korku dialects a hundred and fifty miles away. Are there now magazines, fiction, etc. in Korku? Is the language acceptable in the courts or in other official venues? I do not know.

The Bible translations were prepared before the relative-correlative structure was borrowed. I have not been able to examine them for participial relative usage, so all I can say is Zide and Nagaraja (and presumably Girard, but her grammatical notes do not go that far) know such structures.

The relative-correlative construction uses as relative pronoun (in Nagaraja) *jo* – like the Hindi – and (in Zide) *je* 'who', and uses as correlative *Dij* 'he/she'.

Nagaraja's example:

/jo kitab Tebalaliñen Doken Dikitab iña hoy/

[jo kitab Tebal-àʔ-liñ-èn
which book table-GEN-on.(top.of)-LOC

Doò-ki-en Di-e kitab iñ-àʔ hoy]
put INTENS.REC-INTR.PST that-INAN book I-GEN is
'Which book he put on the table (that book) is mine.'

Zide's equivalent: /je kitab Tebalàliñen Doòken Dii kitab iñ-àʔ/.

Zide observed, not so much in Korku, but more in the South Munda language Gutob, that the ostentatious use of the Indo-Aryan calque (in the Gutob case based on either Desia Oriya or Standard Oriya, usually the former) had the function of showing the linguist – a 'sahib' of some sort – that the informant was educated, did know Oriya. This function was more common – needed – for Gutob speakers (then, c 1975) than it was for Korku speakers.

The relative and relative equivalent structures noted then are the separate – nonrelative – sentences, the Indo-Aryan-like relative-correlative calque, and – more interestingly – the participial relative constructions. For the sentence, 'the teacher who beat the boy went to Amravati' the adjectival form would be '(the) (he) beat the boy teacher' and the nominalized – 'one' (*waalaa*) – form would be '(the) (he) beat the boy one'.

In the former we have a verb form with tense and animate object following the object with accusative suffix:

/poerakhèʔ muNDàkhenej masTar amraotìn olen/

[poera-khèʔ muNDà-ki-è(n)-ej
boy-ACC beat-INTENS-TR.PST-3SG.ANIM

masTar amraoti-èn ol-en]
teacher Amravati-LOC go-INTR.PST
'The teacher who beat the boy went to Amravati.'

In the latter:

/poerakhèʔ muNDàkhenej-miThaj Dii masTar amraotìn olen/

[poera-khèʔ muNDà-ki-èʔ-(n)-ej-mi(n)-Thà-ej
boy-ACC beat-INTENS-TR.PST-3SG.ANIM-one-NOM-3SG.ANIM

Dii masTar amraoti-èn ol-en]
that teacher Amravati-LOC go-INTR.PST
'The one who beat the boy (that teacher) went to Amravati.'

Further examples from Drake:

[am-àʔ bhagia iñ-àʔ-anTi-èn
you.SG-GEN servant I-GEN-for-LOC

saʔ-lì-èʔ-mi(n)-e aol ban]
bring-CISLOC-TR.PST-one-(INAN) good not
'The one your servant brought for me is not good.'

[am ghuRgi-khèʔ i-wa-ej-bà
you horse-ACC give-BEN-3SG.ANIM-PRES/FUT

jhaRa ton-èn Taàkha]
fodder which-LOC is.located.INAN
'Where is the fodder you (will) give the horse?'

This construction is noted just once in Nagaraja, but both Zide and Drake find these fairly commonly, so that I assume they are old and still in use. (Girard's grammatical appendix to her dictionary does not get to such topics.) This sort of construction is not rare in the languages of the world (and is found in Dravidian, among other language groups). I suggest that some educated speakers – e.g. my tutor and chief informant – have begun to use this participial adjective construction more frequently in formal literary discourse, i.e. to use this and not the relative-correlative construction.

A number of new scripts were devised for the various Munda languages in the first half of the twentieth century. Two of these have survived and flourished: the Ol Ceme't or Ol Ciki script of Raghunath Murmu for Santali, and the VaraN Kshiti script of Lako Bodra for Ho. (Not that all Santals or all Hos have adopted these scripts.) Korku – like Mundari and Kharia – has been satisfied to use the Devanagari script, the standard script for Hindi. The introduction of the new Santali and Ho scripts came with a great deal of cultural 'enlargement', and the increasingly promoted literacy involved some attempts at standardization of the language, e.g. with regard to relative clauses. Korku has no such complications.

