Mon-Khmer Studies

VOLUME 42

The journal of Austroasiatic languages and cultures

Established 1964

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published in 2013 by:

Mahidol University (Thailand)

SIL International (USA)

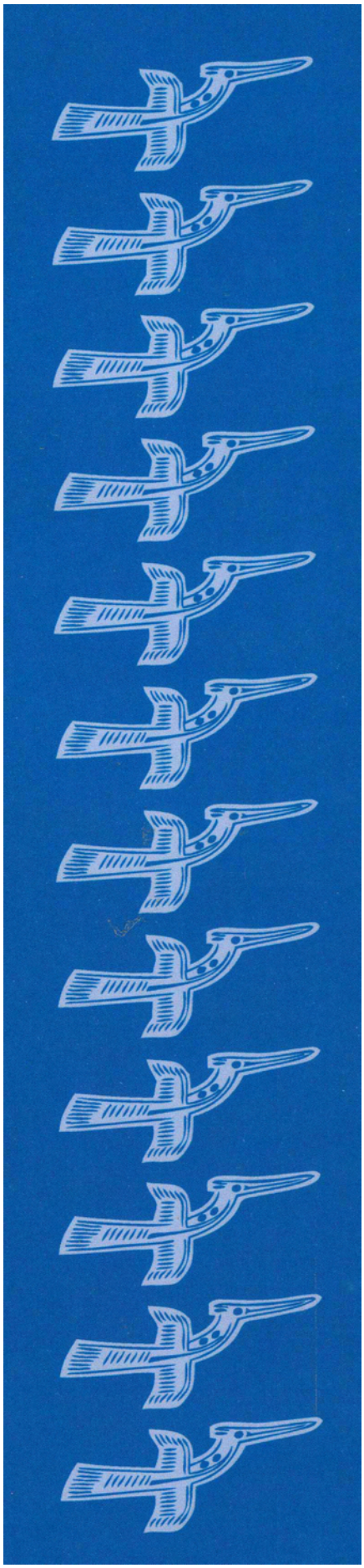
Contents

Papers (Peer reviewed)

Notes, Reviews, Data-Papers

Mon-Khmer Studies

Volume 42



Authors: K. S. NAGARAJA, Paul SIDWELL, Simon GREENHILL
Title: *A Lexicostatistical Study of the Khasian Languages: Khasi, Pnar, Lyngngam, and War*
Pages: 1-11
Date Received: 26/1/2013
Revised Text Accepted: 8/5/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A Lexicostatistical Study of the Khasian Languages: Khasi, Pnar, Lyngngam, and War.[1]

K. S. Nagaraja
Retired, Mysore, India

Paul Sidwell
Australian National University

Simon Greenhill
Australian National University

Abstract
This paper presents the results of lexicostatistical, glottochronological, and Bayesian phylogenetic analyses of a 200 word data set for Standard Khasi, Lyngngam, Pnar and War. Very few works have appeared on the subject of the internal classification of the Khasian branch of Austroasiatic, leaving the existing reference literature disappointingly incomplete. The present analysis supports both the strong identity of Khasian as a unitary branch, with an internally nested branching structure that fits neatly with known historical, geographical and linguistic facts. Additionally, lexically based dating methods suggest that the internal diversification of Khasian began roughly between 1500 and 2000 years ago.
Keywords: Lexicostatistics, Bayesian phylogenetics, language classification
ISO 639-3 language codes: kha, lyg, pvb, aml

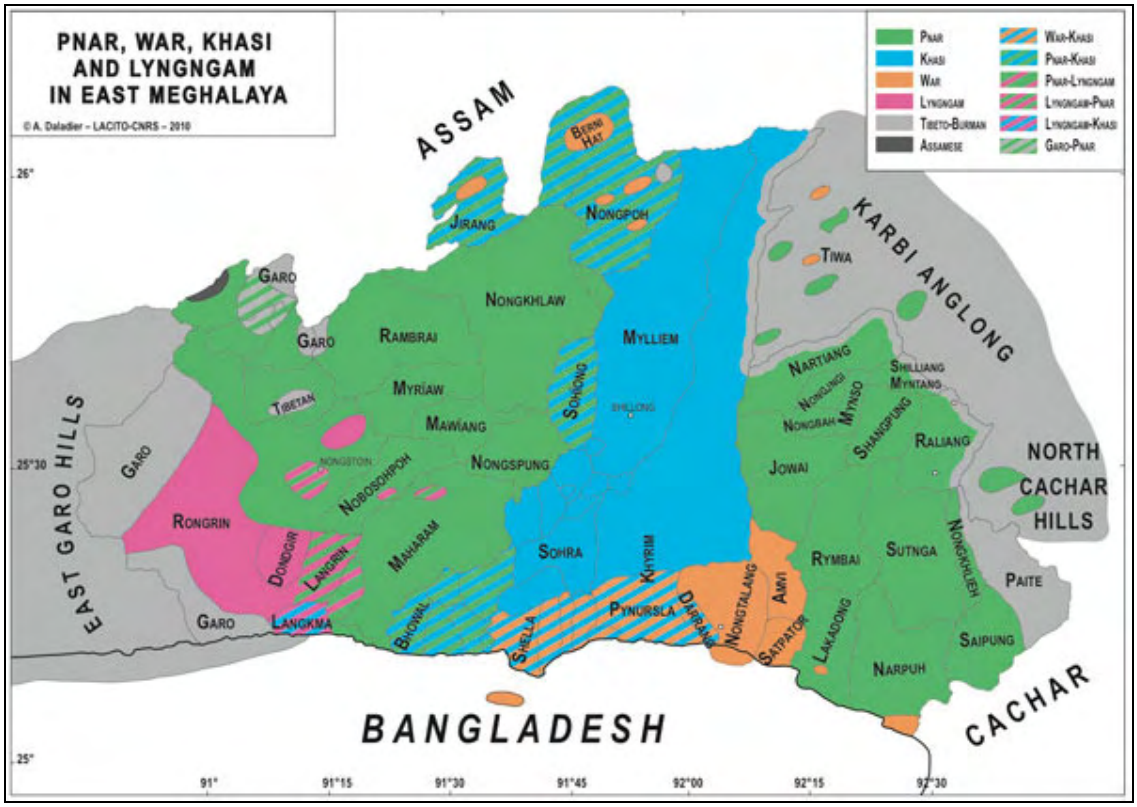


Figure 1: Map of Khasian varieties from Daladier (2010)

[1] The present paper extends a 2004 study by K. S. Nagaraja "A Lexico-statistic study of Khynrian and Lyngngam dialects of the Khasi language" (The NEHU Journal 2.1:43-56). Special thanks are due to Hiram Ring and Mankular Gashnga for assistance with data and analyses in the preparation of this paper.

K. S. NAGARAJA, P. SIDWELL, S. GREENHILL. 2013.
A Lexicostatistical Study of the Khasian Language. *Mon-Khmer Studies*. 42:1-11

1. Background: previous studies

Comparative-historical analysis of Khasian remained underdeveloped through the 20th century, primarily because attention has traditionally focused on the standardized variety, which enjoys official status and widespread use in religious contexts in Meghalaya state. We see an indication that language attitudes were well entrenched already in the 1800s in this extract from Roberts (1891) Khasi grammar:

> In this work, the dialect of Cherrapoonjee is taken as the standard, because it is the purest, as universally acknowledged by the natives, besides being more amenable to systematical arrangement than the patois of the smaller villages.
> (Roberts 1891, xiv)

Robert's text also includes an extensive list of "ugly" (presumably quite popular) non-standard pronunciations that speakers were urged to avoid at all costs. In striking contrast, *The Linguistic Survey of India* (Grierson (1903) correctly recognized four languages which correspond to the four speech varieties analysed in this paper (Khasi (Khynrium), Pnar (Synteng), War, and Lyngngam[2]) and provided some useful comparative lexical and syntactic examples. However, Grierson's data suffered from limitations in the transcription and other gaps that made it difficult to provide a basis for linguistic analyses, and it would be approximately a century before improved data, similarly systematically organized, would start to become available.

Of the present authors, Nagaraja collected data for both Standard Khasi and Lyngngam in 1988, and published a paper on the status of Lyngngam in 1996. That paper made various observations on the grammatical, lexical and phonological correspondences between Khasi and Lyngngam, including the important observation that, "around forty percent of Lyngngam's vocabulary seems to be unrelated to Khasi." The same author followed up with a lexicostatistical study in 2004, based on a 200 word list that was subsequently used as the basis for our more recent and extensive analysis that is the main topic of this paper. That study found 43.9% of cognates between Khasi and Lyngngam, and using Lee's (1953) glottochronological method calculated a separation date of 1,890 years. Subsequently, other lexicostatistical studies have been conducted independently.

Brightbill et al. (2007) conducted a sociolinguistic survey of Khasian villages in Bangladesh, and at a couple of locations within Meghalaya, presenting their wordlists and lexical analyses in their online report (see URL in the references). Although focusing on the War varieties within Bangladesh, Brightbill et al. provided useful lexical data for a number of Khasian varieties, in the form of a comparative lexicon with more than 300 items. On the basis of that list they calculated the lexicostatistical matrix reproduced as Fig. 2 ("lexical similarity chart" in their terminology):

Magurchora (I)										
92	Barenga (E)									
90	91	Niralapunji (A)								
90	89	88	Singur (D)							
90	91	88	88	Aliachora (B)						
88	86	88	86	87	Dabolchora (C)					
89	86	85	87	87	83	Amlarem (J)				
35	35	35	35	35	34	36	Noksia (F)			
18	18	18	18	18	17	17	31	Lyngngam (H)		
33	33	32	34	32	31	34	49	30	Jaintiapur (G)	
25	25	26	26	25	24	28	51	34	48	Shella (K)
29	29	29	30	29	27	32	55	36	53	75 Shillong (L)

Figure 2: Lexicostatistical matrix of selected Khasian varieties from Brightbill et al. (2007:17)

[2] Another apparent substantial Khasian speech community is Maram, to the west of the main Khasi area, corresponding to the green Pnar (!) area on the western side of Daladier's map. Maram is not treated here due to lack of suitable data, but we can report that impressionistically it is very similar to Standard Khasi.

In Brightbill et al.'s scheme the wordlists are identified mostly by place names: the first six above are War varieties spoken in Bangladesh, while *Amlarem* is a War dialect from Meghalaya, the *Noksia* and *Jaintiapur* are Pnar varieties, the *Shella* is ambiguously explained as being "Khasi-War", and the *Shillong* is from a speaker of Standard Khasi. The main result is that War lects in Bangladesh are clearly identified as varieties of one language with percentages all above 80%. However, the other figures are more difficult to interpret, especially in respect of the particularly low percentages that Lyngngam shares with other lists - as low as 17% - well below what we might anticipate given the analysis of Nagaraja (1996). Their calculations appear to be heavily skewed by a failure to allow for missing items in the lists compared, and are included here mainly for the sake of completeness in reviewing the lexicostatistical data on Khasian.

Another of the present authors, Sidwell, attempted his own lexicostatistical study of Khasian, which is presented in his (2009) survey of Austroasiatic classification. That study used the standard 100 word Swadesh list, aggregating items from the following sources:

- Lyngngam data from Nagaraja (1996),
- Khasi from standard dictionaries,
- Amwi from Weidert (1975),
- Pnar (Noksia) and War (Amlarem) from Brighthill et al. (2007).

Cognates were identified manually and a matrix generated (Fig. 3) automatically using Jacques Guy's Glotpc.exe program:[3]

Lyngngam				
63	Khasi (Shillong)			
54	75	Pnar (Noksia)		
41	55	57	War (Amlarem)	
37	53	51	80	Amwi (Weidert)

Figure 3: Lexicostatistical matrix for five Khasian varieties, by Sidwell (2009)

The above figures were interpreted as indicating that the languages fell into two sub-groups: War versus a Khasi-Pnar-Lyngngam group, with the latter having an ambiguous structure. Generally the main finding that the War varieties form a distinct sub-branch is supported strongly by comparative phonology. War is strongly marked by historical vowel restructuring that saw many mergers with high front vowels, and dissimilatory restructuring of diphthongs. Some examples can be seen in the following table (Fig. 4) of data extracted from Lyngngam from Nagaraja (2004), Brightbill et al. (2007), and Amwi from Weidert (1975).

Gloss	**Lyngngam**	**Khasi (Shillong)**	**Pnar (Noksia)**	**War (Amlarem)**	**Amwi**
'two'	*aːr*	*ʔaːr*	*ʔaːr*	*ʔi*	*ʔũ*
'chicken'	--	*ʔiar*	*ʔiar*	*siʔi*	*sʔi*
'fish'	*kʰa*	*kʰa*	*kʰa*	*hi*	*hi*
'red'	*ənsaw*	*saw*	*sao*	*sia*	*sia*
'stone'	*maw*	*maw*	*mao*	*ʃmia*	*ʃmia*

Figure 4: Comparative data illustrating phonological innovations in War

It is apparent that the lexicostatistical studies conducted so far have been very limited in scope, and conducted with differing data sets that make their result difficult to compare and assess. In this context it was decided to extend Nagaraja's (2004) study, by adding data representing Pnar

[3] Figures on branches are words retained per 1000.

and War to the 200 word list already used for Lyngngam and standard Khasi, and additionally to add Palaung data - Palaung representing a more distantly related Austroasiatic language - to securely root the tree and test overall coherence of Khasian.

2. The present study

The present study takes the data set of Nagaraja (2004) to which are added data items for:

- Pnar, Jaintia dialect from Ring (2012).
- War, Lamin dialect from Gashnga (forthcoming).
- Palaung, Namshan dialect from Shorto 2013.

All the data are provided in a table as an appendix to this paper. Cognates are scored in the rightmost column of the table using letter codes, according to the method specified by Guy (1994) in which members of the same etymon are given the same letter, loans and empty fields are given *. Nagaraja's (2004) cognate assignments were reassessed in the light of the new data, and Sidwell's ongoing proto-Khasian phonological reconstruction,[4] resulting in some changes. The scores were then processed with Guy's GLOTPC.EXE to count the pair-wise percentages, generating the table at (Fig. 5):

	Khasi	Pnar	Lyngngam	War	Palaung
Khasi		74	62	52	20
Pnar	74		55	54	19
Lyngngam	62	55		41	18
War	52	54	41		20
Palaung	20	19	18	20	

Figure 5: Lexicostatistical table for Khasi, Pnar, Lyngngam, War and Palaung.

Overall the matrix indicates straightforward nesting branching relations within Khasian, and unambiguous rooting based on the strikingly consistent 18~20% cognacy with Palaung. Further analysis with Guy's GLOTED.EXE indicates that the real percentages diverge from theoretically predicted percentages by no more than 2% in respect of any pair-wise comparisons, so we can have some confidence that the analysis is not significantly distorted by drastic differences in rates of lexical change. As regards to inferring interference by borrowing, it appears that the cognacy rates with War are indicative; we know that War and Lyngngam speakers are geographically separated, and can assume that the figure of 41% counted between them is not significantly distorted by loans. On the other hand, Khasi and Pnar show higher percentatages against War (52% and 54% respectively) and the somewhat higher agreement between Pnar and War, which are known to be in contact, is surely indicative of some mutual borrowing, which has not been identified and scored so in our dataset. Thus, although the pair-wise comparions of Pnar-Lyngngam and Pnar-War show similar values (55% and 54% respectively) we can assume that the latter figure is likely to be high because of undiagnosed borrowing (as borrowing between Pnar and War is far more likely than between either and Lyngngam[5]). Similarly, the higher agreement between Khasi-Lyngngam (62%) versus Pnar-Lyngngam (55%) is likely to be indicative of some borrowing of Standard Khasi words into Lyngngam. Of course, it must be acknowledged that it is possible that the above patterns are largely the result of differences in rates of change, but logically it is difficult to see how that would produce such a tidily branching nested tree, as opposed to a more random pattern.

The figure of 74% agreement between Khasi and Pnar is strikingly consistent with 75% figures independently obtained by Brightbill et al. (2007) and Sidwell (2009), and provide significant comfort to the view that they are more or less indicative of the real distance between the two languages. The figures indicate that a high degree of mutual lexical intelligibility is to be predicted, approaching the threshold for treating them as dialects of the same language. Clearly

[4] At the time of writing a 2012 version of this reconstruction is available online at sealang.net/monkhmer. It is expected that this will be replaced with an extensively revised version later in 2013.

[5] If anything, all three are likely to share unrecognised loans from Standard Khasi.

Khasi and Pnar sub-group closely, Lyngngam then appears to sit above Khasi-Pnar, and all three are more distantly related to War, which (as noted above) also is known to have a divergent phonological history.

For those who are bold enough to pursue the question, it is also possible to apply glottochronological calculation to our figures, in full awareness of the harsh critiques of glottochronology, especially since Bergsland and Vogt. (1962). We have done so, using Lee's (1953) formula **t = log C / (2 log r)** and his retention rate for the 200 item list of 80.5% per thousand years. Applying this formula to the lowest pair-wise percentage at each apparent node, we get the following tree with divergences dated in years (y) before present at Fig. 6.

```
Khasi ------:74%/694y----:55%/1378y----:41%/2054y----:18%/3951y
Pnar -------'             |             |             |
Lyngngam ----------------'              |             |
War ------------------------------------'             |
Palaung ----------------------------------------------'
```

Figure 6: Family tree with glottochronological dating of divergences table for Khasi, Pnar, Lyngngam, War and Palaung.

Further computational analyses were carried out on the dataset. Firstly a neighbor net was generated using SplitsTree v4.11.3 (Bryant & Moulton 2003) by Simon Greenhill, here at Fig. 7. The result, displayed below, is quite straightforward, and is largely consistent with the lexicostatistics: the close relation between Khasi and Pnar is reproduced, and the marginally closer relation of them to Lyngngam versus War is evident. No disproportionate inferring signals are evident.

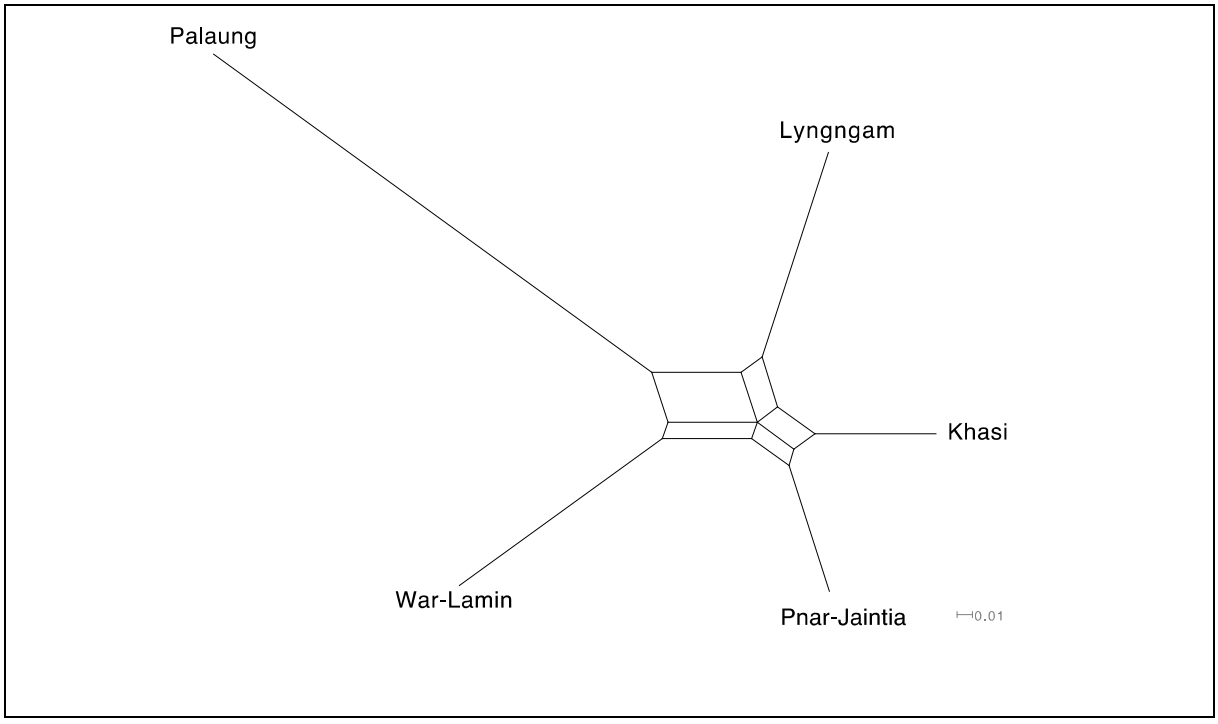


Figure 7: Neighbor Net for Khasi, Pnar, Lyngngam, War and Palaung.

Next, a Bayesian Phylogenetic analysis was run by Greenhill using BEAST v1.7.4 (Drummond et al. 2012). Here, a simple Continuous-Time Markov Chain model was used to analyse the binary presence/absence of cognates implementing a strict clock for inferring rates of cognate gains and losses. The analysis was run for 2,000,000 generations, sampling 1,000. The first 200,000 generations were discarded as "burn-in" after inspection of the traces showed that this was sufficient time for the chain to stabilize (c.f. Greenhill, in press). The results are similarly consistent with the lexicostatistics; the number 1 at each node indicates 100% probability of the

branching, as the program consistently generated the same tree with every pass through the data. Additionally the tree is constrained to indicate a time depth of 500 years BP for the Khasi-Pnar split, for the sake of generating a calibrated tree. In so far as we are able to offer any objective bases for calibrating the chronology, the Buranji chronicles of the Ahom kingdom apparently reference the Pnar kingdom at Jaintia about 500 years before present, which suggests a floor under the separation of Khasi and Pnar (e.g. Gait 1906:255 lists Jaintia kings from approximately 1500 AD onwards). In this context, the glottochronological calculation of 694 years for Khasi-Pnar separation seems quite realistic, although still admittedly speculative. The Bayesian analysis estimates the age of the Khasi-Pnar split to be similar – but younger – with a mean of 535 years (95% Highest Posterior Density Interval = 500-603 years), see Fig. 8. In terms of the age of the Khasian subgroup itself, glottochronology estimates the age at 2054 years, while the Bayesian analysis places the origin of this subgroup at a younger median of 1350 years (95% HPD = 1028-1737 years). Given that the Bayesian dating estimate is calibrated to the youngest possible age indicated by our meager historical sources, the estimate of 1350 years is quite likely to be an underestimate, although as such it establishes a reasonable minimum parameter for speculations about pre-Khasian migration into Northeast India.

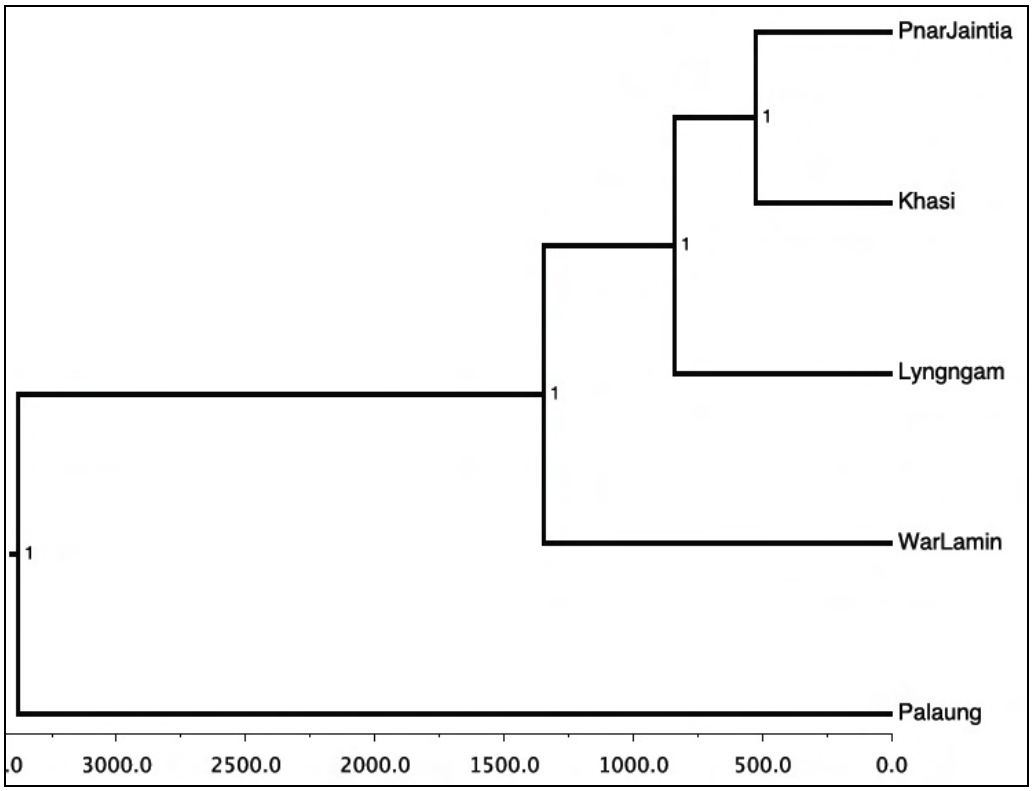


Figure 8: Bayesian Phylogenetic analysis for Khasi, Pnar, Lyngngam, War and Palaung.

3. Concluding remarks

The present study makes a further contribution to the emerging field of comparative Khasian linguistics, with a quantitative analysis of lexical correspondences that supports both the unity of the Khasian branch, and a strong nested internal structure. Within Khasian, the War language(s) form the highest branching node, consistent with indications of historical phonological restructuring. The remaining languages form a tightly linked subgroup, with Lyngngam placed outside a Khasi-Pnar core. Whilst these results are intriguing, fine-grained lexical, grammatical and phonological analyses should be applied to further rigorously infer the subgrouping of the Khasian languages.

References

Bergsland, Knut, Hans Vogt. 1962. On the Validity of Glottochronology. *Current Anthropology* 3.2:115-153.

Brightbill, Jeremy, Amy Kim, and Seung Kim. 2007. *The War-Jaintia in Bangladesh: a sociolinguistic survey*. SIL Electronic Survey Reports.

Bryant, David & Moulton, Vincent. 2003. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. *Molecular Biology and Evolution*, 21(2), 255–265.

Daladier, Anne 2010. (web resource) Carte du groupe pnarique-war-lyngngam http://lacito.vjf.cnrs.fr/themes/tibeto-birman/images/East-Meghalaya-PWKL600.jpg

Drummond, Alexei J., Suchard, Mark A., Dong, Xie, & Rambaut, Andrew. 2012. Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution*, 29(8), 1969-1973.

Gait, E. A. 1906. *A History of Assam*. Calcutta: Thacker, Spink & Co.

Gashnga, Mankular. forthcoming. *Lamin(War)-English Dictionary*. Canberra, Asia-Pacific Linguistics.

Greenhill, Simon J. In Press. Evolution and Language: Phylogenetic Analyses. In *The International Encyclopedia of the Social and Behavioral Sciences*. Wright, JD (Ed). Elsevier.

Grierson, G.A., 1903, *Linguistic Survey of India*. Vol. (iii) Part (ii) (Reprint 1967). Delhi. Banarashidaas.

Guy, Jacques. 1994. GLOTTO (web resource) http://linguistlist.org/issues/5/5-630.html

Lees, Robert B. 1953. On the basis of glottochronology. *Language* 29: 113-127.

Nagaraja, K.S. 1996. Status of Lyngngam, *Mon-Khmer Studies*, Vol. 26: 37-50.

Nagaraja, K.S. 2004. A Lexicostatistic study of Khynriam (St.Khasi) and Lyngngam dialects of the Khasi language. *The NEHU Journal* II.1:43-56.