Here we have a different problem, and one that the data do not allow us to get very far in answering: the factors determining the choice among three possibilities. Why and when do which speakers use non-relative – two-sentence – grammar, and, when relatives are selected, when is the borrowed

relative-correlative used and when the participial adjective? And this presupposes that we are talking of well-formed complete sentences, since there are alternative ways of conveying the information. What we need to say requires us to take a closer look at our text materials and how they were acquired. To my surprise, all of Nagaraja's texts (all older tales), and Girard's texts – and these included personal narratives, and came from Korku women as well as men – did not use relatives at all. So we ask, who needs relative constructions in Korku? What are they good for? Where do the examples we have (in Drake and Zide) come from? How were they elicited? An obvious source of bias is that the language (and context) of elicitation was standard Hindi (perhaps Marathi for Nagaraja), and if one asks for a translation of the standard Hindi relative clause – a relative-correlative construction – what one will get in Korku – at least at first – is a calque on that. That seems to the informant to be what is being asked for. Then the linguist asks "Is there any other way you can translate that?", and the skilled, cooperative informant is likely to think that a complex sentence – not two simple sentences – is being asked for, so if he comes up with any alternative, he gives Sahibji the participial adjective. What sort of 'natural language' is this? Lots of free conversation (Zide has some of this) would help, as would the elicitation of more 'formal' – literary – material. One would like to see a translation of the sort of request that a Korku literate in Hindi would have written in Korku. How would it come out in Korku? Presumably more formal-literary – Korku is more in use now that it was forty years ago, but I have no such newer materials, and Nagaraja does not give us any. The establishment of Jharkhand, a separate 'tribal' state in what used to be south Bihar, and the official recognition of Santali-Mundari-Ho as a 'national language,' the use of these languages in primary education, and the offering of these languages as subjects at Ranchi University have made a difference, and those developments would have become known to some educated Korkus. Earlier there had been publications – stories, magazines – in Santali and Mundari, and these languages were clearly far ahead of Korku in the development of 'literary' styles. So without much information on all this I cannot say much about what motivates the selection of the (old) participial adjective construction versus the (borrowed) relative-correlative construction. My surmise is that the one or two of the literate Korku I knew who would have found a 'literary style' useful would have preferred the – native – participial adjective to the borrowed relative-correlative construction, But there could be others who would make the opposite choice. The sort of extensive oral (recorded) corpora and written materials that Anderson (personal communication, 2005) has collected for Ho should be revealing in telling us about relative clauses and their information-providing non-relative equivalents in North Munda.

5. Conclusions

Nagaraja's book is full of fascinating material, but some of it is so confusingly presented that a serious reader needs to rewrite the book to make full use of that material. One must extract grammatical information from the texts to expand and revise the earlier grammatical presentation. To do this, it

helps to know a fair amount of Korku (and devote a great deal of time to the work). Unfortunately, this limits the utility of the book.

The Tokyo Institute supported Nagaraja's write-up, and put out – printed and bound – an elegant book, perhaps the most attractively printed grammar of a Munda language I know. What they should have done as well was to provide Dr. Nagaraja with a good editor, an editor who worked through and improved the whole book, not just some of its sections. (The brief sections on compound sentences and complex sentences are well edited. The exposition here is clear, concise and in good English, and the examples – Korku sentences – well chosen.) But many of the sections of the book were poorly edited, e.g. various of the example sentences are not translated, fully or at all, and without some help in a number of these cases it is not clear what the sentences could mean and what the grammatical points being made are. (See the section on imperatives for this sort of mess.) Much of the exposition is unclear, and would have benefited from rewriting. Some of it is incoherent (again, see the section on imperatives). Any close reader of the book wants – what are lacking – full indices, morpheme indices as well as topic indices. I would have been saved hours and no little irritation if 'pronominal verb suffixation, object agreement, subject agreement' and some other key topics had been indexed.

The English (starting with the title) is marred by numerous mistakes. Most of these are transparent; one understands what is meant. Some are not. Thus, when Nagaraja translates /Dij khadjen/ (Zide's transcription) 'he became young', it is not Celtic-like myths one must think of, but Hindi and Indian English translation. In Hindi *javaan* does usually mean 'young', but *vah javaan ho gayaa* does not mean 'he became young' – a good editor would have questioned such a definition – but 'he became mature, grown up'. Apparently /kad-khad/ – 'big', in Korku – as used by Nagaraja's informants, not mine – invites similar misunderstanding. The reader must retranslate /Dij khadjen/ from Nagaraja's Indian English 'he became young' (here, literally 'big') into standard English 'he became mature, grown up'. In some of the texts Nagaraja gives up on English and just gives us the Hindi as the English translation: /kula mama/ in English should be 'Uncle Tiger', not 'kula mama'; /re/, a vocative particle, can only be usefully translated 're' if he tells us in English something about *re*, et cetera.

We should be grateful to the Tokyo Institute for underwriting Nagaraja's work in Japan, and paying for the publication of the volume. Nagaraja put a good deal of work into the book. Probably there would be NO book without the support of the Institute, but if they had extended their assistance to providing good editing it would have been a better book.

6. Minor Cavils.

The brief – perfunctory – definitions in the glossary should have been made fuller and more accurate. (Compare the much better definitions in Girard's dictionary.) The item glossed can be a verb stem, reduplicated verb infinitive, non-reduplicated verb infinitive, bimorphemic complex; none of

these is labeled or analyzed, and the reader cannot know what he is getting unless he knows a fair amount of Korku, some but not all of that knowledge coming out of Nagaraja's grammar. To translate /Taàkha(ʔ)/ as 'have' is not completely wrong, but is – to say the least – misleading. /Taàkha(ʔ)/ means 'to be (located)'. The dubious justification of it as 'have' would be that to say, for instance, 'he has three sons' one says /Dij-èn aphài kon-ku Taàkha-ku/, literally 'in him three sons are'. Is the reader supposed to know or intuit this? A close reading of some of Nagaraja's texts would provide the information.

Nagaraja does not indicate which words in the glossary are loanwords – there are many – or their sources. That information would have been useful, particularly for those words borrowed from Dravidian (e.g. /bo/ 'let's go', /poTTa/ 'boy'), which Nagaraja is particularly well qualified to trace. The more than 300 Mundari cognates are useful and appropriate, but ideally the source might have been the *Encyclopaedia Mundarica*, much fuller and more accurate than Bhaduri.

Since this book would be the natural place to look for a full Korku bibliography, the bibliography should have been more complete. The booklets on Korku put out by the Madhya Pradesh government should have been listed, and the Bible translations and some of the ephemeral publications that Nagaraja listed in his Austroasiatic bibliography should also have been listed. Why Nagaraja's own earlier bibliography (which includes useful annotations on many Korku listings) is omitted here is a mystery. Several papers of mine might have been listed too – and made use of in the text, e.g. Zide (1968) on pronouns, Zide (1978) on numerals, Zide (1991) on dictionaries, the (early) Zide data on the verb in Pinnow (1966), which is listed, and, more important, Zide (1991) on demonstratives.

Nagaraja's texts, all tales, are much appreciated, and contain a good deal of dialogue, but we could have used some real (recorded) conversations, too. The reader gets some rough idea of discourse properties from the texts in the book, but a greater variety of text types would have added to these ideas. Certain sentence types, e.g. the Hindi calque relative clause, could have used some comment on, e.g. the limited discourse contexts (I assume) it occurred in, in Nagaraja's elicitations. That such constructions are not found in the sort of tales he gives us in his texts is not surprising.

REFERENCES

Bhaduri, M.B. 1931. *A Mundari-English Dictionary.* Calcutta: Calcutta University Press.

Bhattacharya, Sudhibhusan. 1975. *Studies in Comparative Munda Linguistics*. Simla: Indian Institute of Advanced Studies..

Biligiri, H.S. 1965. "The Sora verb, a restricted study." *Indo-Pacific Linguistic Studies II* 231-250.

Bodding, Paul Otto. 1923-29. *Materials for a Santali Grammar.* 2 vols. Benagaria: Santal Mission of the Northern Churches.

Deeney, John. 1975. *Ho Grammar and Vocabulary.* Chaibasa: Xavier Ho Publications.