Ring, Hiram. 2012. A phonetic description and phonemic alanlysis of Pnar. *Mon-Khmer Studies* 40:133-175.

Roberts, H. 1891. *A Grammar of Khassi language*, London.

Sidwell, Paul. 2009. *Classifying the Austroasiatic languages: history and state of the art*. Munich, Lincom Europa.

Shorto, H. L. 2013. *Palaung Word List: Based on materials collected from Pan Shwe Kya, Namhsan, Sept-Oct, 1957*. Canberra, Asia-Pacific Linguistics.

Weidert, Alfons K. 1975. *I tkong Amwi. Deskriptive Analyse eines Wardialekts des Khasi*. Wiesbaden: Harrassowitz

Appendix: Lexical data and cognate scores

Lexical data and cognate scores (* marks loanwords and missing forms)

	Gloss	Khasi	Lyngngam	Pnar/Jaintia	War/Lamin	Palaung	Cognate scores
1	all	*roʔ*	*prok*	*warɔʔ*	*bərɒʔ*	*paj*	aaaab
2	and	*baːd*	*nam*	*wa*	*waː*	-	abcd*
3	animal	*mraːd*	*mraːd*	*mraːd*	*mrat*	*to*	*****
4	ashes	*dpej*	*əpaw*	*tpai*	*tvɔ*	*kəhvaŋ*	abacd
5	at	*ha*	*he*	*ha*	*ti*	--	aaab*
6	back (anat.)	*dien*	*bad don*	*rɲkʰi*	*təmpɒŋ*	*krɔŋ*	aabcd
7	bad	*sɲiew*	*kɨncʰa*	*siʔ*	*kɒ̃m*	*kʰu (?)*	abcde
8	bark (of tree)	*snep*	*snieʔ*	*sneiʔ*	*snɪəʔ*	*gɔʔ*	aaaab
9	because	*namar*	*amte*	*neibʰaʔ*	*kaʔ*	*(?)*	abcd*
10	belly	*kpoʔ*	*ləwbaʔ*	*kpɔʔ*	*pɒʔ*	*vɛʔ*	abaac
11	big	*heʔ*	*kɨmba*	*hɛʔ*	*mɪa*	*daŋ*	ababc
12	bird	*sim*	*sim*	*sim*	*ksem*	*sim*	aaaaa
13	to bite	*daʲt*	*kɨnnap*	*dait*	*hit*	*gaʔ*	abacd
14	black	*joŋ*	*ɨŋŋoŋ*	*jɔŋ*	*priŋ*	*jəm*	aaabc
15	blood	*snam*	*snam*	*snam*	*rnʊə*	*hnam*	aaaaa
16	blow	*pɨrsat*	*pʰinnur*	*slu*	*pet*	*put*	abcdd

K. S. NAGARAJA, P. SIDWELL, S. GREENHILL. 2013.
A Lexicostatistical Study of the Khasian Languages. *Mon-Khmer Studies*. 42:1-11

17	bone	*ʃʔeŋ*	*cʔeŋ*	*ʧʔeiɲ*	*ʧiəʔ*	*kəʔaŋ*	aaaaa
18	breathe	*riŋ mɨnsiem*	*riŋ ɨnsom*	*rɪŋ mɲseim*	*reŋ hənsʊ*	*pʰum*	aaaaa
19	burn	*tʰaŋ*	*ɨntʰaŋ / tʰɨnnəŋ*	*tʰaŋ*	*ɵaŋ*	*gut*	aaaab
20	child	*kʰun*	*kʰon*	*kʰɔn*	*hʊn*	*kuən*	aaaaa
21	cloud	*lʔoʔ*	*lʔoʔ*	*lʔɔʔ*	*ləmpem*	*ut*	aaabc
22	cold	*kʰriat*	*bɨnsir*	*kdʒam*	*ktjam*	*kat*	abccd
23	come	*wan*	*lɨnnar*	*wan*	*va, van*	*hluh, rɔt*	abaac
24	count	*ɲjaw*	*cʰɨnnan*	*niaw*	*ʧaː*	*dir*	abacd
25	cut	*kʰap*	*kʰɨnnap*	*aʔ*	*pam*	*set*	aabcd
26	day	*sŋi*	*sŋej*	*sŋi*	*ʃŋaː*	*səŋi*	aaaaa
27	die	*jap*	*ɲinnap*	*jap*	*jip*	*jəm*	aaaab
28	dig	*tiʔ*	*tɨnniet*	*tiʔ*	*tɪəʔ*	*puər*	aaaab
29	dirty	*ɟaboʔ / ɟakʰlia*	*dɨmmut / ɟɨmbaʲt*	*taroj*	*ʧəmet*	*ɲu ɲu (?)*	abcde
30	dog	*ksew*	*ksu*	*ksaw*	*ksɪa*	*sɔ*	aaaaa
31	drink	*diʔ*	*dɨnniet*	*diʔ*	*deʔ*	*teəŋ*	aaaab
32	dry	*rkʰiaŋ*	*riəŋkʰoŋ*	*raw*	*rhɪəŋ*	*raj, roh*	aabac
33	dull, blunt	*isiʔ*	*ɟosmoʲt*	*tʰla*	*len*	--	abcd*
34	dust	*pum-pum*	*pum-pum*	*dʒlpʰoʔ*	*tjəlpʰʊʔ*	*kərboh*	a*bbc
35	ear	*ʃkor*	*ləkur*	*ʧkɔr*	*tərəŋ*	*hjɔʔ*	aaabc
36	earth	*kʰɨndew*	*kmiəŋ*	*kʰndaw*	*pərθa*	*kətɛ*	abc*d
37	eat	*baːm*	*bɨnnəŋ*	*bam*	*bʊə*	*hap*	aaaac
38	egg	*pɨlleŋ*	*pɨlliŋ*	*pl̩leiɲ*	*sʔi*	*kətəm*	aaabc
39	eye	*kʰmat*	*kʰmat*	*kʰmat*	*mat*	*ŋaj*	aaaab
40	to fall	*hap*	*eŋaj*	*hap*	*hərem*	*rar*	abacd
41	far	*ɟŋaj*	*ɟŋi*	*dʒŋai*	*ʧŋʊ*	*səŋaj*	aaaaa
42	fat-grease	*sŋaːʲd*	*ɨmmir*	*kʰlaŋ*	*ləʔɒt*	*kəmu*	abcde
43	father	*kpa*	*pa*	*pa*	*pa*	*kun*	aaaab
44	to fear	*ʃeptieŋ*	*tieŋ dait*	*teʲɲ*	*ktɪəŋ*	*jɔ*	aaaab
45	feather	*sner*	*snir*	*tʰawanɛr*	*θəbənɪar*	--	aaaa*
46	few	*kʰɨn-diat*	*tah-diat*	*kʰadʒʲak*	--	*bre*	aab*c
47	to fight	*jaʃoʔ / jadat*	*jamuʔ*	*jatʃɔʔ*	*jaʔ dat*	*taik*	aba*c
48	fire	*diŋ*	*ədoŋ*	*diɲ*	*ʃmen*	*ŋər*	aaabc
49	fish	*doʔkʰa*	*kʰa*	*dakʰa*	*hi*	*ka*	aaaaa
50	five	*san*	*san*	*san*	*ran*	*pʰən*	aaaaa
51	to float	*per*	*raŋ*	*pɛr*	*sper*	*plur*	abaac
52	to flow	*tuːʲd*	*sɨn-toːʲd*	*toːʲd*	*pɒr*	*hlaj*	aaabc
53	flower	*sɨntiew*	*sɨntew*	*sn̩tu*	*khlʊə*	*poh*	aaabc
54	to fly	*her*	*kɨndej*	*pn̩hɛr*	*pɪar*	*pər*	abacc
55	fog	*dum-lʔoʔ*	*nioŋ nia*	*lʔɔʔ kʰn̩daw*	*dʊm*	*aj*	abcad
56	foot	*kɟat*	*kɟat*	*kdʒat*	*nɪa*	*ɟɯŋ*	aaabc
57	four	*saw*	*saw*	*so*	*rɪa*	*pʰon*	aaaab
58	freeze	*ʃoʔ tʰaʔ*	*tiŋŋam / bɨnsier*	*tʰaʔ*	*tʰaʔ*	*krɔʔ*	abaac
59	fruit	*soʔ*	*suʔ*	*sɔʔ*	*sɒʔ*	*ple*	aaaab
60	give	*aj*	*ɨnnaj*	*e*	*ʔaː*	*dɛh*	aaaab
61	good	*bʰa*	*mɨrrʰiaŋ*	*bʰa*	*mɪət*	*laʔ*	*a*bc
62	grass	*pʰlaŋ*	*pʰlaŋ*	*pʰlaŋ*	*smʊt*	*kərban*	aaabc
63	green	*ɟɨrŋam*	*sɨŋiəŋ*	*jɾŋam*	*ʧərŋam*	*ɲər*	aaaab
64	guts, intestines	*snier*	*snor*	*snɛr*	*nɒr*	*rɛŋ*	aaaab
65	hair	*sɲiuʔ*	*sɲiək*	*sɲʲoʔ*	*sʊʔ*	*huʔ*	aaaaa
66	hand	*kti*	*ktej*	*kti*	*taː*	*ti*	aaaaa
67	he	*u*	*ɟutuʔ / umi*	*u*	*u*	*ən*	aaaab
68	head	*kʰlieʔ*	*kʰliʔ*	*kʰleiʔ*	*kʰlɪa*	*kiŋ*	aaaab
69	hear	*sŋap*	*sŋu*	*sniaw*	*sãʔ*	*ɟu*	aaabc
70	heart	*kloŋ snam*	*kloŋ snam*	*klɔŋ snam*	*klɒŋ rnʊə*	*nuər*	aaaab
71	heavy	*heʔ*	*kenbaʔ / kʰɨnnia*	*kʰia*	*stʊʔ*	*ɟən*	abcde
72	here	*haŋ ne*	*haniʔ*	*heini*	*tineː*	--	aaaa*
73	hit	*tied*	*udaʔ*	*dat*	*dat*	*tum*	abccd
74	hold-take	*ʃɨm*	*tʰom*	*tʃim*	*lʊm*	*lɛ*	abacd

75	how	*kumno*	*naŋ net*	*kammɔn*	*kinjaʔ*	*kʰuj mɔ*	abacd
76	hunt	*beʔ (mraːd)*	*wuŋ na*	*lai siɛt dɔʔ*	*pətar*	*ɻɔm*	abcde
77	husband	*tŋa / lok*	*koraŋ*	*lɔk*	*lʊk*	*rəleh*	abbac
78	I	*ŋa*	*nə*	*ŋa*	*ŋə*	*ʔɔ*	aaaab
79	ice	*tʰaʔ*	*tʰaʔ-əlliʔ*	*tʰaʔ*	*tʰaʔ*	*je ge*	aaaa*
80	if	*lada*	*lede*	*lada*	*nimə*	--	aaab*
81	in	*ha*	*he*	*ha*	*ti*	*naə*	aaab*
82	kill	*pɨnjap*	*pɨnŋap*	*pɲjap*	*pənjip*	*piəm*	aaaab
83	know	*tip*	*heʔkən*	*tɪp*	*tɒʔ*	*nəp*	abaac
84	lake	*puŋ*	*puŋ*	*puŋ*	*sʊ*	*nɔŋ*	aaab*
85	laugh	*rkʰie*	*ɨllom*	*rkʰai*	*rʊ*	*jum*	abac*
86	leaf	*sla*	*sla*	*sla*	*sli*	*hla*	aaaaa
87	left side	*diaŋ*	*tɨmmiəŋ*	*tidiɛŋ*	*di par ta dɪəŋ*	*ʔi-ve*	aaaab
88	leg	*kɻat*	*kɻat*	*kdʒat*	*nɪa*	*ɻuŋ*	aaabc
89	to lie, deceive	*tʰok*	*tʰɨlloʲt*	*dʒlɛr*	*pənrɒʔ*	*cɔʔ*	abcde
90	live	*im*	*ɨnnim*	*ɪm*	*pʔem*	*im*	aaaaa
91	liver	*doʔnuːd*	*noːd*	*noːd*	*kθim*	*kərtɔm*	aaabb
92	long	*ɻroŋ*	*ɻirroŋ*	*dʒrɔŋ*	*kərɒŋ*	*hluŋ*	aaaab
93	louse	*ksi / ɻɨnreiŋ*	*sɨlliet*	*ksi*	*ksa*	*si*	abaaa
94	man-male	*ʃinraŋ*	*kʰonkoraŋ*	*cɳraŋ*	*tərma*	*ime*	aaabb
95	many	*bun*	*bon*	*bon*	*ʃibʊə*	*kun, bərcu*	aaaab
96	meat-flesh	*doʔ*	*meʲm*	*dɔʔ*	*dɒʔ*	*jəŋ*	abaac
97	mother	*kmie*	*gma*	*bei*	*maː*	*ma*	aaaaa
98	mountain	*lum*	*dom*	*lom*	*pdeŋ*	*sor*	aaabc
99	mouth	*ʃintur*	*gap*	*ktein*	*tkɒŋ*	*mur*	abcde
100	name	*kɨrteŋ*	*kɨrteŋ*	*prtuid*	*tvɪəŋ*	*ɻu*	a*bc*
101	narrow	*raʃiŋ / bakʰim*	*bakʰim*	*kʰɪm*	--	*ɔp*	a*a**
102	near	*ɻan*	*ɻŋan*	*dʒan*	*tjan*	*dət*	aaaab
103	neck	*rɨndaŋ*	*kraŋ*	*rdaŋ*	*rdaŋ*	*rəmɔŋ*	abaac
104	new	*tʰɨmmaj*	*tʰɨmmaj*	*tʰɱme*	*θmaː*	*kənmɛ*	aaaaa
105	night	*miet*	*sənŋu*	*meit*	*ləmaʔ*	*sum*	abaac
106	nose	*kʰmut*	*leumut*	*kʰmut*	*mərkɤ̃ŋ*	*muh, mur*	aaaba
107	not	*ɨm*	*ɨŋi*	*ɱ*	*tɒʔ tə*	*kə*	aabbc
108	old	*rim*	*rim*	*rɪm*	*sərem*	*prim*	aaaaa
109	one	*uwej*	*uwew*	*wi*	*mi*	*u*	aaaab
110	other	*ki-wej*	*marber*	*kɔpsar*	--	*laj*	aa**b
111	person	*briew*	*brü*	*bru*	*tjəpreʊ*	*bi*	aaaab
112	to play	*leʔ kaj*	*kʰellaj*	*kɳdeiʔ ke*	*khirʊ*	*kəvɛʔ*	abacd
113	to pull	*tan*	*rɨnnieŋ*	*tan*	*pətɪaʔ*	*ruit, tut*	abacd
114	to push	*kʰɨnniaʔ*	*kɨncʰew*	*ŋiat*	*khən jit*	*con*	abacd
115	to rain	*slap*	*slap*	*slap*	*slaː*	*ɻuŋ*	aaaab
116	red	*saw*	*ənsaw*	*so*	*sɪa*	*kʰo, ni*	aaaab
117	right-correct	*dej*	*dew*	*tɔʔ*	*tɒʔ*	--	aabb*
118	right side	*mon*	*tɨm-mon*	*timun*	*di par ta mɒn*	*kʰwa*	aaaa*
119	river	*waʔ*	*por*	*waʔ*	*ʔam*	*om*	abacc
120	road	*linti*	*twar*	*sərɔk*	*sərɒk*	*deŋ*	abacd
121	root	*tɨnraj*	*tɨrraj*	*tʰeid*	*ʃɨt*	*riər*	aabba
122	rope	*tɨllaj*	*laŋnaj*	*tɭle*	*tərʊ*	*vər*	abacd
123	rotten	*pʔut*	*ɻillit*	*pʔut*	*khvi*	*əm*	abacd
124	to rub	*kɨrʃut*	*kɨrcʰut*	*kɽtʃut*	*kəntʃʊt*	*sut*	aaaaa
125	salt	*mluʔ*	*maluk*	*bloʔ*	*pnʊʔ*	*sɔʔ*	aaaab
126	sand	*ʃʔiap*	*cʰʔjap*	*tʃʔɛʌp*	*sərvʊ*	*saj*	aaabc
127	to say	*oŋ*	*ɨnnoŋ*	*ʔɔŋ*	*ʔɒŋ*	*dah*	aaaab
128	to scratch	*truːd*	*tɨrrut*	*tʃbɔt*	*kʰəbɒʔ*	*pɔʔ*	aabbc
129	sea	*duriaw*	*duriaw*	*duriaw*	*dʊrɪəʊ*	--	a*aa*
130	to see	*joʔi*	*muɻoʔ*	*pait*	*maʔ*	*ju*	aabca
131	seed	*sɨmbaj*	*ɻellej*	*sɱbe*	*tjʊsba*	*kəʔaŋ*	abaac
132	to sew	*suʔ*	*sɨnnek*	*sor*	*sʊ*	*ɻiŋ*	aaaab

K. S. NAGARAJA, P. SIDWELL, S. GREENHILL. 2013.
A Lexicostatistical Study of the Khasian Languages. *Mon-Khmer Studies*. 42:1-11

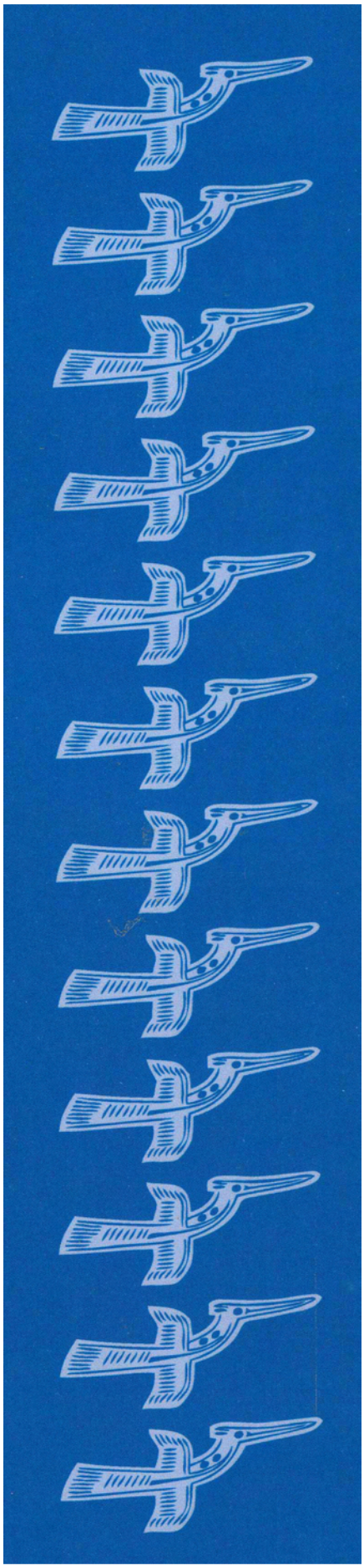
133	sharp	*nɐp*	*ɨntaʔ*	*nɛp*	*nɐp*	*ləm*	abaac
134	short	*liŋkot*	*tɨmban*	*tbien*	*ʧrit*	*ɛm*	abbcd
135	to sing	*rwaj*	*rɨŋwi*	*rwai*	*rvʊ*	*ŋir*	aaaab
136	to sit	*ʃoŋ*	*cʰoŋ*	*tʃɔŋ*	*ʃkɪə*	*məʔ*	aaabc
137	skin	*snieʔ doʔ*	*snieʔ mejm*	*sneiʔ*	*snɪəʔ*	*hur*	aaaab
138	sky	*bneŋ*	*brej*	*bneiŋ*	--	*pleŋ*	aba*a
139	to sleep	*tʰiaʔ*	*ɨŋŋiŋ*	*tʰiaʔ*	*θɪəʔ*	*ʔit*	abaac
140	small	*rit*	*doh-dit*	*kʰeiŋ*	*sbɪət*	*diət*	abcdb
141	to smell	*sma*	*ɨnnaw*	*sma*	*rʔɪəŋ*	*ʔur*	abacd
142	smoke	*tdem*	*ɨntʰak*	*tdɛm*	*tdem*	*tuk*	abaac
143	smooth	*ɟliʔ*	*ɟɨmpaj*	*jaliʔ*	*tjəlliʔ*	*kleət*	abaac
144	snake	*bseŋ*	*bseŋ*	*pseiŋ*	*psen*	*hiŋ*	aaaaa
145	snow	*jor*	*jor*	*tʰaʔ*	*ksɪəŋ məŋ*	--	a*bc*
146	some	*kʰɨndiat*	*taʔ-diat*	*kʰaɟiak*	*ʃitʃĩŋ*	*pərdi*	aabcd
147	to spit	*biaʔ*	*ɟɨrtʰew*	*mn̥tʰu*	*pəθeʊ*	*bɛʔ*	abbba
148	to split	*pʰiaʔ*	*tʰɨillaʔ*	*pʰiaʔ*	*phit, khlɪəʔ*	*ploh*	abaac
149	to squeeze	*kʰem*	*kʰɨnnim*	*ksiʔ*	*ʧəpɪə*	*piət*	aabcc
150	to stab-pierce	*duŋ*	*daneŋ*	*duŋ*	*təndoŋ*	*bruh*	aaaab
151	to stand	*ieŋ*	*ñiəŋ*	*jeiŋ*	*rəŋ*	*ɟəŋ*	aaaaa
152	star	*kʰlur*	*kʰlor*	*kʰlor*	*khlʊə ʃmen*	*səmiŋ*	aaabb
153	stick (of wood)	*dieŋ*	*ədiəŋ*	*deiñ*	*pərnɪa*	*hviət*	aabcd
154	stone	*maw*	*maw*	*mo*	*ʃmɪa*	*mo*	aaaaa
155	straight	*biet*	*lɨmpʰar*	*beit*	*bɪt*	*pʰiəŋ*	abaac
156	to suck	*kɟit*	*kɨnɟok*	*bu*	*tjʊr*	*bu, but*	aabcb
157	sun	*sŋi*	*sŋej*	*sŋi*	*njʊŋa*	*səŋi*	aaaaa
158	to swell	*at*	*ɨnnat*	*ad*	*ʔat*	*gɯ*	aaaab
159	to swim	*ɟŋi*	*ɟɨnnaj*	*jmpa*	*rɪəŋ*	*lɔj*	aabac
160	tail	*tdoŋ*	*kdoŋ*	*tdɔŋ*	*tdɒŋ*	*staʔ*	aaaab
161	that	*-ta / -tej*	*ga-tej*	*katai*	*ke/u tʊn*	*taj*	aaaba
162	there	*katʰie*	*gatʰoʔ*	*heitai*	*tʊ tʊn*	--	aabc*
163	they	*ki*	*gniʔ*	*ki*	*jə*	*ge*	abaca
164	thick	*rben*	*rɨmbin*	*rbɛn*	*rben*	*hət*	aaaab
165	thin	*staŋ*	*sɨntaŋ*	*staŋ*	*staŋ*	*hrer*	aaa*b
166	think	*pɨrkʰat*	*pɨrkʰat*	*pr̥kʰat*	*pərkhat*	*tʰaŋ*	a***b
167	this	*-ne*	*ga-niʔ*	*kani*	*ke/u ne*	*ʔɯ*	aaaab
168	thou	*me / pʰa*	*mi / pʰe*	*me / pʰa*	*ɪəm (m) / ɪəhe (f)*	*mi*	aaaba
169	three	*laj*	*laj*	*le*	*la*	*ʔuəj*	aaaab
170	to throw	*kawaŋ*	*lɨntʰew*	*pakʰɔt*	*phədat*	*rup*	abcde
171	to tie	*teʔ*	*tɨnnak*	*kdɔʔ*	*kʊt*	*toʔ*	aabca
172	tongue	*tʰɨillieɟ*	*tʰɨlloʲt*	*tʰl̥leiʔ*	*khlit*	*kərtaʔ*	aaaab
173	tooth	*bniat*	*moʲn*	*lʌmeiŋ*	*ləmen*	*hraŋ*	abbbc
174	tree	*dieŋ*	*diəŋ*	*deiŋ*	*tvɪa*	*he*	aaabc
175	to turn	*kɨlla*	*kɨlla*	*dɔŋ*	*khərvi*	*pən*	a*bcd
176	two	*aːr*	*ar*	*ar*	*ʔõə*	*ʔar*	aaaaa
177	to vomit	*prei*	*pɨrraw*	*prai*	*hərɒʔ*	*hɯr*	aaabc
178	to walk	*jaːʲd*	*dinniʔ*	*lai kdʒat*	*lɪa*	*pʰət*	abcd*
179	warm	*sʔaːʲd*	*ɨnʃɨt*	*cit*	*dɒt*	*kəʔur*	abbcd
180	to wash	*sait*	*sait*	*sait*	*ksi*	*kʰoj, kəta*	a*aab
181	water	*um*	*gum*	*um*	*ʔam*	*ʔom*	aaaaa
182	we	*ŋi*	*jew*	*i*	*ʔɪə*	*ʔɛ*	abccc
183	wet	*ɟʰieʔ*	*ɟimbaʲt*	*dʒheiʔ*	*tjəriəʔ*	*om pjo pjo*	abaac
184	what	*-ej*	*umet*	*ilɛʔ*	*i a*	*mɔ*	abcad
185	when	*lano*	*minnet*	*n̥nu*	*daŋ njaʔ*	*jam*	abac*
186	where	*haej*	*hanet*	*ʧeiwʌn*	*ti njaʔ*	*mɔ*	ab*cd
187	white	*lieʔ*	*əlliʔ*	*leiʔ*	*slaŋ*	*blɔʔ*	aaabc
188	who	*-no*	*jət*	*u/ka ji*	*u/ke ʔaː*	*paj*	abbcd
189	wide	*jar*	*ɨniər*	*jar*	*hɪaŋ*	*vah*	abcde
190	wife	*tŋa*	*kontʰaw*	*lɔk*	*kə lɒk*	*pənle*	abcde

K. S. NAGARAJA, P. SIDWELL, S. GREENHILL. 2013.
A Lexicostatistical Study of the Khasian Languages. *Mon-Khmer Studies*. 42:1-11

191	wind	*lʔer*	*lʔier*	*lʔɛr*	*srʊə*	*kur*	aaaab
192	wing	*tʰapŋiaŋ*	*tʰapnir*	*tʰawanɛr*	*θəbənɪar*	*peəŋ*	aaaab
193	to wipe	*ŋiad*	*innat*	*ɲiam*	*ʃʔɪəʔ*	*kʰut, kʰuit*	abcde
194	with	*bad*	*nam*	*wa*	*bəʔ*	--	abca*
195	women	*kɨntʰej*	*rawkmaw*	*kn̩tʰai*	*hənθa*	*ɪpən*	abaac
196	woods, forest	*kʰlaw*	*ləwtəp*	*kʰlo*	*kərmɪa*	*bri*	abacd
197	worm	*wieʔ*	*wiak*	*weiʔ*	*khvi*	*riər*	aaaab
198	ye > you (pl.)	*pʰi*	*pʰjaw*	*pʰi*	*hi*	*pɛ*	aaaaa
199	year	*snem*	*snim*	*snɛm*	*snem*	*sənəm*	aaaaa
200	yellow	*stem*	*sɨntim*	*stɛm*	*tŋʊə*	*teŋ*	aaabc

Mon-Khmer Studies

Volume 42



Author: Michelle MILLER
Title: A Description of Kmhmu' Lao Script-Based Orthography
Pages: 12-25
Date Received: 26/6/2012
Revised Text Accepted: 14/5/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A Description of
Kmhmu' Lao Script-Based Orthography

Michelle MILLER
SIL International

Abstract
Kmhmu' is a language of the Mon-Khmer language family. Extensive linguistic research and analysis of the varieties of Kmhmu' spoken in Southeast Asia has led to the grouping of Kmhmu' into three dialect categories, generally referred to as Northern, Western and Southern (Svantesson 1989). The orthography described in this paper was developed for the Southern dialect and utilizes a Lao-based script. Suksavang and Preisig (Suksavang et al 1994) were instrumental in refining this orthography. This description of the Southern Kmhmu' orthography explains how the Lao script is used and/or adapted to represent the Kmhmu' phonemes, presents orthographic conventions for writing words of various structural types and summarizes teaching/learning experiences observed in mother-tongue Kmhmu' speakers.
Keywords: Austroasiatic, orthography, sesquisyllables
ISO 639-3 language codes: kjg

1. Introduction

Kmhmu' is a language of the Mon-Khmer language family. There has been extensive linguistic research and analysis of the varieties of Kmhmu' spoken in Southeast Asia, which has led to grouping of Kmhmu' dialects into three major dialect categories, generally referred to as Northern, Western and Southern (Svantesson 1989). The orthography described in this paper was developed for the Southern dialect, which is spoken in Phongsali, eastern Udomsay, Luang Prabang, Hua Phanh, Xieng Khouang, Sayabuli, Vientiane and Bolikhamsay provinces in Lao PDR (Suksavang et al. 1994, Svantesson 1989), in Diên Biên Phu', So'n La and Nghê An provinces in North Vietnam and some villages of Sipsongpanna in China (Suwilai 2002)". Preisig proposes that this orthography can also be used by speakers of other dialects with some explanation on how alternate pronunciation could be associated with the various graphemes and orthographic conventions of this orthography (Suksavang et al 1994), an assertion that requires further testing.

Systematic linguistic analysis and an effort to develop a written form of Kmhmu' began in the 1950s with the work of William Smalley. According to an account by Preisig[1], he was the first to use the Lao script to write Kmhmu', though this script did not come into wide-spread use at that time. Perhaps the earliest attempt at writing Kmhmu' was made by Mrs. C.H. Crooks who used the northern Thai script to print a translation of the Gospel of Mark (Svantesson 1983:1, Smalley 1963:75). Another of the first pieces of literature produced in Kmhmu' utilized a Roman script, a Kmhmu'-French dictionary developed in 1964 by a French teacher for use in the classroom (Suksavang et al 1994). A French priest, Father Bonometti, also used a Roman script-based orthography in his translation of Scripture portions printed during the 1960s.

Preisig and Suksavang were instrumental in refining the Kmhmu' Lao script-based orthography. The first description of this orthography was prepared by Preisig in 1990 as an unpublished manuscript entitled The Kmhmu' Orthography Dialect of Xieng Khouang, Luang Prabang and Sam Neua, and the orthography has not been changed since that paper was written (Preisig 1990). The Kmhmu' Lao script orthography was first used in an official publication in 1994, the Kmhmu'-Lao-French-English Dictionary (Suksavang et al 1994). This orthography statement is an effort to formalize the description prepared by Preisig (1990). Section 2 describes how the Lao script is used to represent the Kmhmu' phonemes. Special attention is given to explaining the rationale behind the various solutions employed for representing Kmhmu' phonemes that are not found in Lao. Section 3 contains a description of Kmhmu' word structure, an

[1] personal communication 2010

MILLER, Michelle. 2013. A description of Kmhmu' Lao script-based orthography.
Mon-Khmer Studies. 42:12-25

explanation of the orthographic conventions for writing words and a description of conventions that have been adopted for punctuation. Section 4 contains a description of how the Kmhmu' orthography is being used followed by a review of teaching/learning experiences with this orthography. Conclusions and outlook are in Section 5. The language data used in this description comes from Osborne (2013) and Suksavang et al (1994).

2. Kmhmu' orthography

The Southern dialect of Kmhmu' has 36 consonant phonemes. All of them occur in initial position and 16 of them form codas. There are 17 initial consonant clusters. The Kmhmu' orthography uses the Lao graphemes to represent the same sound-symbol correspondence in Kmhmu' and Lao where possible. Because the sound inventories of Lao, a Tai-Kadai language, and Mon-Khmer Kmhmu' differ considerably, it was necessary to adapt the use of the Lao graphemes and writing conventions to accommodate a Kmhmu' orthography. Specific adaptations are described in the sections below, grouped according to initials, clusters, codas, and vowels.

2.1 Initial consonants

Table 1 below presents a summary of all 36 initial consonant phonemes. Various solutions were employed to represent these phonemes, and explanations of these solutions are presented in this section.

Table 1: The 36 Initial Kmhmu' consonant phonemes and graphemes; consonants not found in Lao are in shaded cells.

Manner of articulation	Point of articulation									
	bilabial		alveolar		palato-alveolar		velar		glottal	
voiceless unasp. stops	p	ป	t	ต	ʨ	จ	k	ก	ʔ	อ
voiceless asp. stops	pʰ	ผ	tʰ	ท	ʨʰ	ฉ	kʰ	ข	--	
voiced stops	b	บ	d	ด	ɖ	ฮ	g	ฆ		
pre-glottalized nasals	ʔm	ม่	ʔn	น่	--		ʔŋ	ง่		
voiceless nasals	m̥	หม	n̥	หน	ɲ̊	หย	ŋ̊	หง	--	
voiced nasals	m	ม	n	น	ɲ	ย	ŋ	ง		
pre-glottalized approx.	ʔw	อ่	---		ʔj	ย่	--		--	
voiceless approximants	w̥	หอ			j̊	หย				
voiced approximants	w	อ			j	ย				
fricatives	--		s	ส	--		--		h	ห
voiceless lateral	--		l̥	หล	--		--		--	
voiced lateral			l	ล						
voiceless trill	--		r̥	หอ	--		--		--	
voiced trill			r	อ						

There are 17 initial Kmhmu' consonant phonemes that are not found in Lao. The solutions used to represent these non-Lao sounds can be grouped into three categories, namely 1) creation of a new grapheme, 2) assignment of a new value to an existing grapheme or diacritic and 3) the use of special characters. The sections below describe how these solutions were applied.

2.1.1 Creation of a new grapheme

Voicing is a Kmhmu'-specific contrastive feature not found in Lao. Lao does not have a voiced velar stop, /g/, or a voiced palato-alveolar stop, /ɖ/. In the Lao grapheme inventory there are not any available graphemes for representing additional velar phonemes, and thus a completely new grapheme was created to represent the voiced velar stop in Kmhmu'. The new grapheme was

created to resemble the Lao grapheme for the consonant-initial voiceless velar stop /k/ which is {ก}. The grapheme created for /g/ in Kmhmu' is {ก̚}. The representation of the voiced palato-alveolar stop was derived via a different solution (see section 2.1.1.2).

Example word in Kmhmu': /gul/ ກູລ 'to be fat'

2.1.2 Assigning new value to an existing grapheme or diacritic

Existing Lao graphemes and diacritics with phonemic values foreign to Kmhmu' were used to symbolize preglottalisation, another Kmhmu'-specific contrastive feature, as well as several other non-Lao phonemes, namely the voiced and voiceless palato-alveolar stops, and the voiceless sonorants.

2.1.2.1 Voiceless palato-alveolar stop /ʨʰ/

The Lao grapheme {ຊ} is used to represent the voiceless aspirated palato-alveolar stop. In the Lao orthography this grapheme represents the low-class consonant phoneme /s/. Kmhmu' also has the phoneme /s/, but there is another Lao grapheme available to represent /s/ in Kmhmu', the high-class grapheme {ສ}.

Example word in Kmhmu': /ʨʰɔʔ/ ຊ໋າະ 'to put away in a corner'

2.1.2.2 Voiced palato-alveolar stop /ʥ/

The Lao grapheme {ຈ} is used to represent the voiced palato-alveolar stop. In the Lao orthography this grapheme represents the diphthong /ia/. The grapheme {ຈ} was likely chosen to represent initial consonant /ʥ/ because the diphthong /iɑ/ approximates a palatal place of articulation. Admittedly, its use as a consonant in the Kmhmu' orthography may be somewhat confusing for new readers of this orthography, though its use as an initial consonant grapheme is not ambiguous in any context.

It should be noted that using the grapheme {ຈ} as a consonant will necessitate special considerations for font design. Because the 'tail' of this grapheme hangs below the baseline it will collide with the vowel grapheme the /u/ which is written below the consonant grapheme unless adjustments are made in the font.

Example word in Kmhmu': /ʥɔʔ/ ເຈ໋າະ 'great grandchildren'

2.1.2.3 Pre-glottalized initial sonorants /ˀm/ /ˀn/ /ˀŋ/ /ˀj/ /ˀw/

The Kmhmu' orthography uses the Lao tone diacritic ◌̈ to indicate pre-glottalization of three nasals and two approximants shown below. The southern Kmhmu' dialect does not have phonemic tone, so this diacritic is not needed for the purpose of representing tone in Kmhmu'.

nasals: /ˀm/ - {ม̈}, /ˀn/ - {น̈}, /ˀŋ/ - {ง̈}

approximants: /ˀj/ - {ย̈}, /ˀw/ - {ว̈}

Example words in Kmhmu':

/ˀm/ → /ˀmɑŋ/ ม̈ั่ง 'to be hidden'

/ˀn/ → /ˀnɨɑn/ เนื่̈อน 'month'

/ˀŋ/ → /siːm ŋ̊ɛːk/ สีม เง่̈า 'black toucan'

/ˀj/ → /ˀjɑːŋ/ ย่̈าง 'round basket for carrying poultry'

/ˀw/ → /ˀwɑːt ˀwɑːt / ว่̈าด ว่̈าด 'loud booming sound'

2.1.2.4 Voiceless initial sonorants /m̥/ /n̥/ /ɲ̥/ /ŋ̊/ /w̥/ /j̊/ /l̥/ /r̥/

Kmhmu' has eight voiceless sonorant consonants. Lao does not have voiceless sonorant consonants therefore some innovation was required to represent this feature. In the Kmhmu' orthography the Lao grapheme for /h/, {ຫ}, is written preceding the grapheme of the voiced form of the consonant to form a digraph as shown below. In the Lao orthography, the same grapheme, {ຫ}, is used as a diacritic with six of the low class consonants to indicate tone (Becker 2003), therefore new readers of Kmhmu' who can read Lao need to learn the new meaning of the diacritic {ຫ}.

nasals: /m̥/ - {ຫມ}, /n̥/ - {ຫນ}, /ɲ̥ / - {ຫຍ}, /ŋ̊ / - {ຫງ}

approximants: /w̥/ - {ຫວ}, /j̊ / - {ຫຢ}

trill: /r̥/ - {ຫຣ} lateral: /l̥/ - {ຫລ}

Example words in Kmhmu':

/m̥/ → /m̥aŋ/ ຫມັ້ງ 'old'

/n̥/ → /n̥ɨm/ ຫນຶ່ມ 'young'

/ɲ̥/ → /ɲ̥aːm/ ຫຍາມ 'to be used'

/ŋ̊/ → /siːm ŋ̊ɛːk / ຢສີມ ແຫງກ 'black toucan'

/w̥/ → /w̥at/ ຫວັດ 'to throw'

/j̊/ → /j̊aːŋ/ ຫຢາງ 'female' (animal)

/l̥/ → /l̥aːŋ/ ຫລາງ 'classifier for traps'

/r̥/ → /r̥aːŋ/ ຫຣາງ 'tooth'

2.1.3 Use of special characters

In Lao there is a no voiced trill /r/, but because there are many borrowed words particularly from Thai that have a syllable-initial voiced trill, a special character that is no longer part of the official Lao alphabet (Becker 2003) is used to represent the voiced trill in borrowed words, namely {ຣ}. The Kmhmu' phoneme inventory includes both a syllable-initial and syllable-final voiced trill, and the special character {ຣ} is used for this sound in Kmhmu' also.

Example word in Kmhmu': /raːŋ/ ຣາງ 'flower'

2.2 Consonant Clusters

Lao does not have consonant clusters, but Kmhmu' has them in syllable-initial position. Consonant clusters are found only in major syllables; minor syllables do not have consonant clusters (see sections 3.2 and 3.3 for discussion of major and minor syllables). Consonant clusters in Kmhmu' are written with the two consonant graphemes of the cluster in a sequence. There is a restricted set of consonant phonemes found in both the first (C_1) and second (C_2) elements of the cluster forming a total of 16 clusters.

Table 2: Consonant cluster (C_1 and C_2) phonemes and graphemes

C_1		C_2 w	◌	l	◌	r	◌
p	ป			pl	ปล	pr	ปร
p^h	ผ					p^hr	ผร
b	บ			bl	บล		
t^h	ท					t^hr	ทร
d	ด					dr	ดร
t^h	ฐ					t^hr	ฐร
ɖ	ฏ					ɖr	ฏร
k	ก	kw	กว	kl	กล		
k^h	ค	k^hw	คว	k^hl	คล	k^hr	คร
g	ฆ	gw	ฆว			gr	ฆร
s	ส					sr	สร

2.3 Codas

Of the 16 Kmhmu' final consonant phonemes six are not found in Lao: /t̚/, /h/, /r/, /l/, /ŋ̚/ and /j̚/. Table 3 below presents a summary of all 16 Kmhmu' final consonant phonemes. The solutions used to represent these non-Lao sounds can be grouped into three categories, namely 1) use of the corresponding Lao consonant-initial grapheme, 2) creation of a new grapheme and 3) assignment of a new value to an existing grapheme. Detailed explanations of these solutions are presented below.

Table 3: The 16 Kmhmu' final consonant phonemes and graphemes

Manner of articulation	Point of articulation bilabial		alveolar		palato-alveolar		velar		glottal	
voiceless unasp. stops	p	บ	t	ด	t̚	จ	k	ก	ʔ	ะ
voiced nasals	m	ม	n	น	ŋ̚	ย	ŋ	ง	--	
voiced approx.	w	ว	--		j	ย	--		--	
fricatives	--		--		j̚	ยห	--		h	ห
voiced lateral	--		l	ล	--		--		--	
voiced trill	--		r	ร	--		--		--	

2.3.1 Use of the corresponding Lao consonant-initial grapheme

Smalley (1963) espouses adherence to the "phonemic principle", namely that "every distinctive sound is represented by one symbol and only one in the writing system" (Smalley 1963:38). This principle was followed in determining graphemes for the final consonants in Kmhmu' that are not found in Lao where possible. The following sections describe how this was applied for four consonant-final Kmhmu' phonemes.

2.3.1.1 Palato-alveolar stop /t̚/

The grapheme chosen to represent /t̚/ is {จ}.

Example word in Kmhmu': /ho:t̚/ โจ๋จ 'to be finished'

2.3.1.2 Voiceless glottal fricative /h/

The grapheme chosen to represent /h/ is {ห}.

Example word in Kmhmu': /tuh/ ตุ๋ห 'old rice field'

2.3.1.3 Voiced lateral approximate /l/

The grapheme chosen to represent /l/ is {ລ}.

Example word in Kmhmu': /kɑːl/ ກ່ອນ 'before'

2.3.1.4 Voiced trill / r /

The grapheme chosen to represent /r/ is {ຣ}.

Example word in Kmhmu': /kɑːr/ ກ່ານ 'to grill'

2.3.2 Creation of a new grapheme

The initial consonant phoneme /ŋ/ in Lao is represented by the grapheme {ໝ}. This grapheme could not be used for the final consonant phoneme /ŋ/, however, because in the Lao orthography the grapheme {ໝ} in syllable-final position represents the phoneme /j/, not /ŋ/ as it does in the initial position. It was therefore necessary to find another character to represent /ŋ/ in syllable-final position. The solution was to design a new character, the grapheme {ᩲ}.

Example word in Kmhmu': /tɑːŋ/ ต่าᩲ 'to weave'

2.3.3 Assignment of a new value to an existing grapheme

The voiceless palatal approximant /j̊/ in syllable-final position is represented by the digraph {ยห}. The voiceless palatal approximant in initial position is represented as {ຫຢ} and therefore some discussion is warranted to explain why different graphemes are used for this phoneme in initial and final position.

The Lao orthographic convention marks word-final /j/ as {ย}, and it is thus logical to use the {ย} in this grapheme. The digraph {ยห} adds devoicing to the word-final palatal semi-vowel. However, there is a phonetic difference between syllable- initial and –final /j̊/. There is a voiced transition from a syllable-initial /j̊/ into the following vowel whereas the syllable-final /j̊/ has a voiced transition caused by the preceding vowel before it becomes fully voiceless. The grapheme {ຫຢ} for syllable-initial /j̊/ and {ยห} for final /j̊/ thus reflect the pronunciation variants for this one phoneme in different positions within a word.

Example word in Kmhmu': /rɨɑj̊/ เรือยห 'to cease (of rain)'.

2.4 Kmhmu' vowels

Kmhmu' has 10 vowel qualities which all occur in both short and long forms (Osborne 2013). There are also three diphthongs (Osborne 2013). Vowel quality contrast is neutralized in open syllables where all vowels and diphthongs are long. Neutralization also occurs in syllables ending on /h/, /ʔ/, /j̊/, and /j/ where all vowels are short (Osborne 2013). All of the Kmhmu' vowels except one and all three diphthongs are found in Lao. The Kmhmu' orthography can therefore use the Lao graphemes in the same sound-symbol correspondence, with the addition of two graphemes {ᩮ} and {ᩯ} representing the long and the short forms respectively of the near-open central vowel /ɐ/. Table 4 presents a complete listing of all vowels and corresponding Kmhmu' graphemes. As in Lao, the representation of some vowels is dependent on syllable structure, with some vowels being written differently in open and closed syllables.

2.5 Special symbols representing vowel-consonant combinations

The Kmhmu orthography has incorporated the use of the Lao special characters representing CV combinations /ɑm/, /am/, and /ɑw/, realized as {ำ}, {ไ}, and {ᩦ} (for convenience these are also included in Table 4).

Table 4: Kmhmu' vowel phonemes and graphemes; shaded graphemes denote the two vowels not found in Lao

	Front Unrounded					
	Short			Long		
	IPA	cv	cvc	IPA	cv	cvc
Close	i	◌ิ	◌ิ◌	i:	◌ี	◌ี◌
Mid	e	เ◌ะ	เ◌็◌	e:	เ◌	เ◌◌
Near- open	ɛ	แ◌ะ	แ◌็◌	ɛ:	แ◌	แ◌◌
Open	--			--		
Diphthongs	ia	เ◌ีย	เ◌ียว	ia:	เ◌ีย	เ◌ียว
Special symbols	am	◌ำ		--		
	Central Unrounded					
	Short			Long		
	IPA	cv	cvc	IPA	cv	cvc
Close	ɨ	◌ึ	◌ึ◌	ɨ:	◌ึ	◌ึ◌
Mid	ə	เ◌ิ	เ◌ิ◌	ə:	เ◌ิ	เ◌ิ◌
Near- open	ɐ	แ◌ิ	แ◌ิ◌	ɐ:	แ◌ิ	แ◌ิ◌
Open	ɑ	◌ะ	◌ั◌	ɑ:	◌า	◌า◌
Diphthongs	ɨɑ	เ◌ือ	เ◌ือ	ɨɑ:	เ◌ือ	เ◌ือ◌
Special symbols	ɑj	◌ไ		--		
	Back Unrounded					
	Short			Long		
	IPA	cv	cvc	IPA	cv	cvc
Close	u	◌ุ	◌ุ◌	u:	◌ู	◌ู◌
Mid	o	โ◌ะ	โ◌็◌	o:	โ◌	โ◌◌
Near- open	ɔ	เ◌าะ	เ◌ั◌	ɔ:	◌ॅ	◌ॅ◌
Open	--			--		
Diphthongs	ua	◌ัวะ	◌ัว◌	ua:	◌ัว	◌ัว◌
Special symbols	ɑw	เ◌ า		--		

3 Orthography conventions

This section describes the conventions for writing Kmhmu' words, borrowed words and punctuation.

3.1 Kmhmu' word structure

A brief description of Kmhmu' word structure will be presented in this section because the rationale for various orthographic conventions for writing Kmhmu' words are related to features of the word structure. Kmhmu' words in terms of their syllabic composition are usually mono- or sesquisyllabic (Matisoff 1973). There are a small number of di- and tri-syllabic words.

Monosyllabic words consist of one syllable, referred to as a major syllable in multisyllable words, which may be open or closed, and can contain the full range of vowel qualities (Osborne 2012).

Sesquisyllabic words consist of a major syllable preceded by a minor syllable, or as described by Gafos as "a heavy syllable which is optionally preceded by a vowelless syllable made

up of one or two consonants" (1999:119). The use of the term 'minor syllable' in this paper is in accordance with the criteria described by Herr (2011) as 1) a syllable that cannot exist independently of the major syllable, 2) has vowel neutralization, 3) a reduced consonant inventory and 4) lacks prosodic features such as vowel length, tone and stress. According to Herr (2011) the term 'minor syllable' should not be used for the initial syllables in multi-syllable words in which there is merely reduced or limited vowel contrast (either in vowel quality or length). She prefers the term "major syllable with reduced vowel contrast" or "reduced syllable" (Herr 2011:25) for such syllables.

It could be argued that in some dialects of Kmhmu', namely one or more of the 'Northern' tonal dialects (Svantesson 1983), the term 'minor syllable' is not completely appropriate according to the definition stated above because, according to Svantesson (1983), there are both tonal and non-tonal minor syllables. Svantesson and Karlsson (2004 cf Herr p. 25) site examples in 'Northern' Kmhmu' dialects in which the tone of the minor syllable is contrastive, although minimally so in that only about ten minimal pairs have been identified (2004:2 cf Herr p. 26). The Southern dialect of Kmhmu' for which this orthography has been designed, has no contrastive tone (Osborne 2012) and thus the term 'minor syllable' is appropriate.

Di- and tri-syllabic words are rare. Disyllabic words have two major syllables. There are a few trisyllabic words that contain two minor syllables preceding a major syllable, or a minor syllable followed by two major syllables in which the second major syllable is reduplicated (Suksavang et al. 1994).

In the remainder of this paper when transcribing syllable structure, a lowercase 'v' will be used to denote the vowel of the minor syllable, given the greatly reduced nature of this vowel in both vowel quality and length, and the uppercase 'V' will be reserved for transcribing vowels with full vowel characteristics. This distinction in transcription is being made in order to eliminate ambiguity in transcriptions of sesquisyllabic or multisyllabic words.

3.2 Monosyllabic words (or major syllables)

Spelling monosyllabic words (or the major syllable in sesquisyllabic or multisyllable words) is quite straightforward in that each phoneme in the word (or syllable) has explicit graphemic representation, i.e., every sound has a corresponding character. Table 5 presents the Kmhmu' syllable patterns for monosyllabic words with examples.

Table 5: Single-syllable words or major syllables

Syllable structure	IPA	Kmhmu'	Meaning
CV	tu:	ຕູ	'to falsely accuse'
CCV	pʰrɨa	ເພຍີ	'fire'
CVC	kuŋ	ກຸ	'village'
CCVC	plɔ:ŋ	ປລອ	'calf' (of leg)

There is potential for ambiguity in pronouncing single syllable words that are of the CCV pattern in which the vowel or part of the vowel is written before the consonant cluster. There are two situations in which ambiguity may occur:

(1) The consonant cluster may be interpreted as two separate phonemes in words in which the part of the vowel symbol is written before the consonant cluster and part is written after (or above), such as is the case with the vowels {ເ◌ະ}, {ແ◌ะ} or {ໄ◌ะ}. For example, consider the word /ple'/ 'fruit', written in Kmhmu' as {ເປລະ} (Suksavang et al). This word could be pronounced as the two-syllable word /pe lɑ'/ {ເປ - ລະ}. However, the two-syllable pronunciation has no meaning. The reader must therefore rely on the meaning of the word in its context in order to determine the correct pronunciation in such situations, albeit rare.

(2) The consonant cluster may be interpreted as two separate phonemes (or conversely two separate phonemes may be interpreted as a consonant cluster) in words with long vowels in which the vowel grapheme is written before the consonant cluster such as with the vowels {ໃ◌}, {เ◌}, or {แ◌}. For example, the word /kool/, written in Kmhmu' as {ໂกล}, could be read as /ko:l/ or as /klo:/. Again, the reader must rely on the meaning of the word in context in order to determine the correct pronunciation. In both of these situations it has been observed that new readers may hesitate when encountering such structures, but with practice they gain automaticity in word recognition in context.

3.3 Sesquisyllabic words

Sesquisyllabic words consist of one major syllable preceded by one minor syllable. As described previously, writing the major syllable of a sesquisyllabic word is quite straightforward because each phoneme is represented by a grapheme. The method for writing the minor syllable, particularly the vowel, is determined by whether the minor syllable is open (Cv) or closed (CvC), and is described in sections 3.3.1 and 3.3.2 below. There is a restricted set of consonants that is found in the initial and final consonant position of the minor syllable, and the inventory of the initial consonant set is also dependent on whether the minor syllable is open or closed. Initial and final consonant inventories of minor syllables are presented in Table 6.

Table 6: Minor syllable consonant phonemes and graphemes

Initial consonants: Cv	Initial consonants: CvC	Codas: CvC
p ป	p ป	p บ
t ต	t ต	t ด
tʰ ฐ	tʰ ฐ	k ก
t จ	t จ	m ม
k ก	k ก	n น
g ฆ	--	ŋ ง
s ส	s ส	ŋ ว
h ห	h ห	l ล
l ล	l ล	r อ
r อ	r อ	--

3.3.1 Open minor syllables (Cv)

In minor syllables of the Cv structure *a vowel grapheme is not written*, only the initial consonant grapheme is written. The major syllable of the word is written connected to the minor syllable to form a single unit.

Example word: Cv.CVC → /tɤ̆ˈ.lɔːŋ/ → จลอง 'boat'

The vowel is not written in the minor syllable of the Cv structure because the vowel quality is reduced and difficult to distinguish, the length of the vowel is always extra short, and writing a vowel grapheme may inadvertently cause the reader to read the syllable as a full-length, accented syllable which it is not. In fact for Kmhmu', Shaw (1993 cf. Gafos 1999:120) proposes that minor syllables with a single consonant be assigned "no mora" in his moraic model of syllables.

The vowel is not written in the minor syllable of the Cv structure because the vowel quality is reduced and difficult to distinguish, the length of the vowel is always extra short, and writing a vowel grapheme may inadvertently cause the reader to read the syllable as a full-length, accented syllable which it is not. In fact for Kmhmu', Shaw (1993 cf. Gafos 1999:120) proposes that minor syllables with a single consonant be assigned "no mora" in his moraic model of syllables. There are variant pronunciations of the non-contrastive vowels in the pre-syllable based on regional accents, though generally the vowel quality is a central unrounded vowel area [ɨ - ə - a] (Svantesson and Karlsson 2004:1; Suksavang et al 1994), and if speakers tried to spell the word as it is pronounced a confusing variety of spellings would likely result.

Because the vowel is not written in the minor syllables of the Cv structure, it is necessary to address the question of potential ambiguities in the reading of multisyllable words of both the Cv.CV(C) and Cv.CCV(C) word structures.

First, with words of the Cv.CV(C) structure, how will the reader know that the two consecutive initial consonants are to be read with a vowel between then and not as a consonant cluster or as a consonant phoneme whose grapheme is a digraph (note the only digraphs are the voiceless sonorant consonants), i.e., how will the reader know the word should be read as Cv.CV(C) multi-syllable word and not as a single-syllable word of the CCV(C) pattern or CV(C) (where C is a digraph)?