Drake, John. 1903. *A Grammar of the Kurku Language.* Calcutta: Baptist Mission Press.

Dwivedi, D.S. 1961. "Korku noun morphology." *Madhyabharati* 10(10).

Dwivedi, D.S. 1963-4. "Korku pronouns." *Madhyabharati* 11-13:89-92.

Girard, Beryl. 1965. *Korku-Hindi-English Dictionary.* Ramakhera and Derhtalai, Nimar District, Madhya Pradesh.

Hoffmann, John (Johann). 1903. *Mundari Grammar.* Calcutta: Bengal Secretariat Press.

Hoffmann, John and Arthur van Emelen. 1928-1978. *Encyclopaedia Mundarica,* 16 vols, Patna: Government Superintendent Printing. [Reprint 1990: New Delhi: Gyan Publishing House.]

Konow, Sten. 1906. "Kurku." In Grierson, George A., ed., *Linguistic Survey of India, vol. 4: Munda and Dravidian languages,* pp. 423-433. Calcutta: Superintendent, Government Printing. [Reprint, New Delhi: Motilal Banarsidass, 1967, 1973.]

Nagaraja, K.S. 1989. *Austroasiatic Languages,* A Linguistic Bibliography. Pune: Deccan College.

Pinnow, Heinz-Jürgen. 1966. "A comparative study of the verb in the Munda languages." In Zide, Norman H., ed., *Studies in Comparative Austroasiatic Linguistics,* 96-193. London: Mouton & Co.

Zide, Norman H. 1960. *Korku: Phonology and Morphophonemics.* University of Pennsylvania PhD dissertation.

Zide, Norman H. 1968. "Some comparative notes on Gataʔ pronouns." In Heesterman, J.C. et al., eds., *Pratidanam: Indian, Iranian, and Indo-European Studies presented to Franciscus Bernardus Jacobus Kuiper on his sixtieth birthday,* pp. 348-358. (Janua Linguarum, series maior, 64.) The Hague: Mouton.

Zide, Norman H. 1978. *Studies in the Munda Numerals.* Mysore: Central Institute of Indian Languages.

Zide, Norman, H. 1991a. "Possible Dravidian sources of some Munda demonstrative bases." In Lakshmi Bai, B. and B. Ramakrishna Reddy, eds., *Studies in Dravidian and General Linguistics (Festschrift for Bhadriraju Krishnamurti),* pp. 349-364. (Osmania University Publications in Linguistics 6.) Hyderabad: Osmania University.

Zide, Norman H. 1991b. "Lexicography of other languages of the Indian subcontinent: the Munda Languages." In Hausmann, Franz J., et al., eds, *Wörterbücher, vol. 3,* pp. 2533-2547. (Handbücher zur Sprach- und Kommunikationswissenschaft 5.) Berlin: de Gruyter.

Received: 10 September 2009

Department of South Asia Languages
and Civilizations, University of Chicago
5801 South Ellis Ave., Chicago, IL 60637
USA
<n-zide@uchicago.edu>

****

CONFERENCE REPORTS

The 3nd SSEASR Conference, 3-6 June 2009, Bali

Reported by Amarjiva LOCHAN
Delhi University, India

Some 567 participants from 63 countries showed interest in the conference, discussing the relationship between water, culture, and religion. The 3rd South and Southeast Asia Association for the Study of Culture and Religion (SSEASR) Conference was auspiciously held at Denpasar, Bali/Indonesia, and successfully organised by the Institute Seni Indonesia (ISI) and Universitas Hindu Indonesia (UNHI) on June 3-6, 2009. There were some late cancellations because of the current swine flu epidemic, but this didn't dampen the jovial atmosphere during the conference.

The Head of the Organising Committee, Professor I Wayan Rai, invited experts from around the world to talk about contemporary problems and to find solutions. The conference in Bali was seen as a follow-up to a conference held in 2007 at Bangkok/Thailand.

SSEASR's President, Dr. Amarjiva Lochan, said that people need to understand not only the role of culture and religion as foundation for life, but water as an important element for life in the universe. "We want to touch on religion and culture respecting water, so there will be awareness to preserve it," he said.