Because there is a restricted set of consonants that can be in the initial consonant position and there is a limited set of consonant clusters and digraphs, we can determine potential ambiguities by looking at the possible consonant combinations that may occur in a Cv.CV(C) word. The combinations that would be ambiguous in terms of how they should be read are those that mimic a consonant cluster or a digraph. Table 7 below presents a summary of possible consonants that may appear in 1) the consonant position of the Cv minor syllable, 2) the C2 position of a consonant cluster 3) consonant clusters and 4) digraphs.

Table 7: Determination of potentially ambiguous consonant combinations in Cv.CV(C) word constructions; shaded consonant clusters and digraphs are potentially ambiguous.

Minor syllable **C**v	Major syllable C**C**V(C)	Consonant clusters **CC**	digraphs ห + C	Summary of potentially ambiguous consonant combinations
p ป		kw กว	m̥ หม	pl ปร
t ต	w ว	k^hw คว	n̥ หน	
t^h ฐ		gw ฆว	ŋ̊ หย	
t จ		pl ปล	ŋ̊́ หง	
k ก		bl บล	m̥ หม	kw กว, kl กล
g ฆ	l ล	kl กล	w̥ หว	gw ฆว, gr ฆร
s ส		k^hl คล	j̊ หย	sr สร
h ห		pr ปร	r̥ หร	hw หว, hl หล, hl หร
l ล		p^hr ผร	l̥ หล	
r ร		t^hr ทร		
		dr ดร		
	r ร	t^hr ฒร		
		ḍr ฎร		
--		k^hr คร		
		gr ฆร		
		sr สร		
		pr ปร		

Of the potentially ambiguous consonant combinations summarized in Table 7, there is no data to support the existence of words with these consonant constructions. We can conclude, then, that there are likely no ambiguities for the reading of Cv.CV(C) words. When reading Cv.CV(C) the appearance of a 'cluster' may initially cause some confusion for new readers of Kmhmu' because in the Lao orthography there are not any unwritten vowels in multisyllable words. However, with practice, Kmhmu' speakers are able to read words of the Cv.CV(C) structure fluently.

Regarding the potential for ambiguities when reading Cv.CCV(C) words, there is no ambiguity as to how to read them because there is only one way to decode a word with three consecutive initial consonants, i.e. the word can only be read by inserting a vowel sound after the first consonant.

Example word of the Cv.CCV(C) pattern: /lɛˈdruŋ/ ລດຮຸ໋ງ ‘to fall down’ (hands and legs spread).

3.3.2 Closed minor syllables (CvC)

In contrast with the Cv minor syllables, a vowel is written in the CvC minor syllable. The same vowel grapheme is written in all closed minor syllables (CvC). The vowel chosen for this syllable structure is the close central vowel /ɨ/, realized as the grapheme {ิ}. There is variation in speaker pronunciation of this vowel sound in the minor syllable, and there is often vowel harmony between the vowel of the minor syllable and that of the main syllable, but the close central vowel was chosen because that is the vowel most frequently used by speakers (Suksavang et al 1994). The use of the same vowel in all CvC minor syllables helps to maintain consistency in spelling. The vowel cannot be left unwritten as in the Cv minor syllable because doing so would allow for an initial consonant sequence of four graphemes in a word, i.e, a two-syllable word could begin with CCCC, making de-coding too difficult. Because the vowel is marked in multi-syllable words with a CvC minor syllable, there is no ambiguity in reading words having CvC.CCV(C) or CvC.CV(C) structures.

Example word with CvC.CCV(C) structure: /tɨnˈdriɑh/ ຈຶ໋ນດຣັ໋ຽຫ ‘comb’

Table 8: Summary of Kmhmu’ word structure (adapted from Osborne 2013)

Frequency of occurrence	Examples			
	Syllable structure	IPA	Kmhmu'	Meaning
common structures	CV	tu:	ตู	to falsely accuse
	CVC	mar	ม้	snake
	CCV	pʰrɨɑ	เพอ๊อ	fire
	CCVC	plɔ:ŋ	ปลอ๋ง	calf (of leg)
	Cv.CV(C)	tɨ'lɛ:	จแล	spade for sowing
	Cv.CCV(C)	sɨ'klo:p	สโกลบ	to clutch many things concurrently
	CvC.CV(C)	sɨr'nɛ:	สึอแน	long iron stick for piercing a hole
	CvC.CCV(C)	tɨm'braʔ	ตึมบอะ	fireplace for cooking
rare structures	CvC.CvC. CVC	pɨnhɨn'koʔ	ปึนซึนโกะ	to dress someone
	Cv.CVC. CVC	tɑˌlɑmˈ pɑ:m	ตล๋ำปฟม	butterfly
	CV.CV	tɔˈrɔ:	จ๋อ๋	to approach

3.4 Loan words, punctuation and word breaks

Words that are borrowed from Lao are written with a Kmhmu’ spelling which reflects the Kmhmu’ pronunciation. Following is an example of a borrowed word written in the Kmhmu’ orthography:

(Lao) /kɑmlɑŋ ກຳລັ໋ງ / ‘in the process of’ vs. (Kmhmu’) /kɨmˈlɑŋ/ ก็มล้ง

The Kmhmu’ script utilizes a specially created half-space to separate words and a full space at the end of a sentence. Lao does not have word breaks, only clause and sentence breaks which serve the function of punctuation. Because the Kmhmu’ orthography does have word breaks it is necessary to introduce punctuation to identify clause breaks and sentence breaks, which are the comma and period respectively. Other punctuation markers were incorporated to indicate expression (question mark and exclamation mark) and to identify direct speech (double quotation mark).

4 The Kmhmu' orthography in use

As described in Section I, Preisig and Suksavang were the primary decision-makers in the development of the current orthography in the late 1980s. They were influential in bringing a degree of standardization to spelling through the production of the Kmhmu'-Lao-French-English dictionary (1994). These individuals continue to be influential in the promotion of the Kmhmu' orthography. The growing community of Kmhmu' that has knowledge of the orthography has found it acceptable and is using it. It is difficult to estimate how many people can read this orthography, but several hundred Kmhmu' can read this orthography with a moderate level of fluency.

The two unique Kmhmu' characters {ꭑ} and {ꭒ} were approved for inclusion in Unicode by iso/iec jtc1/sc2/wg2 in. This will aid in making the Kmhmu' font readily accessible.

4.1 Learning the Kmhmu' orthography

Because of the high degree of over-lap between the grapheme-phoneme correspondences in Lao and Kmhmu' a reading /writing guide that helps readers of Lao learn the Kmhmu' Lao-script based orthography was published in 2010. This guide begins by introducing the Kmhmu' graphemes that have the same sound-symbol correspondences in Lao and then leads the learners through the differences, providing reading practice and phonics exercises focused on to help the learner gain fluency. The contents of this 'transfer guide' can be taught adequately to readers of Lao in about 20 instructional hours, including practice. The elements of the orthography are introduced in the following order:

(1) Consonants and vowels that are the same in Lao/ Kmhmu'

(2) Final consonants in Kmhmu' not found in Lao

(3) Initial consonants in Kmhmu' not found in Lao

(4) Vowels in Kmhmu' not found in Lao

(5) Consonant clusters

(6) Use of the diacritics marking pre-glottalization and de-voicing

(7) Multisyllable words

The most difficult elements of the orthography to learn for Kmhmu' speakers who can already read Lao are the *six final consonants in Kmhmu' that are not found in Lao*. Through experiences teaching the Kmhmu' orthography it has been found that learners need assistance developing phonemic awareness of final consonant sounds in general in the process of learning to read and write. Interestingly, the uniquely Kmhmu' initial consonants prove to be quite easy to learn. Other difficult elements to learn are *consonant clusters*.

4.2 Use of the orthography by speakers of other dialects

Suksavang et al claim that this orthography can be used by speakers of the 'Northern' dialect by assigning new meaning to various elements of the orthography to correspond to the phonologic features of that dialect. No documentation of this having actually been done is available, so it is unknown how viable this approach is. I suspect, though, that given the large variety of sub-dialects it would be difficult to capture all the differences adequately.

However, speakers of other dialects have used the transfer guide with a couple of outcomes. Some speakers, when encountering a word that is pronounced differently in their dialect (for example with tone in stead of de-voicing) the speaker automatically reads the word the way he normally speaks it. Other people have used the transfer guide to *learn* the southern dialect, and are motivated to do so because it is the largest dialect and the one used in most radio broadcasts and available literature. There are significant vocabulary differences between dialects in additional to phonological differences.

4.3 Other Kmhmu' orthographies

There is a second Kmhmu' orthography that has been developed which uses a Roman script (Suksavang et el). Young people in particular are seemingly more motivated to learn the Roman script than the Lao script because it 1) is perceived as being more uniquely Kmhmu' and 2) is perceived to be more prestigious because of its association with French or English. The Roman script-based Kmhmu' orthography is significantly more difficult to learn, however.

Work has been done by Mahidol University and SIL International in conjunction with Kmhmu' communities in Thailand to develop a Thai-based orthography for the one of the northern dialects of Kmhmu' spoken in Chiang Rai province. The extent of its use in the community needs to be explored.

5. Summary and outlook

The Kmhmu' Lao script-based orthography is an efficient representation of the Southern Kmhmu' dialect and is being used by the Kmhmu' community as an effective tool for written communication. This orthography is particularly easy to learn for mother tongue speakers of the Southern Kmhmu' dialect who can also read and write Lao. The elements of the orthography that are not found in Lao, such as the uniquely- Kmhmu' final consonants and consonant clusters, require extra practice to learn in order to gain reading fluency. This orthography could also be used in a tother-tongue-first literacy program and would assist learners bridge into Lao language literacy.

Further study needs to be done to investigate whether or not this orthography can be used by speakers of the Northern Kmhmu' dialects by assigning a number of alternate phoneme-grapheme correspondences.

Acknowledgements

I want to express my sincere appreciation to Dr. Sigrid Lew, SIL International, for her assistance in helping me prepare this manuscript. Her insights and guidance were invaluable. I am also grateful to those who collected the data that is in this paper, particularly Suksavang Simana', Elizabeth Presig and Anne Osborne. Additionally, it is a privilege to work with the Kmhmu' people of Southeast Asia.

References

Becker, Benjawan P. 2003. Lao for beginners. Bangkok: Paiboon Pomsan Publishing.

Gafos, Adamantios I. 1999. The articulatory basis of locality in phonology. New York: Garland Publishers.

Herr, Kristen. 2011. The phonologic interpretation of minor syllables, applied to Lemi Chin. Chiang Mai, THN Payap University M.A. thesis.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. Consonant types and tone, ed. by larry Hyman. Southern California Occasional Papers in Linguistics. 1.71-95.

National Statistics Center of the Lao PDR. 2007. http://www.nsc.gov.la (Accessed 19 October 2009).

Osborne, R. Anne. 2013. Phonology Statement for Kmhmu'. To appear in Mon-Khmer Studies Journal.

Preisig, Elisabeth. 1990. The Kmhmu' orthography: Dialect of Xieng Khouang, Luang Prabang and Sam Neua. Unpublished manuscript.

Smalley, William. 1961. Outline of Khmu structure. An essay of the American Oriental Society No 2, ed. by Edward H. Schafer. New Haven: American Oriental Society.

Smalley, William. 1963. Orthography Studies: Articles on new writing systems. London: United Bible Societies in cooperation with the North-Holland Publishing Company, Amsterdam.

Suksavang Simana', Somseng Sayavong and Elisabeth Preisig. 1994. Kmhmu'-Lao-French-English dictionary. Vientiane, Lao P.D.R.: Ministry of Information and Culture, Institute of Research on Culture.

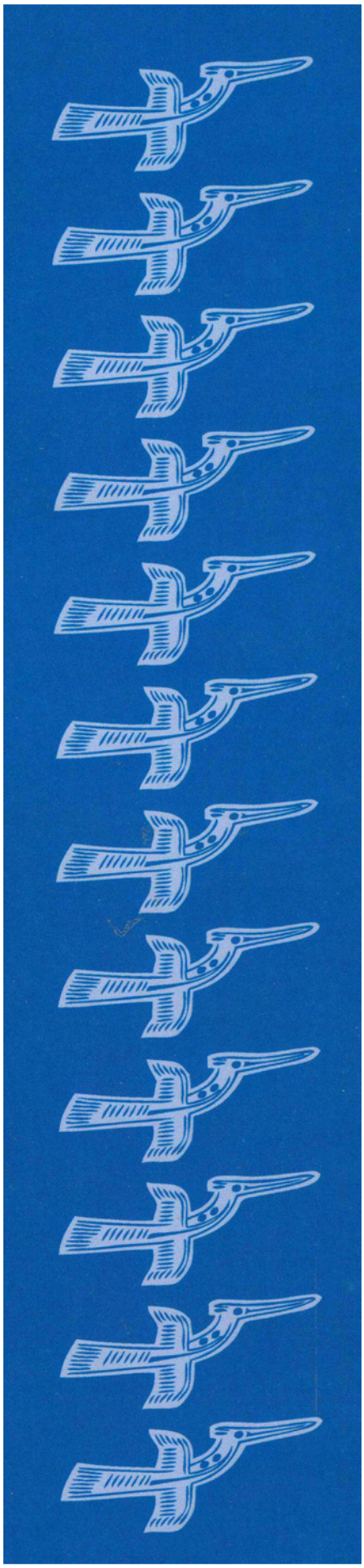
MILLER, Michelle. 2013. A description of Kmhmu' Lao script-based orthography.
Mon-Khmer Studies. 42:12-25

Suwilai Premsrirat. 2002. Thesaurus and Dictionary Series of Khmu dialects in Southeast Asia. Mon-Khmer Studies-Mahidol University Special Publications No.1. Institute of Language and Culture for Rural Development, Mahidol University, Salaya.

Svantesson, Jan-Olaf. 1983. Kammu phonology and morphology. Malmo, Sweden: CWK Gleerup.Svantesson, Jan-Olof. 1989. Tonogenetic mechanisms in Northern Mon-Khmer. Phonetica 46.60-79.Svantesson, Jan-Olaf & Anastasia Mukhanova Karlsson. 2004. Minor syllable tones in Kammu. International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages. Beijing, China, March 28-31, 2004, 1-4.

Mon-Khmer Studies

Volume 42



Authors: Elizabeth HALL
Title: *A phonological description of Muak Sa-aak*
Pages: 26-39
Date received: 3/7/2012
Revised Text Accepted: 14/8/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A phonological description of Muak Sa-aak

Elizabeth HALL
Linguistics Institute, Payap University

Abstract
The Austroasiatic language Muak Sa-aak, belonging to the Angkuic branch of Eastern Palaungic, is a tonal language spoken in Eastern Shan State of Myanmar and in China. This paper provides a phonological description of a variety spoken in Eastern Shan State. Like other Angkuic languages, Muak Sa-aak has undergone a shift whereby proto voiced stops have become voiceless and voiceless stops have become aspirated. However, the language does have the voiced stops /b, d/, due to borrowing. Despite the development of tone, Muak Sa-aak retains contrastive vowel length. Another surprising feature of this language is the phenomenon of final sonorant lengthening for short vowels.
Keywords: Palaungic, Angkuic, phonology
ISO 639-3 language codes: tlq

1. Introduction

Languages of the Palaungic branch of Mon-Khmer are located mainly in Southern China and in Myanmar, as well as in Thailand and also in Laos. Among these, the known Angkuic languages are located primarily in China, although some have previously been documented in other places: Samtao in Myanmar, Mok in Thailand, and Kiorr in Laos (SIL Ethnologue 2009). For Eastern Palaungic language groups, there is published work on Waic languages, including Wa and Plang varieties (such as Diffloth 1980, Paulsen 1992, and Watkins 2002), and some on Lametic languages (such as Narumol 1982 and Conver 1999). There is not, however, very much recent published work available on Angkuic languages. The areas where they live are places that have been, for the most part, difficult for outside researchers to access in recent years. U and Hu have been studied by Svantesson (1988, 1991), who gives a listing of available Angkuic wordlists (1988). Some limited data is also available on Man Met (cited in Diffloth 1991) and on Mok (Wenk 1965, Diffloth 1982[1]). Other Angkuic languages are known only from nineteenth and early twentieth-century wordlists.

The old distinction between the proto *h- and *s- initial consonants, lost in other Palaungic languages, is still maintained in the Angkuic languages (Diffloth 1977). Angkuic languages underwent a Germanic shift in the initial consonants (Svantesson 1991) so that old voiced stops became voiceless, and old voiceless stops became aspirated. Some Angkuic languages have also developed tone and denasalization of final nasals. The latter is seen in P'u-man and Pou Ma as well as in U (Svantesson 1988).

This paper provides a full phonological inventory of the Angkuic language Muak Sa-aak.[2] The Muak Sa-aak people live primarily in the eastern part of Shan State of Myanmar, in Mong Yawng Township. Some of their villages are located near Mong Yawng, and some are near the Chinese border, in what is called Special Region #4. At least two villages are located across the border in China. There are reported to be some Muak Sa-aak people in Thailand as well, although no village locations are currently known; it is not known if they would be the same as the speakers of the language listed in the SIL Ethnologue as Mok (Lewis 2009). The estimated population total is 4,460 in Myanmar and China (Hopple 2007, unpublished).

[1] According to Svantesson (1988: 76), Wenk's "Ya Ang Lawa" is the Mok presented by Diffloth (1982).
[2] A more detailed description of Muak Sa-aak phonology is given in the underlying MA thesis by the author.

Hall, Elizabeth. 2013. A phonological description of Muak Sa-aak.
Mon-Khmer Studies. 42:26-39

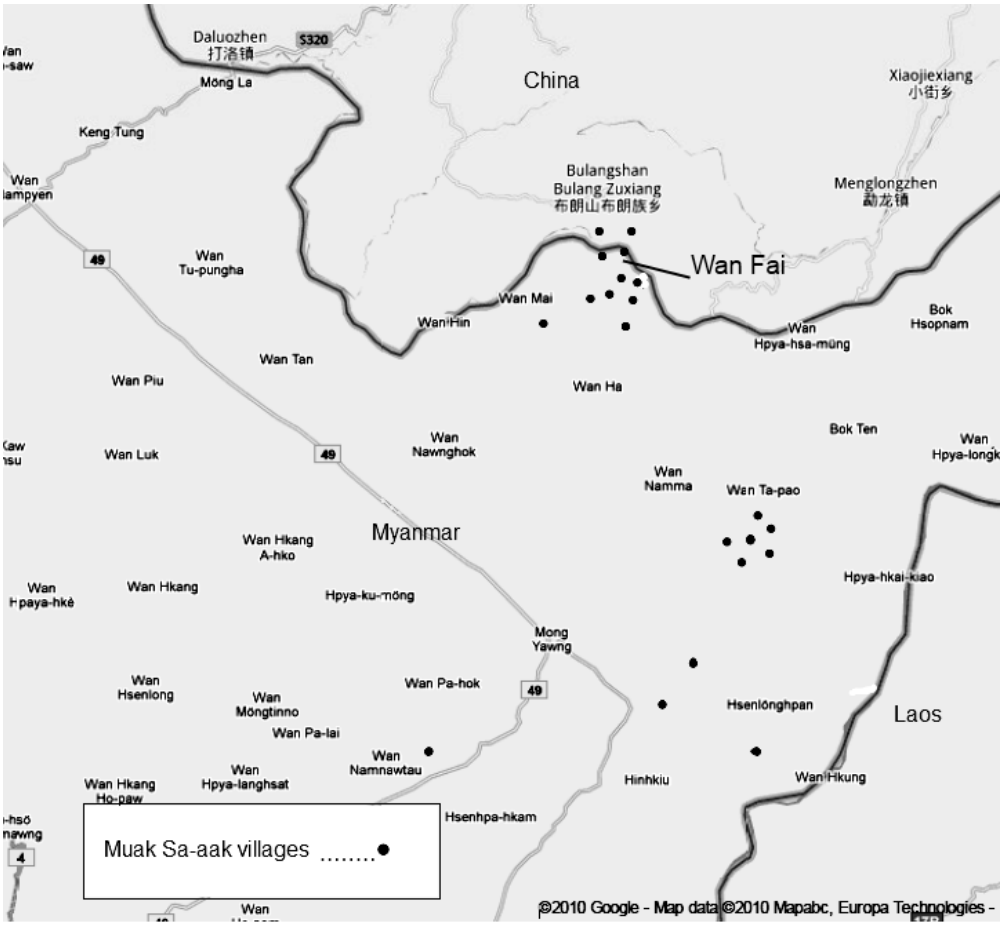


Figure 1. Map of known Muak Sa-aak villages, Eastern Shan State (adapted from Hopple 2007, unpublished)

The position of Muak Sa-aak in relation to other known Palaungic languages is shown in Figure 2.

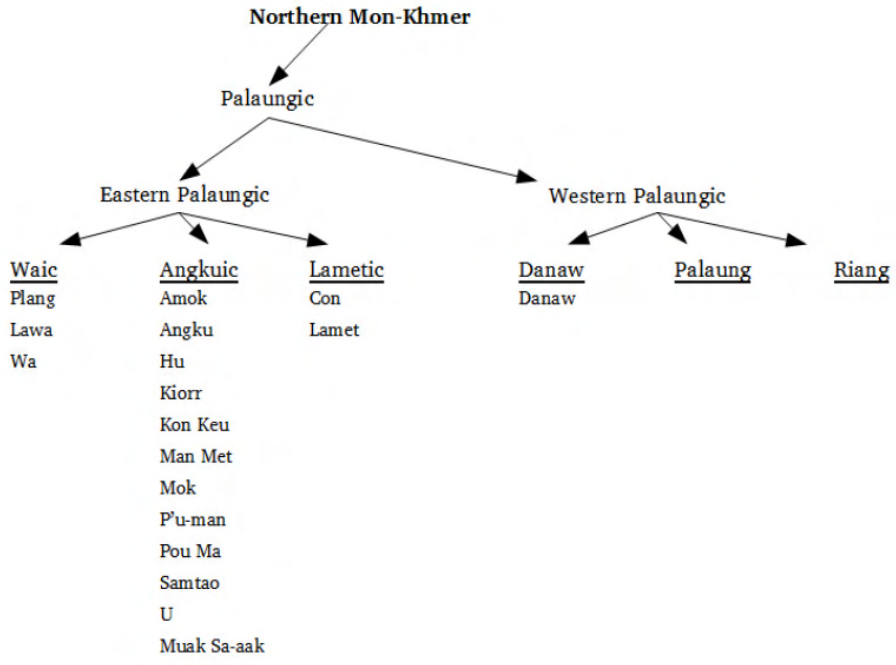


Figure 2. Language family tree for Muak Sa-aak (adapted from SIL Ethnologue 2009)

The data in this paper was gathered from speakers from an old village called Wan Fai in eastern Shan State, Myanmar with a population total of about 620 people (Hopple 2007, unpublished). The 1,659 item wordlist that this analysis is based on was fully or partially recorded from four Muak Sa-aak mother tongue speakers in their early forties to mid fifties in 2007 and 2008.

The Muak Sa-aak are currently listed in the SIL Ethnologue (16th Edition) under two names. The first is "Mok" (in Thailand), for which they are listed as an alternate name. This listing, however, is still uncertain. The second is the name "Tai Loi," for which they are listed as a dialect ("Saneung Muak;" saneun means "language") (Lewis 2009). According to speakers, the Muak Sa-aak are typically called Tai Loi or Tai Doi by outsiders, "Loi" or "Doi" being the Shan and Tai Lue words for mountain. According to Lebar, Hickey and Musgrave, the Shan sometimes call the Palaung "Kunloi," or "mountaineer" (1964: 121) and the Buddhist Wa have been called "Tai Loi" and "Hkun Loi" (1964: 129). Scott also speaks of Tai Loi as being a generic term of reference to hill groups which have become Buddhist, but also principally meaning Buddhist Wa, also called "Wa Küt" (J. George Scott and J. P. Hardiman 1900: 517)

2. Word structure

Muak Sa-aak words follow general Mon-Khmer word structure in being mono- and sesquisyllabic. There is active borrowing from the neighboring Tai Lue (Tai-Kadai language), so many words are loan words. They are realized following Muak Sa-aak phonology and phonotactics. Taking into account main syllable structure, reduced syllables, and tone, the overall word structure can be represented as follows:

$(C).(C)(C)V(C)^{T}$

Reduced syllables in Muak Sa-aak have a limited inventory of onsets /p, p^{h}, t, k, k^{h}, m, s/, and a non-distinctive very short vowel which is not transcribed.

Compounding is very common, as in *mul³.tʰi²* 'silver-hand' =bracelet, or *cʰak².ŋa:j³.laŋ³* 'seed-face-black' = pupil.

Examples for Muak Sa-aak word structure (1-10) are given below.

(1)	CV^{T}	*ci²*	'do'
(2)	CVC^{T}	*puk²*	'rotten'
(3)	CCV^{T}	*kra:³*	'mat'
(4)	$CCVC^{T}$	*kʰrɛp²*	'fish scale'
(5)	$C.CV^{T}$	*k.tɯ²*	'nose'
(6)	$C.CVC^{T}$	*k.can³*	'stand up'
(7)	$C.CCV^{T}$	*t.krɔ²*	'peel/ shell'
(8)	$C.CCVC^{T}$	*t.prɯ:t¹*	'swallow'
(9)	V^{T}	*ɔ:¹*	'cheek'
(10)	VC^{T}	*ɛl³*	'chicken'

3. Consonants

Muak Sa-aak has 21 distinctive consonants. They include oral and nasal stops at the bilabial, alveolar, pre-palatal, and velar points of articulation, plus a glottal stop. Aspiration and voicing are distinctive for the stops, although the voiced stops are not common and mostly occur in borrowed words. Also there are no voiced stops at the pre-palatal or velar points of articulation. This shows that Muak Sa-aak underwent a sound shift comparable to the Germanic shift observed by Svantesson (1988, 1991) for U and Hu.

Hall, Elizabeth. 2013. A phonological description of Muak Sa-aak.
Mon-Khmer Studies. 42:26-39

3.1 Initials

As shown in Table 1, Muak Sa-aak has oral and nasal stops at four points of articulation. The alveolar stop is fronted [t̪].

Table 1. Muak Sa-aak initial consonant phonemes

	Labial	Alveolar	Pre-palatal	Post-palatal
Stop	b p pʰ	d t tʰ	c cʰ	k kʰ
Nasal	m	n	ɲ	ŋ
Fricative	f	s		h
Approximant	w	l, r	j	

3.1.1 VOT of aspirated, tenuis, and voiced plosives

To illustrate the aspirated, tenuis, and voiced contrasts, VOT measures are given in Table 2 below. The voice onset timing (VOT) was measured for initial plosives in six items, the set of *bɛt²* 'fish hook', *pɛ²* 'goat', *pʰe²* '2nd sg pl' and the set of *dɛt²* 'flatten', *tɛk²* 'small', *tʰuːm³* 'love' (see Table 1). Negative VOT for /b, d/ reflects clear voicing. Initial tenuis plosives have very short release noise, with VOTs as short as 0.001s for /p/, or 0.003s for /t/ (because these are initial plosives, VOT during closure cannot be measured). VOT for aspirated plosives, however, varies between 0.061-0.096ms for /pʰ/ and 0.037-0.051s for /tʰ/ and is significantly longer than for the non-aspirated stops.

Table 2. VOT for voiced, tenuis, and aspirated voiceless stops (seconds).

Ref.	Item	Gloss	Speaker 1		Speaker 2	
804	*bɛt²*	fishhook	-0.094	-0.058	-0.072	-0.054
965	*pɛ²*	goat	0.001	0.005	0.004	0.006
1652	*pʰe²*	you (pl)	0.065	0.061	0.096	0.074
1526	*dɛt²*	flatten	-0.064	-0.083	-0.121	-0.138
1509	*tɛk²*	small	0.006	0.007	0.003	0.004
284	*tʰuːm³*	love	0.039	0.038	0.051	0.037

3.1.2 Pre-palatals

Oral and nasal pre-palatal stops /c, cʰ, ɲ/ are realized as alveolopalatals [t̠ɕ, t̠ɕʰ, ɳ̠ʑ][3]. The aspirated plosive occurs in free inter- and intra-speaker variation with a homorganic fricative [ɕ~t̠ɕʰ], as in [ɕak²] versus [t̠ɕʰak²] for 'seed'. In syllable-final position, the alveolopalatal oral and nasal stops /c/ and /ɲ/ are accompanied by a short [i]-like transition, like in other Mon-Khmer languages. Evidence for the interpretation as plosives rather than affricates is provided through VOT measures for the three places of lingual articulation: alveolar, alveolopalatal, and velar (Table 3). The VOT for /c/ and /cʰ/ is in line with that of the other stops. For speaker 1, aspiration of the alveolopalatal stop with 0.070-0.090s is even shorter than for alveolar with 0.076-0.095s. For tenuis stops, the alveolopalatal place of articulation with 0.023-0.024s VOT shows the longest release noise for this speaker, compared to 0.010-0.011s for alveolar and 0.010-0.011s for velar stops. For speaker 2, VOT for alveolar and velar tenuis stops with 0.009-0.011s and 0.006-0.009s is only slightly shorter than for alveolopalatal tenuis stops with 0.011-0.015s.

[3] Since alveolopalatal consonants appear to be a Mainland-Southeast Asian areal feature, Clark represents them with single letters that Chinese researchers use in their description of Hmong-Mien languages (2008). These symbols are chosen for the phonetic transcriptions of the alveolopalatal consonants in this study as they accurately describe the phonetic quality of these sounds; however, the usual palatal symbols will be used for the phonemes since they are widely used in the literature.

The longer release noise of alveolopalatal tenuis stops is caused by the large area of impact of the tongue blade with the alveolopalatal region, and the close approximation of the back of the tongue to the hard palate during closure. The resulting friction upon release is clearly heard with tenuis stops, whereas aspiration superimposes the alveolopalatal release noise. This explains why VOT of initial aspirated alveolopalatal stops does not necessarily differ from VOT measures of initial alveolar or velar stops.

To further support the interpretation of alveolopalatal obstruents as plosives rather than affricates, the length of the sibilant /s/ is also included in Table 3. VOT for the /s/ is considerably longer than any of the stops, including that of alveolopalatal /c/ and /cʰ/, which is evidence for considering these to be stops rather than affricates. For comparison, Thurgood and Demenko (2003), studying Polish affricates, found a duration of about 0.110s for the alveolopalatal affricate, with stop closure of only about 0.050s, leaving the frication duration at about 0.060s. The unaspirated Muak Sa-aak equivalent has a release noise of only 0.011-0.024s.

Table 3. Release noise duration for alveolar, alveolopalatal, and velar stops, and the alveolar fricative (seconds).

Ref.	Item	Gloss	Speaker 1		Speaker 2	
868	*tap²*	army	0.011	0.010	0.011	0.009
17	*tʰa:k¹*	tongue	0.095	0.076	0.068	0.091
1434	*cak²*	pull	0.023	0.024	0.011	0.015
1202	*cʰak²*	seed	0.090	0.070	0.088	0.091
322	*kak²*	obstruct	0.010	0.011	0.009	0.006
206	*kʰat²*	ill	0.082	0.068	0.109	0.084
559	*sak*	tattoo	0.138	0.163	0.080	0.056

For further illustration, spectrograms for initial lingual plosives are provided in Figures 3 and 4, showing similar amplitudes and durations for release bursts and aspiration in all three lingual stops.

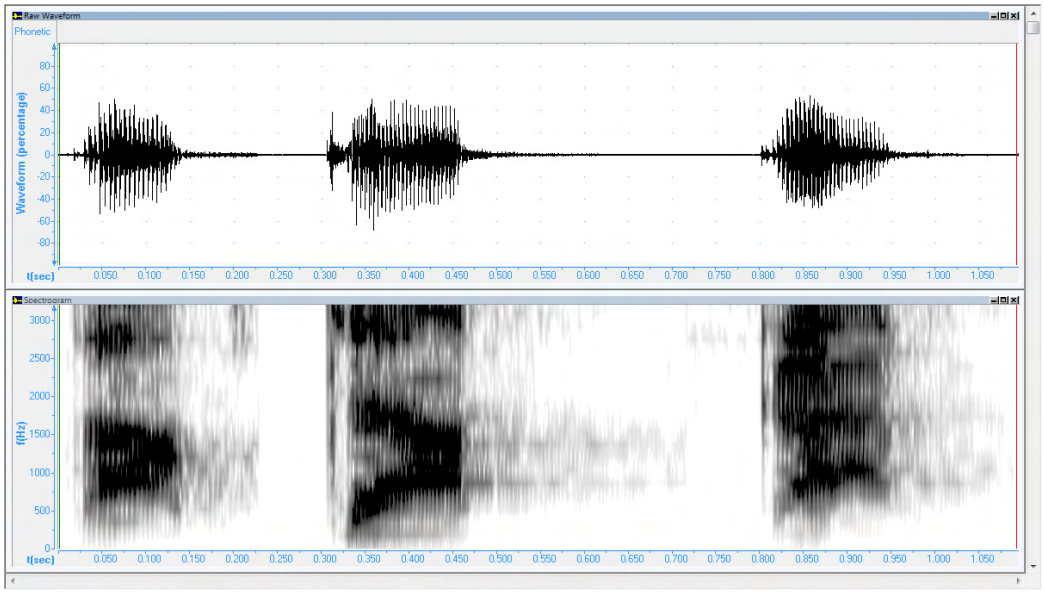


Figure 3. Spectrograms for tenuis stops in *tap²* ‘army’, *cak²* ‘pull’, and *kak²* ‘obstruct’.

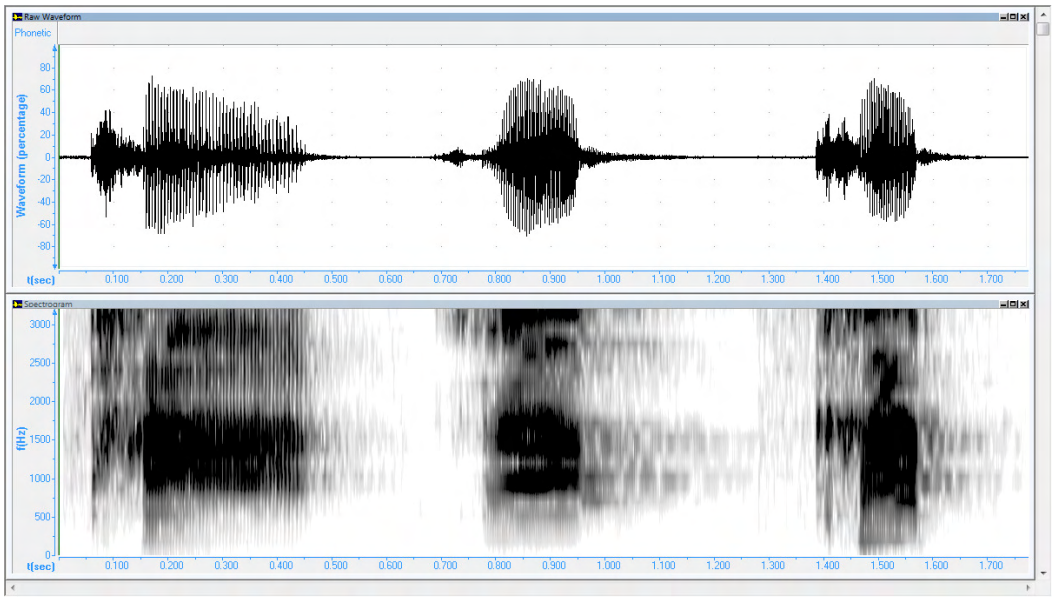


Figure 4. Spectrograms for aspirated stops in *tʰaːkˡ* 'tongue', *cʰakˀ* 'seed', and *kʰatˀ* 'ill'.

3.1.3 Fricatives

There are three voiceless fricatives: labiodental /f/; alveolar sibilant /s/; and glottal /h/. There are only a few occurrences of the labiodental fricative /f/ (23 items in the wordlist, of which nine are borrowed), and some speakers seem to realize it as the frequently occurring aspirated stop [pʰ] instead (the phoneme /pʰ/ is found in 116 items). For example, when referring to the name of their own village, sometimes a speaker would say /faːjˡ/ and other times a speaker would say /pʰaːjˡ/. It is possible that there is a change ongoing caused by Tai Lue borrowings, which do not have the /f/. Analysis of this in more speakers with more data might perhaps resolve this question. The phoneme /h/ occurs in only 31 of 1,658 words in the data, most of them Tai Lue loanwords, all of them monosyllabic, but it does contrast with the sibilant /s/.

3.1.4 Approximants

Muak Sa-aak has four approximants /w, l, r, j/. The labial-velar approximant /w/ is in free variation with a delabialized approximant [β̞] in syllable-initial position, as in [s.wɛŋːˀ] and [s.β̞ɛŋːˀ] for 'flea'. However, in monosyllabic words, [w] commonly occurs before open vowels, particularly the vowel /a/. As the second consonant of a cluster, or as a final consonant, this phoneme is realized as [w]. Word initially, the central alveolar approximant /r/ is in free variation with an alveolar trill. When it is the second consonant of a cluster, it is usually pronounced as an approximant. This phoneme does not occur syllable-finally; /l/, in contrast, occurs syllable-finally but not in clusters, although both occur syllable-initially. The palatal approximant /j/ has palatal and alveolopalatal fricative allophones if preceded by reduced syllables with aspirated onsets. If the reduced syllable vowel is maintained in this environment, the approximant is realized as a voiced fricative [j̟]. If the vowel is entirely omitted, the approximant merges with the aspiration of the minor syllable onset to voiceless [ç, ɕ]. This can be seen in a word pronounced twice by the same speaker: [pʰjaŋˀ]~[p.ɕaŋˀ] for /pʰ.jaŋˀ/ 'fat'. This is in agreement with Svantesson's observations on the palatal approximant in U (1988).

3.1.5 Glottal stop

The glottal stop occurs only phonetically as a predictable vowel onset which may be dropped in continuous speech. In syllable-final position, it is not a consonant, but a suprasegmental feature linked to tone. It occurs only with the constricted Tone 2 in smooth syllables ending in vowels, nasals, or approximants, where consonant clusters are not permitted.

3.2 Consonant clusters

A limited number of consonant clusters are permitted. There are no syllable-final consonant clusters; all are syllable-initial. Only voiceless bilabial and velar stops take the position of the first consonant in a cluster and they can only be followed by /w/ and /r/. The clusters found in the data are given in Table 4.

Table 4. Syllable-initial consonant clusters

	w	**r**
p	pw	pr
pʰ	pʰw	pʰr
k	kw	kr
kʰ	kʰw	kʰr

3.3 Final consonants

The following consonants occur in the syllable-final position: /p, t, c, k, m, n, ɲ, ŋ, l, j, w/.

These finals are the voiceless unaspirated stops, the nasals, and the approximants except for /r/, which does not occur in the syllable-final position. The final stops are unreleased. Final consonants are summarized in Table 5.

Table 5. Final consonants

	Bilabial	Alveolar	Pre-palatal	Velar
Stop	p	t	c	k
Nasal	m	n	ɲ	ŋ
Approximant	w	l	j	

One curious phenomenon regarding finals is found in Muak Sa-aak, namely the occurrence of lengthened final consonants after short vowels. Final sonorants - nasals and the final approximants /l, w, j/- are usually shorter after long vowels, and longer if preceded by short vowels, so that the overall syllable length appears equal (see examples 11-18). This is most easily heard in utterance-final syllables, where there seems to be a preference for a certain syllable length, accomplished through the lengthening of final sonorants if paired with short vowels.

(11) *pʰɤl³* [pʰɤlː³] ‘fly’

(12) *pʰ.juːl³* [pʰ.juːl³] ‘wing’

(13) *jam³* [jamː³] ‘die’

(14) *jaːm³* [jaːm³] ‘cry, weep’

(15) *cʰim³* [cʰɪmː³] ‘bird’

(16) *cʰɯːŋ³* [cʰɯːŋ³] ‘cloth

(17) *kʰaːj³* [kʰaːi³] ‘eat’

(18) *kʰaj³* [kʰaiː³] ‘fat (cow)’

Average lengths for sequences of long vowels followed by short sonorants, and short vowels followed by long sonorants (four tokens each word), are given in Table 6.

Table 6. Comparison of durations of vowel, sonorant, and rime (seconds).

Reference	Item	Gloss	Vowel	Sonorant	Rime
823	*kʰiŋ³*	expensive	0.253	0.285	0.538
1231a	*kʰi:ŋ³*	ginger	0.437	0.212	0.649
134	*ŋaŋ³*	hear	0.182	0.387	0.569
1565	*ŋa:ŋ³*	sweet	0.362	0.198	0.560
249	*jam³*	die	0.139	0.334	0.473
272	*ja:m³*	cry, weep	0.394	0.190	0.584
1267	*kʰum³*	pit	0.244	0.341	0.585
740	*kʰu:ŋ³*	dig	0.435	0.253	0.688

The overall length of the rime is usually slightly longer for syllables with long vowels than for those with short vowels; however, the overall rime duration difference is smaller than the vowel duration difference. For example, the entire average rime length of *kʰiŋ³* and *kʰi:ŋ³* differs by 0.111s, but the length of the vowels differs by 0.184s. For *kʰiŋ³* vs. *kʰiŋ³*, the rime duration differs by only 0.009s whereas the vowel duration difference is 0.180s, similar to the one in the first example. With the exception of two tokens in the data, the sonorant following a short vowel is longer than the sonorant following a long vowel. Averaging these results in the following:

Short vowel 0.204s + Long sonorant 0.337s = 0.541s

Long vowel 0.407s + Short sonorant 0.213s = 0.620s

As can be seen in these averages, short vowels have about half the length of long vowels, and short final sonorants are only about two thirds of long final sonorant duration.

An explanation for this phenomenon might be available from Thai. Brown sees vowel length in Thai as "more a function of where the final consonant begins than where the vowel ends" (1979: 12). He uses a two-fold classification of Thai tones, the one-part tones: falling, low, and high dead in closed syllables, and the two-part tones: rising, mid, and high live in open syllables. After a short vowel, the final consonant begins earlier, in the first part, or head, of the tone; if a vowel is long, the final consonant begins in the second part, the tail (1979).

Rungpat Roengpitya (2002) similarly found that vowel quality and length of final nasal consonants in Thai are secondary markers used to distinguish between short and long vowels. In particular, short vowels have longer nasal finals than long vowels, and a word with a long nasal final was more likely to be identified by the listener as having a short vowel (2002).

Although this regards only nasals, the same phenomenon could be at work with Muak Sa-aak final sonorants, especially since it has long been in direct contact with a Tai language. The question of whether final sonorant lengthening is an Angkuic feature or whether its development is motivated by language contact can only be answered through research on further Angkuic or Palaungic languages.

4. Vowels

Muak Sa-aak has nine monophthongs with contrastive length except for the open-front and open-back vowels. In addition, there are two diphthongs, /ia/ and /ua/. The complete inventory of 18 vowels is shown in Table 7.

Table 7. Muak Sa-aak vowel phonemes

	Front	Back unrounded	Back rounded
Close	i iː	ɯ ɯː	u uː
Close-mid	e eː	ɤ ɤː	o oː
Open	ɛ	a aː	ɔ
Diphthongs	ia	ua	

4.1. Monophthongs

The close unrounded front vowel occurs both as long and short vowels /i, iː/ in Muak Sa-aak. The short vowel has the near-close allophone [ɪ], occuring before final nasals or the lateral approximant. Before all other final consonants and in open syllables, it is realized as the close vowel [i]. The short close rounded back vowel /u/ has two allophones in free variation, [u] and [ʊ]. However, in the data, the near-close allophone [ʊ] usually occurs before final nasals, with checked Tone 2, or falling Tone 3. The allophone [u] is more common, and may occur in all environments (including those where [ʊ] occurs). The sound [ɛː] occurs in free variation with the frequently occurring long mid vowel /eː/ in open syllables. There is a clear contrast between the short vowels /ɛ/ and /e/ in all environments in which they occur. The long [ɔː] and [ɛː], however, appear to have shifted to the diphthongs /ia/ and /ua/ as described below in all other syllable types except for the open syllable.

4.2 Vowel length

Muak Sa-aak does display distinctive vowel length. When trying to describe the difference in some words, one speaker used the terms "heavy" and "light" to refer to syllables with short and long vowels (short being "heavy" and long being "light").

Vowel lengths were measured for minimal pairs or near minimal pairs of [a, i, u], four tokens each: *kʰiŋ³* 'expensive' and *kʰiːŋ³* 'ginger', *ŋaːŋ³* 'sweet' and *ŋaŋ³* 'hear', *jam³* 'die' and *jaːm³* 'cry', *kʰum³* 'pit' and *kʰuːŋ³* 'dig', *kat²* 'burn' and *kaːt¹* 'fasten', and *kut²* 'think' and *kuːk¹* 'stoop'. The average length for short vowels was 0.183s; for long vowels it was 0.393s. Vowel lengths are given in Table 8.

Table 8. Duration of short and long vowels (seconds).

Wordlist reference	Item	Gloss	Vowel durations for each token				Avg.
			Speaker 1		Speaker 2		
823	*kʰiŋ³*	expensive	0.281	0.258	0.258	0.215	0.253
1231a	*kʰiːŋ³*	ginger	0.474	0.470	0.423	0.381	0.437
134	*ŋaŋ³*	hear	0.147	0.177	0.190	0.214	0.182
1565	*ŋaːŋ³*	sweet	0.366	0.376	0.399	0.305	0.362
249	*jam³*	die	0.143	0.128	0.161	0.123	0.139
272	*jaːm³*	cry, weep	0.369	0.390	0.431	0.386	0.394
1267	*kʰum³*	pit	0.280	0.275	0.221	0.200	0.244
740	*kʰuːŋ³*	dig	0.462	0.432	0.472	0.374	0.435
1498a	*kat²*	burn	0.124	0.135	0.131	0.118	0.127
655	*kaːt¹*	fasten	0.356	0.316	0.435	0.459	0.392
252	*kut²*	think	0.168	0.125	0.155	0.165	0.153
172	*kuːk¹*	stoop	0.294	0.389	0.359	0.320	0.341

Waveforms and spectrograms for the pair *jam³* 'die' and *jaːm³* 'cry, weep' are given in Figure 5 to illustrate the difference in vowel length as well as accompanying final sonorant length.

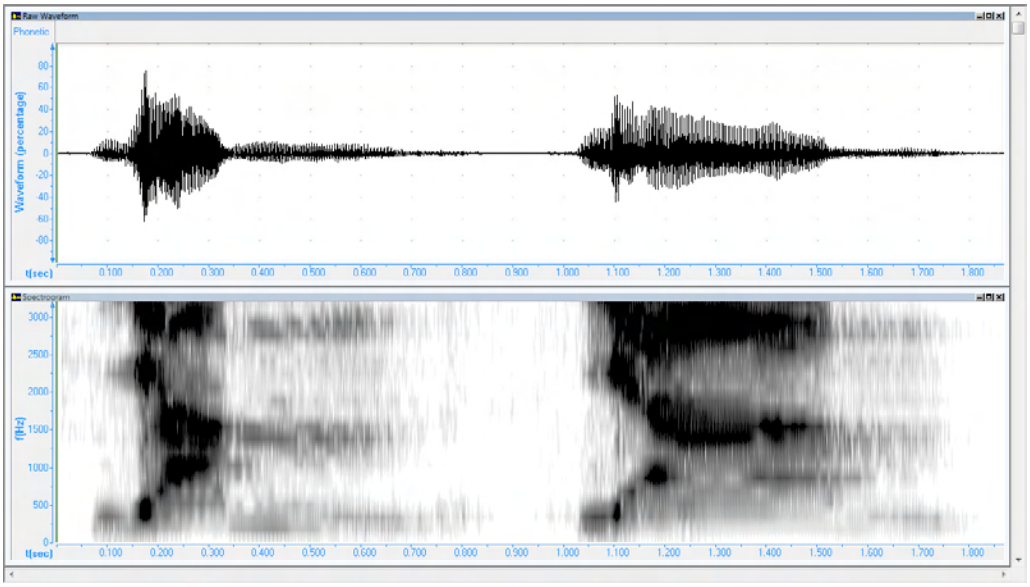


Figure 5. *jam³* ‘die’ followed by *ja:m³* ‘cry, weep’.

4.3 Diphthongs

The diphthongs /ia/ and /ua/ in Wan Fai Muak Sa-aak correspond to the long vowels [ɛ:] and [ɔ:], respectively, in the variety of Muak Sa-aak spoken in Wan Saw [wan¹ sɔʔ²], a Muak Sa-aak variety which does not have these diphthongs. When speakers from Wan Fai try to write their own words in the Tai Lue script, which does not have /ia/ or /ua/, the vowels they choose are ones normally used to write the Tai Lue vowels /ɛ:/ and /ɔ:/. If they then unintentionally read them back as written with long open vowels, they refer to it as “Wan Saw language,” meaning that they are aware of this sound change. This sound change also affects words borrowed from Tai Lue; if the Tai Lue word contains the Tai Lue phoneme /ɛ:/ or /ɔ:/, when borrowed into Wan Fai Muak Sa-aak, it typically becomes /ia/ or /ua/. The diphthongs /ia/ and /ua/ then, if seen as replacements of former [ɔ:] and [ɛ:] in most environments, fill in the empty spaces for the long open front and back vowels (see Table 7, above).

In the Wan Fai Muak Sa-aak variety under study, /ua/ occurs only in closed syllables. The sound [ɔ:] normally occurs in open syllables, although there are a few occurrences in borrowed words with final consonants. This complementary distribution also suggests that the sounds [ua] and [ɔ:] are actually allophones of the phoneme, /ua/. This phoneme could be represented as /ɔ:/ or as /ua/. In this paper, the latter has been chosen because of the limited occurrence of [ɔ:], and for greater symmetry with the diphthong /ia/.

5. Tone

In Muak Sa-aak, there are three distinctive tones: a low Tone 1, a checked Tone 2, and a falling Tone 3. All main syllables have one of these. Presyllables do not display tonal contrast; although speakers labeled them all as Tone 1, the fact that they routinely identified them as the same tone shows that there is no contrast.

Voice quality is not distinctive in Muak Sa-aak but is an accompanying feature of tone. This is most apparent in words with long vowels. Except for the falling Tone 3, these voice qualities are not produced consistently and showed a high degree of both intra- and interspeaker variation.

5.1 Tone 1

The first tone, Tone 1, is a low tone. In one of the recorded speakers, it tends to rise a little. In some words it is realized with stiff voice, a tight, tense phonation type which is more tense than modal voice but less tense than creaky voice[4] such as in *tʰa:k¹* 'tongue', which was often pronounced with even creaky voice, [tʰa̰:k¹]. However, no contrast could be identified based on voice phonation, and this phonation was not even heard consistently with the same word and the same speaker on different occasions. See examples 19-25 for Tone 1 words.

(19) *tʰa:k¹* 'tongue'
(20) *le:k¹* 'pig'
(21) *t.lɤ:¹* 'lizard'
(22) *cʰa:j¹* 'sky'
(23) *li:¹* 'come out, exit'
(24) *rɤ:m¹* 'fade'
(25) *naj¹* 'melt

Tone 1 occurs only in long syllables: syllables with long vowels or diphthongs (regardless of final consonant), or short vowels if followed by a sonorant final.

Presyllables form an exception in that they do have short vowels, and no final consonant; they do not display tonal contrast, but if asked, speakers consistently identify them as having this tone. This suggests that Tone 1 may be the default tone; as Yip describes, tone languages may be considered to have a default, or unmarked, tone, and another tone or tones which are marked (Yip 2002).

The final lateral occurs only rarely and the final pre-palatal nasal does not occur with this tone. Although the other final nasals do occur with this tone, they occur more frequently with the falling Tone 3; those occurring with Tone 1 are commonly borrowed words.

5.2 Tone 2

Tone 2 occurs only on checked syllables, and has two allotones in complementary distribution: high tone on short syllables, and high falling tone on long syllables. The first allotone is a high tone. It occurs only with syllables that have short vowels, with either stop final consonants or a phonetic glottal stop.

The high-falling allotone occurs less frequently, in phonologically open syllables with long vowels, or in closed syllables which have either long or short vowels followed by sonorant finals. This allotone has very creaky voice, and long vowels with this allotone of Tone 2 are slightly shorter than long vowels with either Tone 1 or Tone 3.

With the falling allotone, there are no final stop consonants other than the glottal stop. All syllable types with this allotone can occur with a final glottal stop; however, final glottal stop occurs only with Tone 2 (either allotone), and should be considered a suprasegmental feature of this tone.

There is a complementary distribution between the types of syllables which can occur with the high allotone of Tone 2, and the types which may occur with the falling allotone (Table 9). The two allotones are identified by speakers as being the same tone.

[4] See Ladefoged and Maddieson (1996: 48-50) for further discussion of these voice phonation types.

Table 9. Tone 2 final consonant types distribution by allotone.

	High allotone (short)	High-falling allotone (long)
Open syllable-short vowel	X	
Open syllable-long vowel		X
Stop consonant finals	X	
Nasal consonant finals		X

Examples 26-30 are words with short vowels and the high allotone of Tone 2.

(26) *rɤp²* ‘fishing net’
(27) *tʰi²* ‘arm’
(28) *sut²* ‘smell’
(29) *pa²* ‘have’
(30) *cʰɔ²* ‘dog’

Examples 31-34 take the falling allotone of Tone 2.

(31) *pʰrɤːŋ²* ‘bee’
(32) *cuː¹ ciː²* ‘dung beetle’
(33) *p.niː²* ‘today’
(34) *maːŋ²* ‘destroy, spoil’

All but one of these are probably borrowed words, which is characteristic of words with this allotone of Tone 2.

5.3 Tone 3

Falling Tone 3 has modal voice and is a high falling tone. It does not occur with stop final syllables but only on live syllables. Open syllables do not show a vowel length contrast in this tone; they are all long vowels. The final pre-palatal nasal /ɲ/ occurs only in syllables with this tone, and only with short vowels in the data collected. The majority of words ending with the lateral approximant /l/ also occur with this tone. Examples 35-43 take this tone.

(35) *t.waːj³* ‘tiger’
(36) *kʰaːj³* ‘eat’
(37) *t.poːl³* ‘night’
(38) *kual³* ‘sew’
(39) *ŋaːŋ³* ‘sweet’
(40) *ŋaŋ³* ‘hear’
(41) *p.sɤɲ³* ‘snake’
(42) *feː³* ‘buy’
(43) *tʰuː³* ‘apply, besmear’

Various minimal pairs were found based on tone. Two full sets of minimal pairs, based upon tone plus vowel length, are shown in Tables 10 and 11 below.