Professor Wang Gungwu, Chairman of the East Asian Institute, National University of Singapore, gave the inaugural keynote speech, talking about cultural diffusion and the inter-ocean exchange in the past and present. He postulated that the open seas should stay open and not become stages for future conflicts.

Academic sessions followed, such as about seas, ports, isles and the spreading of religion, ethnicity and religion, and culture and religion along the rivers, among others. Thus, Dr. Paiboon Suthasupa, North Chiang Mai University, and National Research Scholar, Mr. Thanit Wongprasert, introduced Thailand's annual Loi Krathong festival.

Also, on the second day, Professor Ms. Sophana Srichampa, from the Institute of Language and Culture for Rural Development, Mahidol University, Nakhon Pathom, Thailand, explained the symbolism of "dok bua" - the water

flower in Thai culture. In a session about rituals, Dr. Anthony R. Walker, from the University of Brunei Darussalam, went deep into the ritual water practices of the Tibeto-Burman speaking Lahu people of the Yunnan-Northern Southeast Asia borderlands.

Furthermore, the sacred Ganges in India and the mysterious Mekong River in Southeast Asia dominated some other presentations.

The belief that conferences are mainly held because of "networking" is reason enough to continue with this old academic tradition – even in the wake of the world financial crisis. Participants meet people they have never heard of before and to get an impression of who is who in the ivory towers of science and higher learning.

In this context, it was very enlightening to hear the plenary address of Professor Emeritus Frits Staal, University of California at Berkeley, California/USA on the theme of River Foursomes, touching on the worldviews of ancient India and Mesopotamia. From Europe, Professor Dr. Manfred Hutter from the University of Bonn, Institute of Oriental & Asian Studies, Department for Comparative Religion, talked about Christians from South and Southeast Asia in Germany. Prof Rosalind Hackett, the President of IAHR, was also present and chaired a distinguished panel on the role of water and rituals in other world civilisations. Prof Morny Joy, the EC Member of the IAHR, came with two significant panels on Women and Religion which was highlighted in the local media.

Professor Ms. Yolotl Gonzales Torres from Mexico City gave another fascinating talk about temple-mountains in ancient Cambodia and Meso-America. On the third day of the conference, she was also responsible for presenting a bold and rare film from the 1930s by Miguel Covarrubias about his "Island of Bali" - still the most authoritative piece of ethnological research on Balinese culture.

Cultural "gamelan" and dance performances signalled the end of a busy conference, while a cultural afternoon tour on June 6 to the King's Palace of Gianyar and a farewell dinner at Ubud, the cultural heart of Bali, officially ended the event in style.

Finally, there was the important announcement that the future 4th SSEASR Conference will be held in the Kingdom of Bhutan in 2011.

Received: 9 July 2009

Delhi University
INDIA
<amarjiva@email.com>

****

The 19th Annual Meeting of the Southeast Asian Linguistics Society

Organized by the Southeast Asian Linguistics Society and hosted by the University of Social Sciences and Humanities (Vietnam National University in Ho Chi Minh City)

Reported by Dinh Lu Giang
University of Social Sciences and Humanities, Vietnam National University in Ho Chi Minh City, Vietnam

The 19th Annual Meeting of the Southeast Asian Linguistics Society took place on 28, 29 May, 2009

Southeast Asia is an ethnically and linguistically diverse area that hosts a wealth of language contact. Research on Southeast Asian languages informs us about the history, anthropology and culture of the area. For the past 18 years, SEALS has tried to foster and encourage this research. In May 2009, the 19th Annual Meeting of the Southeast Asian Linguistics Society attracted about 120 participants from many countries to Ho Chi Minh City (formerly Saigon), the city where the first three volumes of Mon–Khmer Studies were published some four decades ago. The conference received 112 abstracts and after evaluations by an international panel of linguists, 63 abstracts were selected for oral presentation and another 14 abstracts for poster presentation.

The opening ceremony featured an opening remark from Prof. Vo Van Sen, the rector of the host University. Prof. Vo Van Sen emphasized how honored USSH was to welcome international linguists and called for a greater development of linguistic research in Vietnam.