Table 10. Tone and vowel contrast in nasal final syllables

	Short vowel		Long vowel	
Tone 1	*raŋ¹*	'rich'	*ra:ŋ¹*	'no one there'
Tone 2	*raŋ²*	'field with no-one working it'	*ra:ŋ²*	'leave/ separate'
Tone 3	*raŋ³*	'shining [of the sun]'	*ra:ŋ³*	'flower'

Table 11. Tone and vowel contrast in open syllables

	Short vowel		Long vowel	
Tone 1	(not possible)		*ci:¹*	'sap'
Tone 2	*ci²*	'do, make'	*cu¹ ci:²* *ci:²*	'dung beetle' 'point'
Tone 3	(not possible)		*cu¹ ci:³*	'make a hole'

Tone is closely related to syllable structure, but the effects of borrowing have had an additional impact and thus complicated tonal features in general, and the picture of tonogenesis in particular. The interaction of tone, syllable structure, and borrowing is complex (Hall 2013) and must be addressed separately.

7. Outlook

The findings on the interaction of Muak Sa-aak vowel and sonorant length and their impact on tonal behavior suggest further studies on syllable weight. The phenomenon of final sonorant lengthening would require further investigation in other Palaungic languages in order to decide whether this is an indigenous Austroasiatic feature or the result of borrowing from a Tai-Kadai contact language. To this author's knowledge, nothing else has been published yet on this language, so there are many additional areas for further study, especially dialect survey and diachronic linguistics, as well as grammar and discourse studies.

Since this study examined only the variety of Muak Sa-aak spoken in one village, insights about the variation between villages in this geographic area would be gained by a dialectal comparison. This appears especially promising in regards to laryngeal settings like pitch and laryngeal constriction since these might possibly be influenced by borrowing. Since the speakers in this study were primarily older speakers, it would also be helpful to examine the speech of younger speakers, to see how the language and its phonology might be changing between generations. Several of the speakers involved in this study expressed concern about the possibility of their children or grandchildren losing their language.

A detailed historical comparative linguistic study would be helpful in better defining the relationships between Muak Sa-aak and the other Angkuic languages. The diphthongs /ia/ and /ua/ were not seen in the small amount of data this author has from another Muak Sa-aak village. These diphthongs are also not seen in Tai Lue, the major language influencing Muak Sa-aak. In Muak Sa-aak as spoken in Wan Fai village, however, the diphthongs are seen both in native words and in borrowed words. Another influence from Tai Lue, which does not have labiodental fricatives, might be the occasional realization of the fricative /f/ as the aspirated plosive /pʰ/.

A study of the grammar of Muak Sa-aak remains to be done; in fact there is still very little written on the grammar of any Angkuic languages. Likewise, this author is not aware of any discourse analyses having been performed in these languages to this date.

References

Brown, J.M. 1979. "Vowel length in Thai", in *Studies in Tai and Mon-Khmer Phonetics and Phonology In Honour of Eugénie J.A. Henderson*, ed. Theraphan. L- Thongkum et al., pp. 10-25. Chulalongkorn University Press.

Conver, Lynn C. 1999. A sketch of the phonology of a Lamet dialect. *Mon-Khmer Studies* 29: 35-56.

Diffloth, Gérard. 1977. Mon-Khmer initial palatals and "substratumized" Austro-Thai. *Mon-Khmer Studies* 6: 39-57.

Diffloth, Gérard. 1980. "The Wa languages", in *Linguistics of the Tibeto-Burman Area* 5 (2): 192 pp.

Diffloth, Gérard. 1982. "Subclassification of Palaungic and notes on 'P'uman'" Paper presented at the 15th Sino-Tibetan Conference, Beijing, August 1982.

Diffloth, Gérard. 1991. "Palaungic vowels in Mon-Khmer perspective", in *Austroasiatic Languages, Essays in honour of H. L. Shorto*, ed. J.H.C.S. Davidson, pp. 13-28. School of Oriental and African Studies, University of London.

Hall, Elizabeth. 2013. "An analysis of Muak Sa-aak tone." Paper presented at the 23rd SEALS Conference, Bangkok, May 2013. http://www.jseals.org/seals23/hall2013analysis.pdf. Retrieved August 14, 2013.

Hopple, Paulette. 2007. Notes re: the visit of Plang from Mong Yawng. ms. (Unpublished)

Ladefoged, Peter and Ian Maddieson. 1996. *The Sounds of the World's Languages*. Malden, Massachussetts: Blackwell Publishing.

Lebar, Frank M, Gerald C. Hickey, and John K. Musgrave. 1964. *Ethnic groups of Mainland Southeast Asia*. New Haven: Human Relations Area Files.

Lewis, M. Paul (ed.). 2009. *Ethnologue: Languages of the World*, Sixteenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com/.

Narumol Charoenma. 1982. The phonologies of a Lampang Lamet and Wiang Papao Lua. *Mon-Khmer Studies* 11: 35-45.

Paulsen, Debbie Lynn. 1992. A phonological reconstruction of Proto-Plang. *Mon-Khmer Studies* 18-19: 160-222.

Rungpat Roengpitya. 2002. A Historical and Perceptual Study of Vowel Length in Thai. *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society*, ed. M. Macken, pp. 353-366. Arizona State University, Program for Southeast Asian Studies.

Scott, J. George and J.P. Hardiman. 1900. *Gazetteer of Upper Burma and the Shan States*, Part 1, Vol. 1. Rangoon: Superintendent, Government printing.

Svantesson, Jan-Olof. 1988. "U." *Linguistics of the Tibeto-Burman Area*, vol. 11, no. 1, pp. 64-133.

Svantesson, Jan-Olof. 1991. "Hu - a language with unorthodox tonogenesis", in *Austroasiatic Languages, Essays in honour of H. L. Shorto*, ed. J.H.C.S. Davidson, pp. 67-80. School of Oriental and African Studies, University of London.

Thurgood, Ela, and Grazyna Demenko. 2003. "Phonetic realizations of Polish geminate affricates." *15th International Congress of Phonetic Sciences, Barcelona*, pp 1895-1898.

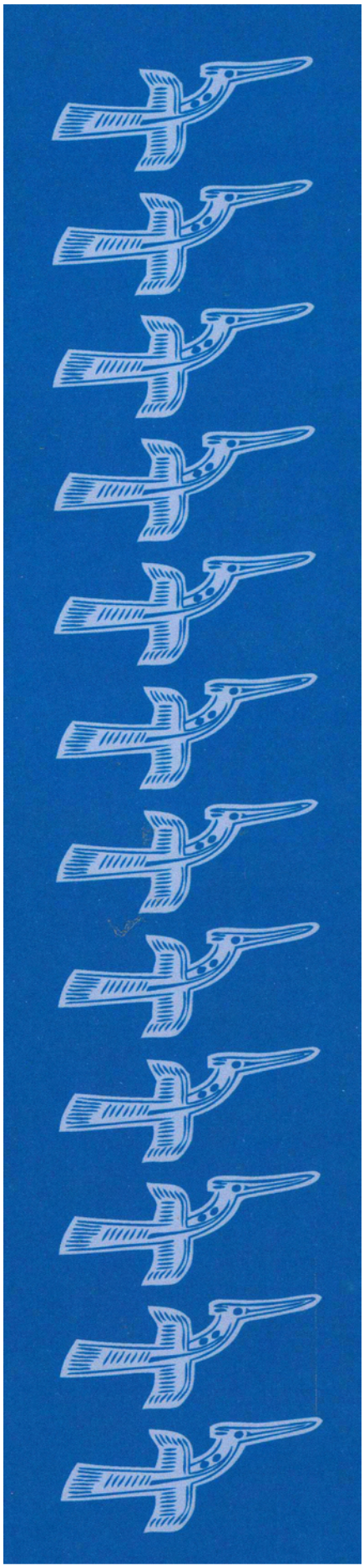
Watkins, Justin. 2002. *The Phonetics of Wa: Experimental phonetics, phonology, orthography and sociolinguistics*. Canberra: Pacific Linguistics.

Wenk, Klaus. 1965. "Drei 'Lăwā'-Vokabularien aus Nordthailand." *Oriens Extremis* 12, 107-127.

Yip, Moira. 2002. *Tone*. Cambridge: Cambridge University Press.

Mon-Khmer Studies

Volume 42



Author: YANIN Sawanakunanon
Title: *Segment timing in certain Austroasiatic languages: implications for typological classification*
Pages: 40-53
Date received: 18/8/2013
Revised text accepted: 29/8/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

Segment timing in certain Austroasiatic languages: implications for typological classification

YANIN Sawanakunanon
Chulalongkorn University

Abstract
This study analyzed segment timing in Mon, Khmer, and Vietnamese. The speech data were segmented into vocalic, consonantal, voiced, and unvoiced intervals. The results showed that the variation of vocalic durations plays an important role in language classification. The different characteristics of vowels in each language led to different timing patterns. Khmer, a restructured language, has vowel length distinctions resulting in the highest variation of vocalic durations. Mon, a register language, follows Khmer with the distinction in phonation types. Vietnamese, a tonal language in which some tones co-occur with phonation, has the lowest variation of vocalic durations. It was noted that suprasegmental features had various levels of influence on segment timing.
Keywords: segment timing, speech rhythm, phonetics
ISO 639-3 language codes: vie, khm, mnw

1. Introduction

Studies investigating segment timing or segment duration have found that many factors affect segment timing. Some of those factors are syllable structure, segment position in the syllable, syllable position in word, phrase, or utterance, stress level, focus, sound environment, speech tempo, articulation process, as well as intrinsic duration of the segment itself (Botinis, Bannert, Fourakis, & Pagoni-Tetlow, 2002; de Jong, 1991, 2004; Greenberg, Carvey, Hitchcock, & Chang, 2003; Suomi, 2005; Warner & Arai, 2001).

There are also studies of segment timing that aim to classify languages according to temporal organization of segments in connected speech. Such studies developed from the study of speech rhythm. Rhythmic units in speech can be determined by the recurrence of stressed syllables, all syllables, or moras and the recurrence of such units are believed to be perceived as approximately equal in duration leading to a rhythmic pattern. It has been suggested that there are three types of speech rhythm: stress-timed, syllable-timed, and mora-timed, based on what units determine the rhythm. The classic examples of stress-timed languages are English and German. The main cited examples of syllable-timed languages are French, Spanish, and Italian (Pike, 1945). Finally, Japanese is a mora-timed language (Laver, 1994).

Acoustic studies, however, fail to support the theory that rhythmic units are perceptually equal in duration (Dauer, 1983; Luangthongkum, 1977; Roach, 1982; Sawanakunanon, 2002; Surinpiboon, 1985; Teeranon, 2000). Dauer (1983) suggested that phonological, phonetic, lexical, and syntactic factors, rather than the speaker's attempt to equalize inter-stress or inter-syllable intervals, may cause rhythmic differences. She further proposed three main differences between stress-timed and syllable-timed languages: the variation and complexity of syllable structure, the presence or absence of vowel reduction, and lexical stress. Stress-timed languages have more types of syllable structures, and those syllable structures are more complex than in syllable-timed languages. In addition, syllable weight plays some role in stress assignment. Heavy syllables tend to be stressed more than light ones.

Vowel reduction is found in stress-timed languages. While vowel reduction is conditioned by phonetic factors in such languages, it seems to be conditioned by phonological environment in languages with syllable-timed rhythm. Moreover, syllable-timed languages do not regularly have reduced variants of vowels in unstressed position. Most stress-timed languages have lexical or word-level stress realized by phonetic characteristics such as high pitch, greater length, loudness, and full vowel quality, which make stressed syllables prominent. It could be argued, therefore, that all syllables tend to be equally prominent in syllable-timed languages. However, some languages have a mixture of characteristics from both rhythmic classes. For example, Catalan, with syllable structures similar to those of Spanish, could be classified as a syllable-timed language, yet it has

YANIN Sawanakunanon. 2013. Segment timing in certain Austroasiatic languages: implications for typological classification. *Mon-Khmer Studies* 42:40-53

vowel reduction. In contrast, Polish, which has a great variety of syllable structures, has no vowel reduction (Nespor, 1990). Although Dauer's proposal may not hold true for every language, it provides alternative ways of explaining factors determining speech rhythm.

Most languages studied in this aspect are European languages and some major eastern languages. Only a few Southeast Asian languages are attested (Grabe and Low, 2002; Romano, Mariano, and Calabro, 2011). Austroasiatic languages, however, have never been analyzed in this fashion. In this study, segment timing of Burmese Mon, Surin Khmer, and Hanoi Vietnamese will be examined. Burmese Mon is a register language in which phonation types of its vowels are phonemically distinctive. Surin Khmer is a restructured language. Its vowel system lost phonation distinction and resulted in phonemic difference in vowel length and various vowel qualities. Unlike the other two languages, Vietnamese is a tonal language.

The three languages have different dominant phonetic and phonological characteristics. In terms of syllable structures, Mon and Khmer are rich in sesquisyllabic words. The stress pattern of light and heavy syllables in Mon and Khmer, thus, could be similar to stress-timed languages and different from Vietnamese which is considered a monosyllabic language. As for vowels, Surin Khmer has vowel length distinction. Vietnamese has one pair of short and long vowel but Mon has none. Suprasegmental features in the three languages are also different. There are six tones in Vietnamese and two of which co-occurred with phonations. Phonation also plays an important role in Mon vowels resulting in two sets of vowels with different phonation types. In this paper, we will see how and whether Mon, Khmer, and Vietnamese, whose phonetic characteristics affect segment duration in different ways, can be classified by their segment timing pattern by following the analyses of the three language classification models explained below in §2.

2. Language classification models

Besides a number of studies in speech rhythm from phoneticians, there are some works in psycholinguistics dedicated to speech rhythm as well. Psycholinguistic studies of speech segmentation reveal infants' ability to determine word boundaries by using rhythmic cues, which are stressed syllables in stress-timed languages, syllables in syllable-timed languages, and moras in mora-timed languages. Adults continue using this ability in second-language acquisition (Mehler, Dommergues, Fraunfelder, & Segui, 1981). Moreover, infants' ability to discriminate languages with different rhythm classes and the ability to group languages with the same type of rhythm suggest that there must be some characteristics in common between languages in the same group which differentiate them from another group (Mehler & Christophe, 1995; Nazzi, Bertoncini, & Mehler, 1998; Ramus & Mehler, 1999). This raises the question as to what those common characteristics are. With the assumption that infants perceive vowels better than consonants because of the higher energy and duration of vowels and that they perceive speech as successions of these high energy sounds (vowels) alternating with noise (consonants), resynthesized speech which replaced all vowels by /a/ and all consonants by /s/ was used in a language discrimination experiment (Ramus & Mehler, 1999). Their results on language discrimination with the use of the resynthesized speech supported the findings of the experiment with natural speech.

Ramus, Nespor, and Mehler (1999) then developed an acoustic model of rhythmic classification. This model incorporates three parameters derived from the duration of vocalic and intervocalic intervals, which are intervals of successive vowels and consonants respectively[1]. These parameters are the proportion of vocalic intervals in the sentence (%V), the standard deviation of the duration of vocalic intervals within each sentence (ΔV), and the standard deviation of the duration of intervocalic intervals within each sentence (ΔC). They found that %V and ΔC show the grouping of languages which supports the theory of three types of speech rhythm. In their study, two languages which had never been classified by speech rhythm were tested. Polish has complex syllable structures, and yet does not have vowel reduction which is claimed to be a characteristic of stress-timed languages. Catalan, on the contrary, has vowel reduction but simple syllable structures. The model groups Polish with English and Dutch while Catalan is grouped with Spanish, Italian, and French. This result suggests that languages in the study are grouped by the variation and complexity of syllable structure, not the existence of vowel reduction.

[1] To illustrate, Ramus et al. (1999) gives an example of the phrase 'next Tuesday on' which can be transcribed as /nɛkstjuzdeiɔn/. The three vocalic intervals from this phrase are the intervals consisting of /ɛ/, /u/, and /eiɔ/. The four intervocalic (or consonantal) intervals are the intervals consisting of /n/, /kstj/, /zd/, and /n/.

However, the use of vocalic and intervocalic intervals raises some questions. How can infants or adult listeners distinguish between a nasal consonant, which is a part of an intervocalic interval, and a nasal vowel, which is a part of a vocalic interval? Should syllabic consonants and glides be included in vocalic or intervocalic intervals? In their study, Galves, Garcia, Duarte, and Galves (2002) found evidence that infants perceive speech signals on the basis of sonority and obstruency. The criterion used to determine sonorant and obstruent sounds in their study is neither articulatory nor phonological but based purely on the acoustic properties of speech. Steiner (2003), using a sonority hierarchy, classified sounds into eight groups: vowel, approximant, syllabic lateral, syllabic nasal, lateral, nasal, fricative, and affricate. The first four groups are included in the vocalic intervals, and the latter four are included in the intervocalic intervals. However, Steiner (2003) suggested that lateral and nasal intervals can classify languages well, and that some classes of consonants might play a more important role than others in language grouping.

Dellwo, Fourcin, and Abberton (2007) took a different approach. They gave an example of the problem in classifying nasal consonants and nasal vowels. They also hypothesized that listeners may be able to distinguish languages based on the difference between voiced and voiceless sounds. Voiced interval (VO), instead of vocalic interval, is used in the parameter %VO, the proportion of voiced interval in the sentence. Voiceless or unvoiced interval (UV), instead of intervocalic interval, is used in the parameter varcoUV, which is the variation coefficient of the standard deviation of unvoiced intervals. Unvoiced intervals are normalized to reduce any effect of speech rate. The results seemed to support the traditional classification. English and German, which are stress-timed languages, are grouped together with high varcoUV values and low %VO. French and Italian, with low varcoUV and high %VO values, are separated from the other two languages. A high varcoUV value can be linked to complex syllable structures, as found in English and German, whereas a low value, as in French and Italian, seems to suggest simple syllable structures.

Not only have there been debates regarding segmentation of vocalic and intervocalic intervals, but also alternative parameters and calculations have been introduced. Low, Grabe, and Nolan (2000) proposed a different calculation of vocalic and intervocalic intervals. In their previous studies (Grabe, Post, & Watson, 1999), English had more vocalic variability than French. They related this finding to vowel quality and explained that English has high variability in vowel durations because it has both full and reduced vowels. French does not have reduced vowels, and that makes the level of vocalic variability lower than that of English. Therefore, they focused on the difference in the variability of vowel duration and computed a Pairwise Variability Index (PVI) which expressed the level of variability in successive measurements. Two versions of PVI were proposed (Grabe & Low, 2002): normalized PVI (nPVI) was used with vocalic intervals, and raw PVI (rPVI) was used with intervocalic intervals. They argued that their PVIs would capture the characteristics of rhythm better than Ramus, et al.'s ΔV and ΔC. Two sets of data of which one had three successive long vowels that followed three successive short vowels, and another which had long and short vowels alternating, would have the same standard deviation of vocalic interval durations although the patterns differed. The results suggested that the vocalic nPVI provided a better separation of languages than the intervocalic rPVI.

The vocalic nPVI values of six languages were also compared by Grabe and Low (2002) with Ramus, et al.'s %V values. English and German, which represent stress-timed languages, have high vocalic nPVI values and low %V values. French and Spanish, representing syllable-timed languages, have low vocalic nPVI values but high %V values. Thus, it seems that these two parameters can reflect a rhythmic characteristic which, in this case, is vowel duration. The conclusion by Grabe and Low (2002) that Thai and Tamil, which have vowel length distinctions, are stress-timed languages because of their high nPVI values are questionable since they also have high %V values which are a characteristic of syllable-timed languages. Therefore, languages which have vowel length distinctions should be carefully examined.

In spite of the varieties of methods used in segmentation and statistical analysis, it can be seen that these studies use timing of segmental intervals to classify languages. They also discuss phonetic and phonological factors shared by groups of languages which make them different from the others. Moreover, this kind of language classification is always compared with the classic rhythm class hypothesis. Whenever unclassified or mixed-rhythm languages are tested, they will be compared with the reference languages, such as English, French, and Japanese, to determine their rhythm class.

3. Method

3.1 Languages, Speakers, and Speech Materials

The three languages analyzed in this study are Burmese Mon, Surin Khmer, and Hanoi Vietnamese. Vietnamese is tonal and rich in monosyllabic words. Mon and Khmer are non-tonal languages and have a great deal of sesquisyllabic words[2]. Moreover, phonation type is contrastive in Mon and Vietnamese. A phonation contrast is found between modal and breathy vowels in Mon and phonation co-occurs with tones in Vietnamese (i.e., a creaky tone and a glottalized tone). Short and long vowels are phonologically different in Khmer. Vietnamese has one pair of short and long vowels. The aforementioned phonetic and phonological characteristics of the three languages can be summarized in Table 1.

Table 1 Phonetic and phonological characteristics of the three languages

Languages	Vowel length	Number of syllables in a word	Phonation contrast	Tonal/ non-tonal
Vietnamese	phonemic (1 pair)	monosyllabic	glottalized and creaky tones	tonal
Mon	non-phonemic	sesquisyllabic	modal and breathy vowels	non-tonal
Khmer	phonemic	sesquisyllabic	-	non-tonal

These three languages will be investigated with the three models of Ramus et al. (1999), Grabe and Low (2002), and Dellwo et al. (2007). The characteristics of the three languages shown in Table 1 have never been the focus of attention before as factors which might contribute to segment timing patterns. Therefore, it is interesting to see whether these characteristics will have some effect on segment timing patterns by using the three language classification models.

The three native speakers of each language ranged in age from 25 to 35 years old. The Vietnamese speakers were from Hanoi and Hai Duong and spoke Hanoi dialect. The three Mon speakers are from Mudon, Myanmar. The Khmer speakers spoke Surin dialect. They are all from Surin Province, Thailand.

Spontaneous speech in stories told by speakers with moderate tempo was sampled at 16 kHz and recorded with a unidirectional microphone directly on a laptop computer hard drive. Approximately 30 seconds of clear speech, not including pauses and hesitations, was selected from each speaker for acoustic analysis.

3.2 Acoustic analysis

The data were segmented and labeled, using the Praat software system (Boersma & Weenink, 2010), into vocalic and consonantal intervals, and voiced and unvoiced intervals. These intervals were identified regardless of syllable and word boundaries. In addition, consonant-vowel and syllable boundaries were also marked for reference. Pauses, as well as syllables preceding and following pauses, were excluded from the analysis. Utterance-final syllables were excluded to avoid lengthening effects. It was also impossible to identify the point where a stop sound ended or began when it occured before and after pauses. Utterance-initial syllables were consequently excluded for consistency. Segmentation was made as accurate as possible despite the fact that there was co-production or coarticulation – that is, overlap in articulatory movements. Particular measurement issues that required careful consideration are discussed below.

Vowels were marked between the points where clear patterns of vowel formants appeared whether the acoustic excitation was voiced or voiceless or both. Other acoustic properties were also used to help identify such points. A vocalic interval was marked between the two points. A consonantal interval was then marked between two vocalic intervals.

For glides, Ramus et al. (1999) included pre-vocalic glides in consonantal intervals and post-vocalic glides in vocalic intervals. Grabe and Low (2002) used formant frequency and amplitude movements to classify glides. They included glides in vocalic intervals unless there were

[2] Sesquisyllabic structure is composed of a minor syllable followed by a major syllable.

observable changes in formant and amplitude of speech signals. In this study, pre-vocalic glides were included in consonantal intervals because constriction in initial position is quite audible. Post-vocalic glides were included in vocalic intervals because there is not enough constriction at the end when the vocal tract is coming to shape 'u' or 'i'. These acoustic criteria for glides then agreed with the measurements of Ramus et al. (1999).

As for voiced and unvoiced intervals, Dellwo et al. (2007) used acoustic cues to locate them. A voiced interval of successive voiced segments, beginning from the onset of a voiced segment to the offset of the next one, was marked across syllable and word boundaries. Similarly, an unvoiced interval was marked from the onset to the offset of an unvoiced segment, or successive unvoiced segments were marked.

A glottal stop closure at the end of a glottalized tone and in the middle of a creaky tone was treated as a consonant. So, it was treated as a part of consonantal or unvoiced intervals. Figure 1 illustrates the segmentation of the four types of intervals in Praat.

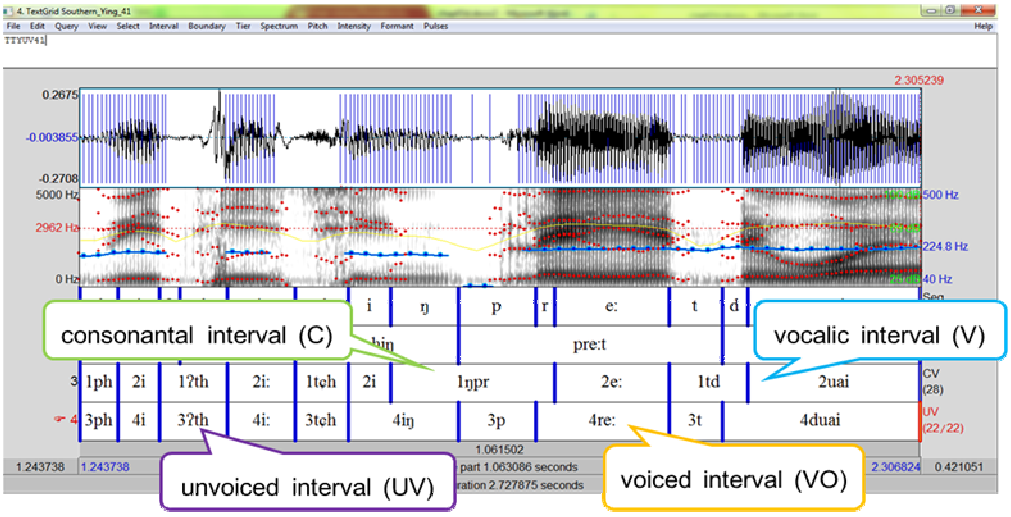


Figure 1: Segmentation of vocalic and consonantal intervals, and voiced and unvoiced intervals

3.3 Language classification parameters

Durations of the four intervals were obtained and expressed as eight parameters in the three models. Duration measurements of vocalic and consonantal intervals were used in the two models proposed by Ramus et al. (1999) and Grabe and Low (2002). The three parameters, which are the proportion of vocalic intervals (%V), the standard deviation of the duration of vocalic intervals (ΔV), and the standard deviation of the duration of consonantal intervals (ΔC), were used in Ramus et al. (1999)'s model. The values of these three parameters were extracted from each utterance. The proportion of vocalic intervals (%V) is the sum of the duration of vocalic intervals divided by the total duration of the utterance. Therefore, %V will show whether the utterance has a proportion of vowels or consonants. The standard deviation of the duration of vocalic intervals (ΔV) and the standard deviation of the duration of consonantal intervals (ΔC) are also calculated per utterance and indicate how the duration of either vocalic or consonantal intervals in each utterance varied. Ramus et al. (1999) found that languages with reduced vowels are likely to have low value of %V and high value of ΔC. Therefore, Mon and Khmer were expected to have such patterns. Although ΔV does not classify languages well compared with %V and ΔC in their paper, it shows the difference between a language with vowel length distinction and languages with no such distinction in my preliminary study. In the current study, Khmer was expected to have high value of ΔV as its short and long vowels are phonemically contrastive. We will also see whether ΔV could capture the difference of phonation types in modal and breathy vowels in Mon.

Grabe and Low (2002) used PVI measurement aiming to show variability of interval duration. However, while the model of Ramus et al. (1999) aims to show variation in each utterance, Grabe and Low (2002) focuses on the difference between duration of two successive intervals. Accordingly, the PVI value represents variability of duration of adjacent intervals, not variability in an utterance. The raw pairwise variability index (rPVI) is used with consonantal

intervals. It shows durational differences between two successive intervals on average. In this paper, it is referred to as rPVI_C for readability and can be computed by using the following formula:

(1) $$rPVI\_C = \left[\sum_{k-1}^{m-1} |d_k - d_{k+1}| / (m-1)\right]$$

In equation (1), 'd' is the duration of the consonantal interval, 'k' is consonantal interval k, $|d_k - d_{k+1}|$ is the absolute value of the durational difference between the preceding and the following consonantal intervals, and 'm' is the number of consonantal intervals in the utterance. The value of rPVI_C is the sum of the durational differences between two successive intervals in each utterance divided by '$m-1$'. In this study, Khmer was expected to have high rPVI_C as its initial consonant cluster is more complex than the clusters of Mon and Vietnamese.

The normalized pairwise variability index (nPVI) which is used with vocalic intervals is referred to nPVI_V in this paper and is calculated by the following formula:

(2) $$nPVI_V = 100 \times \left[\sum_{k-1}^{m-1} \left|\frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2}\right| / (m-1)\right]$$

In equation 2, 'd' is the duration of the vocalic interval, 'k' is vocalic interval k, $|d_k - d_{k+1}|$ is the absolute value of the durational difference between the preceding and the following vocalic intervals, and 'm' is the number of vocalic intervals in the utterance. $(d_k + d_{k+1})/2$ is the average duration of the preceding and the following vocalic intervals and is used to normalized the durational difference between the two vocalic intervals. The value of nPVI_V is the sum of the normalized durational differences between two successive intervals in each utterance divided by '$m-1$' and multiplied by 100.

A high value of PVI shows that there is a great variation between two successive intervals. According to Grabe and Low (2002)'s study, stress-timed languages are likely to have higher value of nPVI_V than that of syllable-timed languages as a result of the durational difference between vowels in stressed and unstressed syllables. In the current study, Mon and Khmer were expected to have higher value of nPVI_V because they have a great number of sesquisyllabic words compared with Vietnamese, which is a monosyllabic language.

Dellwo et al. (2007) proposed sound segmentation into voiced (VO) and unvoiced (UV) intervals instead of vocalic and consonantal intervals as in Ramus et al. (1999) and Grabe and Low (2002). As voiced intervals consist of vowels and voiced consonants, the characteristics of vowels and consonants are responsible for duration and proportion of voiced intervals (%VO). The proportion of voiced interval (%VO) is calculated by the duration of voiced interval in an utterance divided by the duration of the utterance and multiplied by 100.

Unvoiced intervals are only composed of voiceless consonants. The more unvoiced segments occur sequentially, the longer the unvoiced intervals. The variation coefficient of the standard deviation of unvoiced intervals (varcoUV) is computed from equation (3):

(3) $$varcoUV = \frac{\Delta UV}{\overline{UV}} \times 100$$

In equation (3), 'ΔUV' refers to the standard deviation of the duration of unvoiced intervals in the utterance and '$\overline{UV}$' is the average duration of unvoiced intervals in the utterance. The varcoUV value is calculated by dividing the standard deviation of the duration of unvoiced intervals (ΔUV) by the average duration of unvoiced intervals ($\overline{UV}$) and multiplying by 100. The division by $\overline{UV}$ is an attempt to reduce the effect of different speech rate.

As Low et al. (2000) found that there was no effect of speech rate on consonantal intervals, the same result should also be found in the case of unvoiced intervals. Therefore, the standard deviation of the duration of unvoiced intervals (ΔUV), where the duration of unvoiced intervals was not normalized, was analyzed in this study to compare its result with that of varcoUV where the duration of unvoiced intervals is normalized. It is also found in a preliminary study that ΔUV yielded a similar result to ΔC and provided clearer picture of language classification according to statistical analyses.

The values of the three parameters were extracted from each utterance similar to the model of Ramus et al. (1999). Dellwo et al. (2007) found that stress-timed languages were likely to have lower value of %VO and higher value of varcoUV because of their complex syllable structures. The same pattern was expected to be found in Mon and Khmer which have more complex initial clusters than Vietnamese.

The calculation of the eight parameters was done in Microsoft Excel. The results were statistically tested by ANOVA and followed by Tukey's HSD post-hoc test (Tukey's Honestly Significant Difference Test) to ascertain if there was a statistically significant difference.

4. Results

The results of the eight parameters are illustrated below, beginning with the three parameters from Ramus et al. (1999), followed by the two parameters from Grabe and Low (2002) and the last three ones from Dellwo et al. (2007).

4.1 %V, ΔC, and ΔV

The parameters %V, ΔC and ΔV were analyzed following Ramus et al. (1999). The number of vocalic and consonantal intervals, total duration, the average proportion of the duration of vocalic intervals (%V), the average standard deviation of the duration of vocalic intervals (ΔV), and the average standard deviation of the duration of consonantal intervals (ΔC) across all utterances of each language are presented in Table 2.

Table 2 shows that Mon has the highest value for the proportion of vocalic intervals (%V) and Vietnamese has the lowest. This result refutes the hypothesis that Mon and Khmer would have low values of %V because they have reduced vowels in minor syllables of sesquisyllabic words. The ANOVA test shows a significant difference ($p < .05$) and Tukey's HSD test shows that Mon is significantly different from Vietnamese ($p < .05$). The high value of %V, which means a high proportion of vowel duration, in Mon could be because of its breathy vowels, as duration of breathy vowels is found higher than that of modal vowels in some studies (Blankenship, 2002; Kirk, Ladefoged, & Ladefoged, 1984; Luangthongkum, 1990; Wayland & Jongman, 2003). For Vietnamese, a glottal closure in creaky and glottalized tones was treated as a consonant, as mentioned in §3.2. Vowels occurring with such tones are shorter in duration than vowels occurring with non-phonation tones. These tones, hence, contribute to a lower value of vocalic duration in Vietnamese.

Table 2: Total number of vocalic and consonantal intervals, total duration, the proportion of the duration of vocalic intervals (%V), the standard deviation of the duration of vocalic intervals (ΔV), and the standard deviation of the duration of consonantal intervals (ΔC)

Languages	Vocalic intervals	Consonantal intervals	Total duration (Sec)	%V	ΔC	ΔV
KM1	168	175	36.06	50.97	48.88	63.52
KM2	146	155	32.26	49.21	45.57	69.92
KM3	146	149	34.25	54.03	44.86	68.33
KM	460	479	102.57	51.27	46.38	67.42
MN1	142	146	31.11	57.05	42.94	56.10
MN2	134	135	30.22	59.10	41.64	70.69
MN3	177	181	33.22	46.98	44.70	54.25
MN	453	462	94.55	55.43	42.81	62.02
VN1	203	214	33.53	46.68	35.35	39.35
VN2	189	195	36.12	55.81	42.22	46.05
VN3	168	176	33.08	46.18	47.21	47.63
VN	560	585	103.13	49.31	41.05	43.91

KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

The standard deviation of the duration of consonantal intervals (ΔC) was highest in Khmer followed by Mon and Vietnamese. This result supports the hypothesis that Mon and Khmer have

high values of ΔC because their syllable structures are more complex than those of Vietnamese. However, although the numeric pattern of the result supports the hypothesis, the differences were not statistically significant.

Figure 2: Distribution of languages over the %V and ΔC plane (top) and the %V and ΔV plane (bottom)

%V = proportion of vocalic intervals; ΔV = S.D. of vocalic interval duration;
ΔC = S.D. of consonantal interval duration
KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

As for the standard deviation of the duration of vocalic intervals (ΔV), it is found that Khmer has the greatest variation in vocalic interval duration, as expected with its highest value of ΔV among the three languages. Mon comes in second and Vietnamese has the lowest ΔV value. The high ΔV value in Khmer could be a result of the durational difference between short and long vowels. Moreover, the difference between reduced vowels in minor syllables and full vowels in major syllables in sesquisyllabic words might play some role. Although Mon also has a great number of sesquisyllabic words, its ΔV value is lower than that of Khmer. The durational difference between normal and breathy vowels in Mon also resulted in the lower ΔV value than that of Khmer. The result for ΔV, therefore, suggests that the durational difference between short and long vowels is larger than that between reduced and full vowels, as well as normal and breathy vowels. The ANOVA test found a significant difference ($p < .05$) and the Tukey's HSD test found that Mon and Khmer were significantly different from Vietnamese ($p < .05$).

Ramus et al. (1999) suggested that a graph plotted on the %V and ΔC plane gives the best language classification. However, since the ΔC values in Mon, Khmer and Vietnamese are not significantly different, a graph plotted on %V and ΔV plane consequently provides better classification (see Figure 2).

The average values of each language are shown with dotted circles. From Figure 2, the %V and ΔV graph shown in the bottom displays Mon and Khmer together with higher values of %V and ΔV than those of Vietnamese. Mon and Khmer share phonetic characteristics, resulting in similar vowel timing patterns. They are sesquisyllablic languages. The factor that may play the most important role in this part, thus, could be the durational difference between reduced and full vowels in sesquisyllabic words. Vowel length distinction and phonation type distinction in vowels are also important factors causing high values of %V and ΔV.

4.2 PVI Results

PVI measurement is used by Grabe and Low (2002) to show variability of interval duration. The durations of vocalic and consonantal intervals were used to compute the raw pairwise variability index of consonantal intervals (rPVI_C) and the normalized pairwise variability index of vocalic intervals (nPVI_V). The values of both parameters of all speakers of Mon, Khmer, and Vietnamese are presented in Table 3.

Table 3: rPVI_C and nPVI_V values

Languages	**rPVI_C**	**nPVI_V**
KM1	56.46	64.97
KM2	52.84	78.40
KM3	49.76	67.26
KM	53.01	70.72
MN1	46.83	48.11
MN2	45.65	66.54
MN3	47.26	71.97
MN	46.42	62.14
VN1	41.27	54.59
VN2	45.18	48.25
VN3	47.22	58.03
VN	44.27	53.71

KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

From Table 3, the value of rPVI_C in Khmer is higher than Mon and Vietnamese, as expected. Its more complex initial cluster resulted in a high value of rPVI_C, which represents more variability in two successive consonantal intervals. However, the ANOVA test does not show a statistically significant difference, in contrast to the result for ΔC in the Ramus et al. (1999) model.

The value of nPVI_V can be interpreted in the same way. Khmer, again, has the highest value which suggests that adjacent vocalic intervals in Khmer have greater variability than in the other two languages. The result in this part is as expected and similar to the result of ΔV in the model of Ramus et al. (1999), as Mon has the second highest nPVI_V value and Vietnamese comes last. The ANOVA test shows a significant difference ($p < .05$) and the Tukey's HSD test shows that Khmer and Vietnamese are significantly different. Figure 3 shows the three languages plotted on rPVI_C and nPVI_V plane.

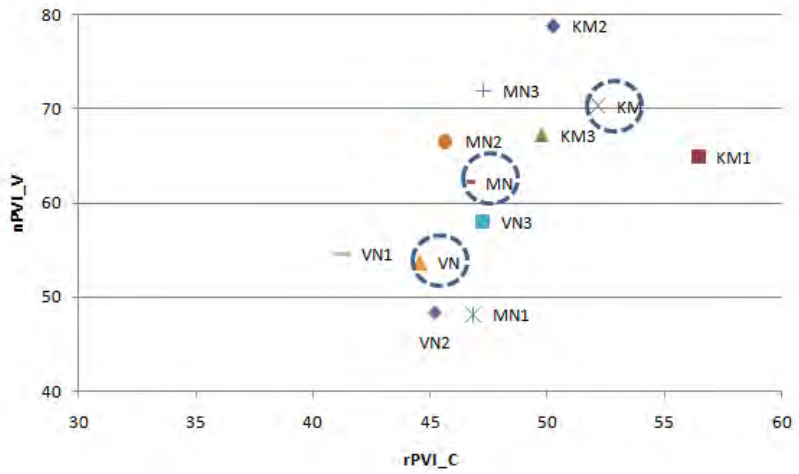


Figure 3: Distribution of languages over the rPVI_C and nPVI_V plane

rPVI_C = raw pairwise variability index in consonantal interval duration;
nPVI_V = normalized pairwise variability index in vocalic interval duration;
KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

The language distribution in Figure 3 shows Khmer with high values for both rPVI_C and nPVI_V. Mon is between Khmer and Vietnamese. The result in this part, thus, supports Grabe and Low's (2002) claim that both parameters reflect the difference between reduced vowels in unstressed syllables and full vowels in stressed syllables as found in Khmer and Mon. Moreover, the values of rPVI_C and nPIV_V of Vietnamese in the current study are similar to those of Romano et al. (2011).

4.3 %VO, varcoUV, and ΔUV

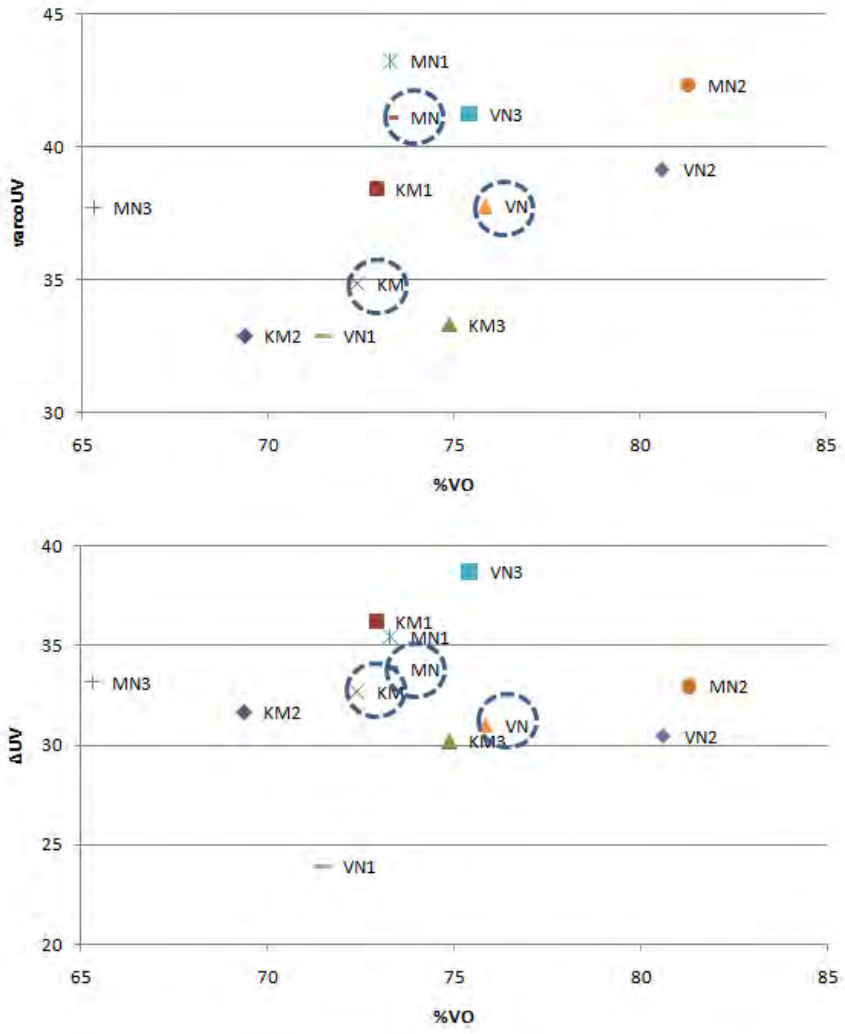
This section describes the analysis of durations of voiced (VO) and unvoiced (UV) intervals following the model of Dellwo et al. (2007). The total number of voiced and unvoiced intervals, the proportion of the duration of voiced intervals (%VO), the variation coefficient of the standard deviation of unvoiced intervals (varcoUV), and the standard variation of unvoiced intervals (ΔUV) are presented in Table 4.

Table 4: Total number of voiced and unvoiced intervals, the proportion of the duration of voiced intervals (%VO), the variation coefficient of the standard deviation of unvoiced intervals (varcoUV), and the standard variation of unvoiced intervals (ΔUV)

Languages	Voiced intervals	Unvoiced intervals	%VO	varcoUV	ΔUV
KM1	113	112	72.91	38.43	36.24
KM2	100	98	69.38	32.88	31.66
KM3	99	95	74.87	33.29	30.21
KM	312	305	72.20	34.75	32.64
MN1	104	102	73.28	43.24	35.44
MN2	81	72	81.29	42.36	32.96
MN3	139	134	65.30	37.72	33.16
MN	324	308	74.79	41.47	33.79
VN1	133	129	71.47	32.89	23.92
VN2	101	97	80.60	39.16	30.45
VN3	90	85	75.40	41.26	38.72
VN	324	311	75.44	37.34	30.41

KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

From Table 4, the %VO values of the three languages can be seen to be not much different. However, it is noticeable that Vietnamese, which has the lowest value of %V as shown in §4.1, has the highest value of %VO. As the difference between the two parameters is that voiced consonants are included in %VO, the higher value of %VO suggests that the Vietnamese data have more voiced consonants than Mon and Khmer. Mon has the highest values of varcoUV and ΔUV. The results of the three parameters for Mon and Khmer, which have more complex syllable structures, show lower values of %VO and higher values of varcoUV than those of Vietnamese, as expected. The graph plotted on %VO and ΔUV plane in Figure 4 shows Mon and Khmer are grouped closer. Nevertheless, the ANOVA tests of the three parameters do not show statistically significant differences across languages.



%VO = proportion of voiced intervals; varcoUV = variation coefficient of the standard deviation of unvoiced interval duration; ΔUV = S.D. of unvoiced interval duration;
KM = Khmer; MN = Mon; VN = Vietnamese; 1 = Speaker 1; 2 = Speaker 2; 3 = Speaker 3

Figure 4: Distribution of languages over the %VO and varcoUV plane (top) and the %VO and ΔUV plane (bottom)

5. Discussion

The eight parameters analyzed in this study were derived from the durations of vocalic-consonantal intervals and voiced-unvoiced intervals. Vocalic and consonantal intervals consist of vowels and consonants respectively. Voiced intervals include not only vowels but also voiced consonants, while unvoiced intervals consist of only voiceless consonants. Although there are some differences between these two groups of intervals, the analyses of the eight parameters were mostly based on the durations of consonantal and vocalic intervals.

According to the results in §4, the parameters acquired from consonantal intervals (ΔC, rPVI_C, varcoUV and ΔUV) do not show statistically significant differences between Mon, Khmer, and Vietnamese. This means that even though Mon and Khmer have more complex initial clusters than Vietnamese, the durational differences are not that large.

On the other hand, there are significant differences between Mon, Khmer, and Vietnamese in the vocalic interval parameters (i.e. %V, ΔV and nPVI_V). Therefore, this finding suggests that the three languages are more different in terms of vowels, especially in their durational variation. The largest differences among the three parameters are found in ΔV and nPVI_V, which show the durational variation of vocalic intervals. The parameter ΔV represents overall variation of vocalic intervals and nPVI_V measures variation of successive vocalic intervals. According to the results of the ANOVA and the Tukey HSD tests shown in §4.1 and §4.2, the value of ΔV in Vietnamese is significantly different from those of Mon and Khmer ($p < .05$) and the value of of nPVI_V in Vietnamese is significantly different from that of Khmer ($p < .05$). Figure 5, with the values of these two parameters plotted, shows that Mon and Khmer are grouped closer with higher values of ΔV and nPVI_V and leave Vietnamese in another corner of the graph with lower values of both parameters.

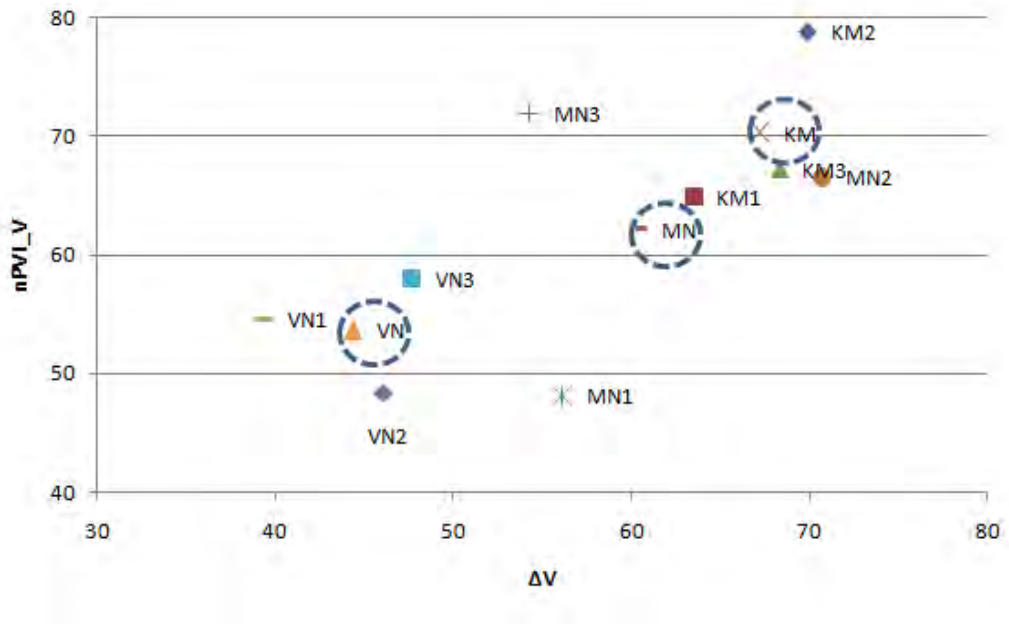


Figure 5: Distribution of languages over the ΔV and nPVI_V plane

As explained in §4, similar phonetic and phonological characteristics should result in similar segment timing patterns. Mon and Khmer are both sesquisyllabic languages and this characteristic differentiates them from Vietnamese, which is a monosyllabic language. The greater difference between the durations of reduced vowels in minor syllables and full vowels in major syllables leads to higher values of ΔV and nPVI_V. The characteristics of vowels themselves also matter. The durational difference between modal and breathy vowels in Mon is another factor that causes high values of both parameters. Moreover, vowel length distinction in Khmer enhances the durational variation of vocalic intervals.

According to the language classification models followed in the current study, values of parameters plotted on graphs can be considered as speech rhythm continuum. Greater variation in the durations of vocalic intervals is a characteristic of stress-timed languages, while less variation is an attribute of syllable-timed languages. Applying this concept to the two parameters plotted in Figure 5, the rhythm continuum would lie from syllable-timed rhythm at the bottom left of the graph to stress-timed rhythm at the top right. Although there are no exact reference points in the graph to determine the region of each type of rhythm, it can be said that Vietnamese, at one end of the continuum, has a characteristic of a syllable-timed language and Khmer, on another end of the continuum, seems to have stress-timed rhythm. As for Mon, it falls in the middle of the continuum but is closer to the stress-timed rhythm end.

6. Conclusion

This study analyzed segment timing in Mon, Khmer, and Vietnamese. The speech data were segmented into vocalic, consonantal, voiced, and unvoiced intervals. The interval durations were then measured and converted into eight parameters (%V, ΔV, ΔC, nPVI_V, rPVI_C, %VO, varcoUV, and ΔUV) following the previous works of Ramus et al. (1999), Grabe and Low (2002), and Dellwo et al. (2007).

The results showed that the variation of vocalic durations plays an important role in language classification. The different characteristics of vowels in each language led to different timing patterns. Khmer, a restructured language, has vowel length distinction resulting in the highest variation of vocalic durations. Mon, a register language, followed Khmer with the distinction in phonation types. Vietnamese, a tonal language in which some tones co-occur with phonation, has the lowest variation of vocalic durations. It was noted that suprasegmental features had various levels of influence on segment timing.

The ΔV and nPVI_V graph shows language classification echoing the statistical analyses. The distribution of languages in the graph resembles a speech rhythm continuum where Khmer is on the stress-timed end, Vietnamese is on the syllable-timed end, and Mon is in the middle close to Khmer.

Acknowledgments

This research was supported by the Royal Golden Jubilee program (RGJ), the Thailand Research Fund (TRF) (PHD/0186/2549). I am deeply grateful to my advisors, Professor Theraphan Luangthongkum and Professor Arthur S. Abramson for their guidance and support. I would also like to thank the reviewers for their invaluable comments. Thanks to all the nine speakers. Without their help, this research project would not have been possible.

References

Blankenship, B. (2002). The timing of nonmodal phonation in vowels. *Journal of Phonetics*, 30(2), 163-191.

Boersma, P., & Weenink, D. (2010). Praat: doing phonetics by computer (Version 5.1.43). Retrieved from http://www.praat.org/

Botinis, A., Bannert, R., Fourakis, M., & Pagoni-Tetlow, S. (2002, April 11-13, 2002). *Crosslinguistic Segmental Durations and Prosodic Typology*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.

de Jong, K. J. (1991). An Articulatory study of consonant-induced vowel duration changes in English. *Phonetica, 48*, 1-17.

de Jong, K. J. (2004). Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics, 32*(4), 493-516.

Dellwo, V., Fourcin, A., & Abberton, E. (2007, August 6-10, 2007). *Rhythmical classification based on voice parameters*. Paper presented at the 16th ICPhs 2007, Saarbrücken, Germany.

Galves, A., Garcia, J., Duarte, D., & Galves, C. (2002, April 11-13, 2002). *Sonority as a basis for rhythmic class discrimination*. Paper presented at the Speech Prosody 2002, Aix-en-Provence, France.

Grabe, E., & Low, E. L. (2002). Durational variability in speech and the rhythm class hypothesis. In C. Gussenhoven & N. Warner (Eds.), *Laboratory Phonology 7* (pp. 515-546). Berlin: Mouton de Gruyter.

Grabe, E., Post, B., & Watson, I. (1999, August 1-7, 1999). *The acquisition of rhythm in English and French*. Paper presented at the 14th International Congress of Phonetic Sciences, San Francisco.

Greenberg, S. M., Carvey, H., Hitchcock, L., & Chang, S. Y. (2003). Temporal properties of spontaneous speech - a syllable-centric perspective. *Journal of Phonetics, 31*, 465-485.

Kirk, P. L., Ladefoged, P., & Ladefoged, J. (1984). *Using a spectrograph for measures of phonation types in a natural language*. UCLA Working Papers in Phonetics 59.

Laver, J. (1994). *Principles of phonetics*. Cambridge: Cambridge University Press.

Low, E. L., Grabe, E., & Nolan, F. (2000). Quantitative characterisations of speech rhythm: Syllable-timing in Singapore English. *Language and Speech*, 43(4), 377-401.

Luangthongkum, T. (1977). *Rhythm in standard Thai*. (Unpublished doctoral dissertation). University of Edinburgh, Edinburgh.

Luangthongkum, T. (1990). The interaction between pitch and phonation type in Mon: Phonetic implications for a theory of tonogenesis. *Mon-Khmer Studies*, 16-17, 11-24.

Mehler, J., & Christophe, A. (1995). Maturation and learning of language during the first year of life. In M. S. Gazzaniga (Ed.), *The Cognitive Neurosciences* (pp. 943-954). MIT Press: Bradford Books.

Mehler, J., Dommergues, J. Y., Fraunfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 298-305.

Nazzi, T., Bertoncini, J., & Mehler, J. (1998). Language Discrimination by Newborns: Toward an Understanding of the Role of Rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.

Nespor, M. (1990). On the rhythm parameter in phonology. In I. M. Roca (Ed.), *Logical issues in language acquisition* (pp. 157-175). Dordrecht: Foris.

Pike, K. L. (1945). *The intonation of American English*. Ann Arbor: The University of Michigan Press.

Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues: A study based on speech resynthesis. *Journal of the Acoustical Society of America*, 105(1), 512-521.

Ramus, F., Nespor, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73, 265-292.

Roach, P. (1982). On the distinction between “stress-timed” and “syllable-timed” languages. In D. Crystal (Ed.), *Linguistic controversies* (pp. 73-79). London: Edward Arnold.

Romano, A., Mairano, P., & Calabro, L. (2011, August 17-21, 2011). *Measures of speech rhythm in East-Asian tonal languages*. Paper presented at the The 17th International Congress of Phonetic Sciences (ICPhS 17), Hong Kong.