The tone of the conference was set by two keynote presentations

• Dr. Paul Sidwell, Managing editor of Journal of Southeast Asian Linguistics Society and Director of Mon-Khmer Studies Project, argued for a reanalysis of Mon-Khmer subgroupings, proposing a flat structure with little internal organization, in a talk entitled “The Austroasiatic Central Riverine Hypothesis”.

• Prof. Scott Delancey from University of Oregon (US) presented interesting data on the spread of the Tibeto-Burman family of SEA languages in a paper entitled “Language replacement and the spread of Tibeto-Burman”.

These were followed by the three parallel sections with 63 presentations.

Section 1: Phonetics and Phonology

This section included 21 oral presentations, concentrating on different issues: tones and registers (Justin Watkins, Kanjana Thepboriruk, James P.

Kirby, Honda Koichi, Marc Brunelle); reduplication (Neil Olsen, Nguyễn, Thi Anh-Thu and John Ingram), dialectal differences (Andrew Carson, Andy Castro, Gu Chaowen), Phonological units, variation and perception (Phinnarat Akharawatthanakun, Daryl Chow, Marlys Macken, Tran Thuy Hien, Noellie Bon, Nguyen Van Hue), language reconstruction (David R. Mortensen and James A. Miller); prosody (Kieu Phuong-Ha)

Section 2: Syntax and Lexicology

This section included 20 oral presentations on the following issues: grammaticalization (Alice Vittrant, Danh Thành Do-Hurinville, Kachen Tansiri); discourse (Natchanan Yaowapat, Pasha Siraj); cases and aspects (Nguyen Hoang Trung, Danh Thanh Do-Hurinville, Jake Terrell); particles and classifiers (Amara Prasithrathsint, Nattaya Piriyawiboon, Jennie Tran); sentence and clause construction (James Chancharu and Yanin Sawanakunanon, Jason Lin, Tam Nguyen, Natsuki Matsui) …

Section 3: Applied Linguistics and other fields

This section included 20 oral presentations on diverse topics: language contact and policy (Mark Alves, Kirk Person, Songgot Paanchiangwong, Yukti Mukdawijitra, David Gil); sociolinguistics (Dinh Lu Giang); psycholinguistics (Ly Toan Thang, Parinya Wongtawan), pragmatics (Kim Ngoc Quang, Feng-Fu Tsao, Pham Dinh Trong, Grace Chew Chye Lay, Suthatip Mueanjai); comparative linguistics (Siriwong Hongsawan, Tobias Weber, Su'ad Awab)

Besides, the poster section included 14 poster presentations on various linguistic topics.

This is the first time that SEALS conference was organized in Vietnam and it was not only an opportunity for SEAL researchers to share experiences and research results, but also a chance for friends and colleagues to reunite and for students to meet senior researchers.

On May 29, 2009, in the afternoon, before the last section of presenters, the 19th SEALS business meeting took place in the large conference auditorium. It was well attended by over 50 conference participants. The minutes were taken by Alves Mark, who reported the formation of the SEALS standing committee as follows: “In the interest of making sure that there is specific committee responsible for making future SEALS happen, it was requested that a standing committee be voted on, with the expectation that at least one member would be from hosting institutions. Several SEALS participants, including former committee members and others, volunteered for the position. It was recommended that the committee include one person of the hosting institution. After the names were collected, a general show of hands was given to indicate acceptance of each of the volunteers. The current SEALS committee consists of the following members: Dinh Lu Giang,

Marc Brunelle, Paul Sidwell, Uri Tadmor, Martha Ratliff, Mark Alves, Justin Watkins, and Alice Vittrant. ”

At the end, Dr. Paul Sidwell thanked George Bedell for his service on the committee and announced the first JSEALS publication, noting how it is no longer a proceedings but rather a refereed journal. This involves maintaining an editorial board and an extended pool of reviewers. Dr. Sidwell then summarized the process and timing of the publications, and noted that papers received last year should be published by the end of the year 2009.

At the end of the business meeting, the decision became final to hold SEALS 20 in Europe for the first time in Zurich, Switzerland in 2010.

Paper abstracts from the conference are downloadable at the http://vns.edu.vn/viet, and can also be accessed via the JSEALS site at: http://jseals.org.

Received: 10 May 2009

University of Social Sciences and
Humanities (USSH), Ho Chi Minh City.
VIETNAM
<lugiangdinh@gmail.com>

****