Sawanakunanon, Y. (2003). A comparison of rhythm in Thai tracheoesophageal and normal speech. In T. L-Thongkum (Ed.), *The Accent of Alaryngeal Thai Speakers*. Bangkok: Chulalongkorn Printing House. (in Thai)

Steiner, I. (2003). *A refined acoustic analysis of speech rhythm*. Paper presented at the LingColl 2003, Budapest, Hungary.

Suomi, K. (2005). Temporal conspiracies for a tonal end: Segmental durations and accentual f0 movement in a quantity language. *Journal of Phonetics, 33*, 291-309.

Surinpiboon, S. (1985). *The accentual system of polysyllabic words in Thai*. Unpublished master's thesis. Chulalongkorn University, Bangkok. (in Thai)

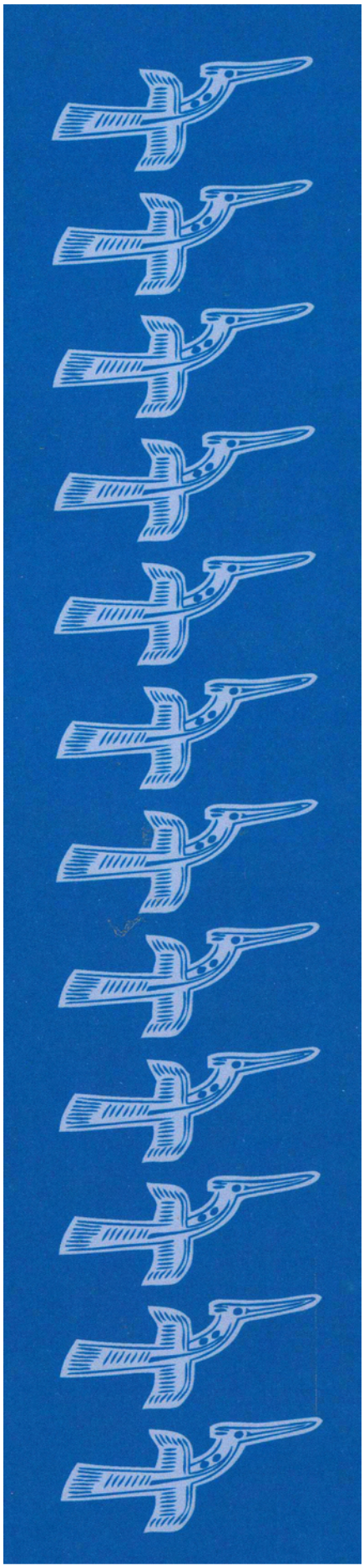
Teeranon, P. (2000). *Rhythmic unit and tonal variation in Thai connected speech*. Unpublished master's thesis. Chulalongkorn University, Bangkok. (in Thai)

Warner, N., & Arai, T. (2001). Japanese mora-timing: A review. *Phonetica, 58*, 1-25.

Wayland, R., & Jongman, A. (2003). Acoustic correlates of breathy and clear vowels: the case of Khmer. *Journal of Phonetics*, 31(2), 181-201.

Mon-Khmer Studies

Volume 42



Author: Narinthorn Sombatnan BEHR
Title: *A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types.*
Pages: 54-80
Date received: 1/5/2013
Revised text accepted: 17/9/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types[1]

Narinthorn Sombatnan BEHR
Department of Linguistics, Faculty of Arts, Chulalongkorn University

Abstract
Previous acoustic studies on a variety of Thai Mon (TM) (Luangthongkum, 1988a; 1990) have found salient pitch patterns, which would seem to indicate a tendency to evolve into a tonal language. However, no acoustic analyses have been undertaken in Burmese varieties of Mon (BM). This research is a synchronic study of vowel systems with an acoustic analysis of vowels in four TM and four BM varieties. A number of vowel phonemes and characteristics were found to be slightly different in TM and BM. H1-A1 and F0 values show a clear distinction between clear vowels and breathy vowels in TM and BM. Conversely, on-gliding and off-gliding vowels were mainly found in BM varieties. Overall, TM and BM are register languages with a pitch pattern. Nevertheless, in the future, TM may become a solely tonal language, while BM seems to tend towards becoming either a tonal language or a restructured one.
Keywords: phonetics, register, vowels
ISO 639-3 language codes: mnw

1. Introduction

The term "register" was first mentioned by Henderson (1952) to describe Cambodian phonology in association with complexes of laryngeal features. First register vowels with a clear voice are more open and have a higher pitch than those of the second register which have a breathy voice. Multidimensional laryngeal features or bundles of laryngeal parameters, resulting from complex laryngeal activity during the phonation process, might be best described as "register complexes". Register complexes comprise several phonetic characteristics, such as phonation type, pitch, loudness, vowel quality and vowel length. Theoretically, one of these parameters could dominate the others due to register distinction; however, more than one feature has been found to be salient due to a possible tendency towards language change. For example, in the case of Kui, a register language, phonation type and pitch play a central role in the language, according to the results of significant H1-H2 and F0 values (Luangthongkum, 1989). Later in 2004, a pitch pattern was found to occur in Kui (Abramson et al., 2004) as well as Khmu (Premsrirat, 2004; Abramson et al., 2007). Perception tests support the idea that Kui and Khmu speakers use pitch as a cue to differentiate word meaning (Abramson, et al., 2004; 2007). These languages, Kui and Khmu, may possibly become tonal languages. Nevertheless, due to the loss of register, complex vowel quality on the co-occurrence of different degrees in glide, height and length can compensate for previous clear vowels and breathy vowels. For example, clear vowels may occur with off-gliding and lower quality in which on-gliding and higher quality may appear with breathy vowels. The vowel system can become restructured with the vowels changing in position and diphthongisation. This can lead to a variety becoming a restructured language in the same way as Khmer (Huffman, 1985), Bru (Phillips, Phillips and Miller, 1976), and Haroi (Mundhenk and Goschnick, 1977).

The Thai Mon (TM) and Burmese Mon (BM) varieties are said to be the same language due to minor differences in their vowel systems (Huffman, 1987-1988). However, previous acoustic studies on Thai Mon (Luangthongkum, 1988a) have revealed significant F0 values which exhibit pitch patterns as a possible salient exponent. This variety is possibly evolving into a tonal language. Meanwhile, Shorto (1966) explains that vowels in the head register are characterised by a clear voice with peripheral quality whereas vowels in the chest register are in breathy voice more centralised. Vowel quality may eventually become dominant in some Mon varieties. Nevertheless,

1 This research is part of my Ph.D. Dissertation "A comparison between the change of vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types" submitted to the Department of Linguistics, Faculty of Arts, Chulalongkorn University.

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

no acoustic analyses have been performed on vowels in BM varieties. It is possible that TM varieties could be distinct from BM varieties, internal and external factors are taken into account.

2. Objective

This research investigated the vowel systems of four Thai Mon varieties in comparison with four Burmese Mon varieties. In addition, single vowels with register contrast in these varieties were acoustically analysed using four parameters: phonation type (the difference of relative amplitude), pitch (F0 values), vowel quality (F1 and F2 values) and vowel length (vowel duration) in order to display the prominent components that could demonstrate whether TM and BM exhibit a tendency to change towards different language types in the future.

3. Data Collection

3.1 Language consultants

The language consultants were male native speakers of Mon born and raised using four different Thai Mon varieties: Ban Kho (TM1), Ban Muang (TM2), Ban Bangkhanmak (TM3), Ban Nong Duu (TM4), and four different Burmese Mon varieties: Mokaneang (BM1), Tancanuʔ (BM2), Sapuʔ (BM3) and Kawbein (BM4). The data was collected in Thailand. Due to the small number of Thai Mon native speakers, the age range of TM speakers was between 50-70 years, while that of the BM speakers was between 30-40 years, due to the fact that they mainly migrated to Thailand looking for work.

3.2 Word lists

In order to analyse the vowel systems, three sets of word lists were used to interview the language consultants. The first one, with 500 vocabulary items, was adapted from the 436 SIL word list (SIL, 2006). The second one, of 300 items, consisted of items selected from Shorto (1962) and Diffloth (1984). The final list consisted of 112 words from Bauer's unpublished dialect word lists.

For acoustic analysis, citation forms consisted of single vowels with register contrast of each variety. The selected monosyllabic and sesquisyllablic words were words used in the speakers' daily life. Syllable structures included open syllables (CV), syllables with glottal finals (CVh), syllables with stop finals (CVT) and syllables with nasal finals (CVN) with mostly voiceless initials. The number of test words varied for each variety according to vowel phonemes: 109 words in TM1, 107 in TM2, 103 in TM3 and TM4, 106 in BM1, 104 in BM2, 103 in BM 3 and 100 in BM4, a total of 835 words.

3.3 Acoustic analysis

3.3.1 Recording

To record the citation form, three native speakers from each variety pronounced each word three times in randomised sequence through a ECM-719 SONY microphone connected to a laptop with 22500 sampling rates. The test tokens totaled 7,515 items.

3.3.2 Acoustic measurement

Vowels in stressed syllables were selected to be measured. To avoid any influence of consonant voicing on the vowels, initials and finals were omitted by visual identification. Each register contrast between clear vowels and breathy vowels were analysed and compared via their phonetic parameters: phonation type, pitch, vowel quality and vowel length, by using "Praat" version 5.2.27. The significant differences of each parameter were statistically analysed by t-test at $p < 0.05$. The four parameters were investigated as follows:

(1) *Relative amplitude of harmonic*: the difference in decibel (dB) between relative amplitude as H1-H2, H1-H3, H2-H4, H1-A1, H1-A2, H1-A3 were measured at five time points of vowel duration: 0% 25% 50% 75% and 100%.

(2) *Fundamental frequency (F0):* F0 values in Hertz (Hz) were measured at five time points of vowel duration as 0% 25% 50% 75% and 100%. Hertz was later converted to semitones using the formula $P_{semitones} = 3.32 \times 12 \times \log_{10} ((F0_{Hz})/base)$ in order to normalise the F0 range across the speakers.

(3) *Formant frequency (F1,F2):* F1 and F2 values were analysed in Hertz at 50% in steady state of vowel.

(4) *Duration:* The onset to offset of vowel was measured in milliseconds (ms).

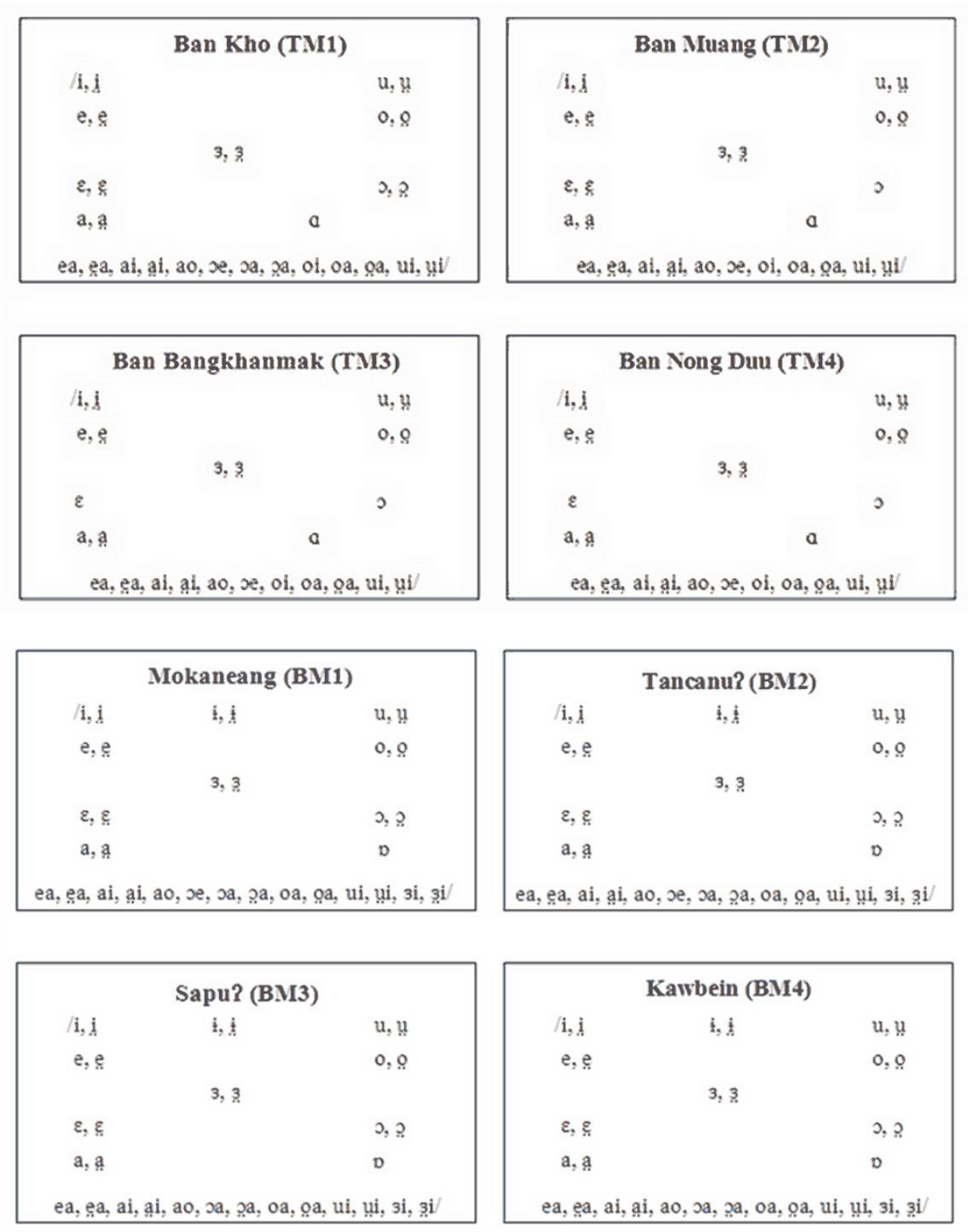


Figure 1: Vowel systems of four Thai Mon and four Burmese Mon varieties

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

4. Results

4.1 Vowel system

Most of vowel inventories in Thai Mon and Burmese Mon varieties occur with register contrast. Based on impressionistic data collection, the phoneme inventories of four Thai Mon and four Burmese Mon varities are illustrated in Figure 1.

From Figure 1, there are 17 monophthongs and 13 diphthongs in TM1 while there are 16 monophthongs and 11 diphthongs in TM2, and 15 monophthongs and 11 diphthongs in TM3 and TM4. Some examples of minimal register contrast are:

/ki/	'bark'	/həkḭ/	'centipede'
/cut/	'put in/ put on'	/cṵt/	'bone'
/sai/	'bee'	/sại/	'thin'

In addition, in Thai Mon varieties, /ɛ̰/ is found in TM1 and TM2 as an example of /həʔɛm/ 'to clear one's throat' vs. /tɛ̰k/ 'tie', and /ɔ̰/, /ɔa/ and /ɔ̰a/ appear in TM1 as examples of /sɔt/ 'fruit' vs. /tɔ̰p/ 'hatch' and /ʔədɔa/ 'inside' vs. /jɔ̰a/ 'sick' respectively but not in TM2, TM3 and TM4.

In BM varieties, there are 19 vowel phonemes in all four varieties for monophthong while 14 diphthongs are found in BM1 and BM2, and 13 diphthongs in BM3 and BM4. The vowel phoneme which occurs only in BM1 and BM2 is /ɔe/ such as/ hətɔe/ 'sand'. Some examples of minimal and analogical register contrast are:

/həmot/	'ant'	/mọt/	'eye'
/kɨt/	'bite'	/həkɨ̰t/	'bedbug'
/ʔədɔa/	'inside'	/jɔ̰a/	'sick'
/tɜiʔ/	'over there'	/sɜ̰ih/	'deep'

In all varieties, monophthongs appear in open and closed syllables. While most diphthongs occur only in open syllables, /ea-ḛa/ in TM and BM varieties, /ɜi-ɜ̰i/, and /ao/ in BM varieties also appear in closed syllables.

Overall, the number of vowel phonemes and phonetic realisation are slightly different between Thai Mon and Burmese Mon varieties. Nevertheless, some vowel phonemes might indicate whether a variety belongs to TM or BM varieties as /ɑ/ and /oi/ which occur only in TM varieties whereas /ɒ/, /ɨ/, /ɨ̰/, /ɜi/ and /ɜ̰i/ appear in BM varieties.

4.2 Acoustic analysis

4.2.1 Relative amplitude

To produce phonation contrast, glottal stricture can vary along the glottal continuum, i.e. breathy voice with more open glottal constriction, creaky voice with tight constriction and modal voice with moderate one (Ladeforged, 1971). Thus, air passing through the glottis is modified differently. The energy difference demonstrates phonation contrast. This can be measured by examining the differences in relative amplitude of a harmonic to that which precedes it, in other words H1 (first harmonic or F0) - H2 (second harmonic), H1-H3 (third harmonic), H2-H4 (fourth harmonic) and the relative amplitude of the first harmonic to that of the strongest peak of formant as H1-A1 (amplitude of F1), H1-A2 (amplitude of F2) and H1-A3 (amplitude of F3). Some of these measurements can successfully distinguish phonation contrast in certain languages. Keating et al. (2010) reveal that H1-H2 distinguishes phonation contrast in Gujarati, White Hmong and Southern Yi, and Esposito (2006) mentions eight other languages. While H1-A1 and H1-A3 differentiate phonation contrast in Gujarati, Jalapa Mazatec and Southern Yi, H1-A3 indicates significant differences in voice quality in Chong (DiCanio, 2009). To distinguish between clear (modal) phonation and breathy phonation, Esposito (2006) shows H1-A2 and H1-A3 to be a successful measurement. In some languages, both H1-H2 and H1-An (H1-A1, H1-A2 and H1-A3) differentiate phonation contrast; for example, H1-H2 and H1-A2 clearly distinguish contrast in

Mazatec (Blankenship, 2002), and H1-H2 and H1-A1 characterize clear (modal) vowels and breathy vowels in Chanthaburi Khmer (Wayland and Jongman, 2003).

The relative amplitude difference of each phonation should differ due to the presence of distinct glottal stricture. In this study, the greater difference occurs in breathy vowels. Figures 2 and Figure 3 show the relative amplitude of H1-H2, H1-H3, H2-H4, H1-A1, H1-A2 and H1-A3 at 0%, 25%, of vowel duration and those of 50%, 75% and 100% are shown in Appendix 1. The results of H1-An (H1-A1, H1-A2 and H1-A3), notably H1-A1, show a significant distinction between clear vowels and breathy vowels ($p < 0.05$) at every time point in most varieties; meanwhile, H1-H2 and H1-H3 show significant differences at 0% and 25% and H2-H4 shows a few significant differences after 50% of vowel duration.

4.2.2 Fundamental frequency

F0 values of clear vowels are higher than those of breathy vowels as seen in Table 1. The results show that F0 values at every time point are significantly different ($p < 0.05$) in TM2, TM3, TM4 and all BM varieties. Meanwhile, the F0 values of TM1 can be distinguished at 0% - 75%, as also seen in Table 1.

The time-normalized average F0 contours of clear vowels and breathy vowels are plotted on a semitone scale in Figure 4. There are similar contours between clear vowels and breathy vowels. The vowel onset rises slightly and then gradually falls to the offset. The slope of clear vowels is higher than that of breathy vowels. However, the slope of breathy vowels in TM1 falls abruptly at the end. A large difference of semitones between clear vowels and breathy vowels occurs in BM and TM1 with a scale range of 3.7-6.1 semitones, while the difference in other TM varieties is 1.4-2.9 semitones. Pitch differences is apparently greater in BM and TM1.

4.2.3 Formant Frequency

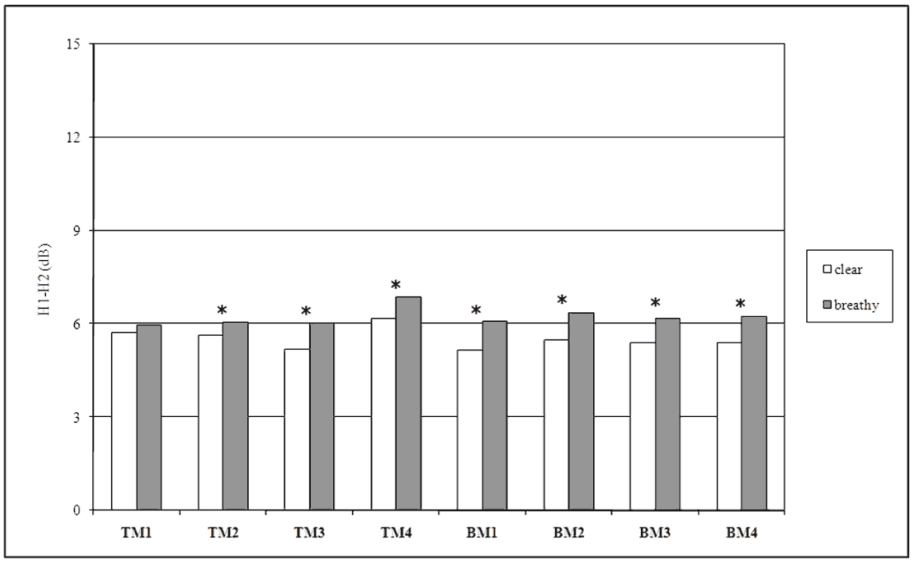
F1 and F2 values of some clear vowels and breathy vowels are significantly different at $p < 0.05$ as shown in Tables 2-3. From Table 2, it can be seen that F1 values in Thai Mon exhibit significant differences for /ɛ-ɛ̰/ in TM1, /o-o̰/ and /u-ṵ/ in TM3, and /a-a̰/ in TM4; meanwhile, a significant difference occurs in the F2 values of /i-ḭ/ and /ɜ-ɜ̰/ in TM1, /e-ḛ/, /o-o̰/ and /u-ṵ/ in TM2, /i-ḭ/ and /ɜ-ɜ̰/ in TM3, /i-ḭ/, /o-o̰/ and /u-ṵ/ in TM4. Both the F1 and F2 values in some pairs of vowels exhibit significant differences as /u-ṵ/ in TM1, /i-ḭ/ in TM2 and /e-ḛ/ in TM3. In BM varieties, F1 values are significantly different for /i-ḭ/, /a-a̰/ and /o-o̰/ in BM1, /i-ḭ/, /e-ḛ/ and /a-a̰/ in BM2, /i-ḭ/ and /ɨ-ɨ̰/ in BM3 and /a-a̰/ in BM4. The difference of F2 values is found significantly in /ɜ-ɜ̰/ of BM1, /o-o̰/ and /u-ṵ/ of BM3, /ɛ-ɛ̰/, /ɜ-ɜ̰/ and /ɨ-ɨ̰/ of BM4 as shown in Table 3. In addition, F1 and F2 values of /ɔ-ɔ̰/ in BM1, BM2, BM4 and /o-o̰/ in BM2 are significantly different.

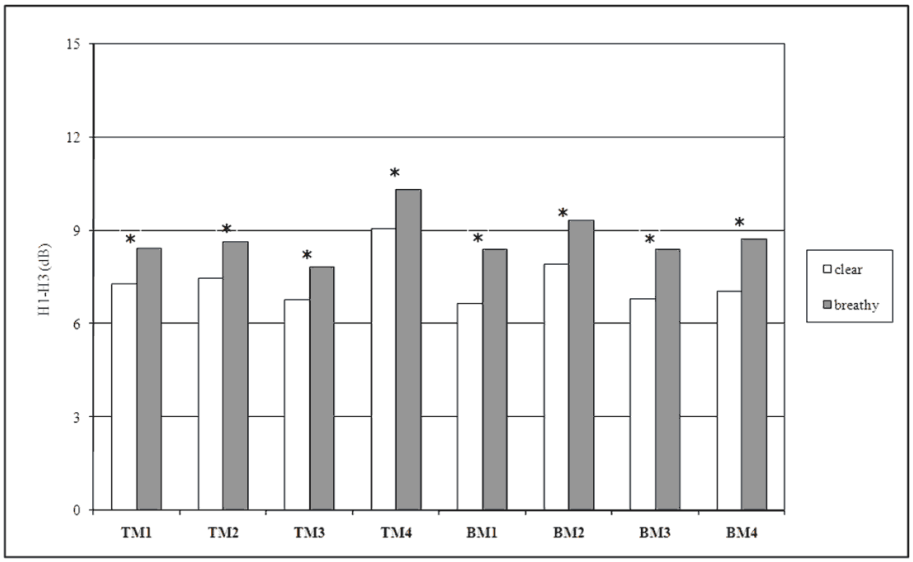
Notwithstanding these values, the difference of F1 and F2 values between clear vowels and breathy vowels are not systematic. No obvious patterns indicate vowel quality difference in vowel space. Neither clear vowels nor breathy vowels are more open or more close, or more front or more back as can be seen in Appendix 2.

4.2.4 Duration

Most breathy vowels are longer than clear vowels, but the duration of breathy vowels in TM2 and TM4 is shorter than that of clear vowels in CVN, as shown in Appendix 3. Overall, the longest duration of breathy vowels in CV is 32-42 ms. and that of clear vowels is 27-38 ms. In other syllable types such as CVh, CVT and CVN, the duration range between clear vowels and breathy vowels is 14-19 ms. and 18-24 ms., 12-19 ms. and 14-23 ms., 17-24 ms. and 15-27 ms. respectively.

This study reveals that a small number of clear vowels and breathy vowels can be significantly distinguished at $p < 0.05$. The significant difference between clear and breathy vowels occurs in CV of TM2, TM3, TM4 and BM3, CVh of BM1 and BM4, and CVN of TM1, TM3 TM4 and BM1. No significant difference is found for CVN syllables except in TM4, whose clear vowels are longer than breathy ones.





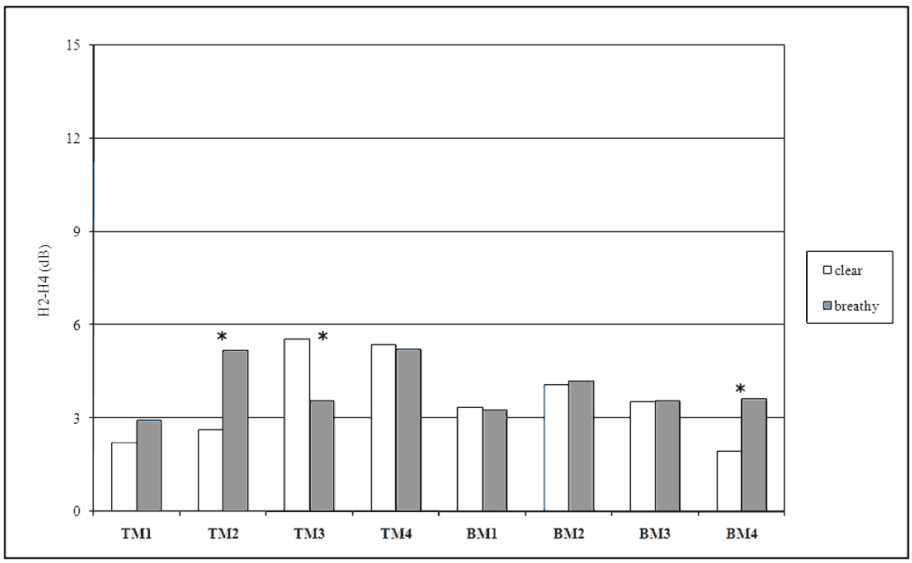


Figure 2.1: Mean values of H1-H2, H1-H3 and H2-H4 (in dB) at 0% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels (An asterisk indicates the values that are significantly different.)

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

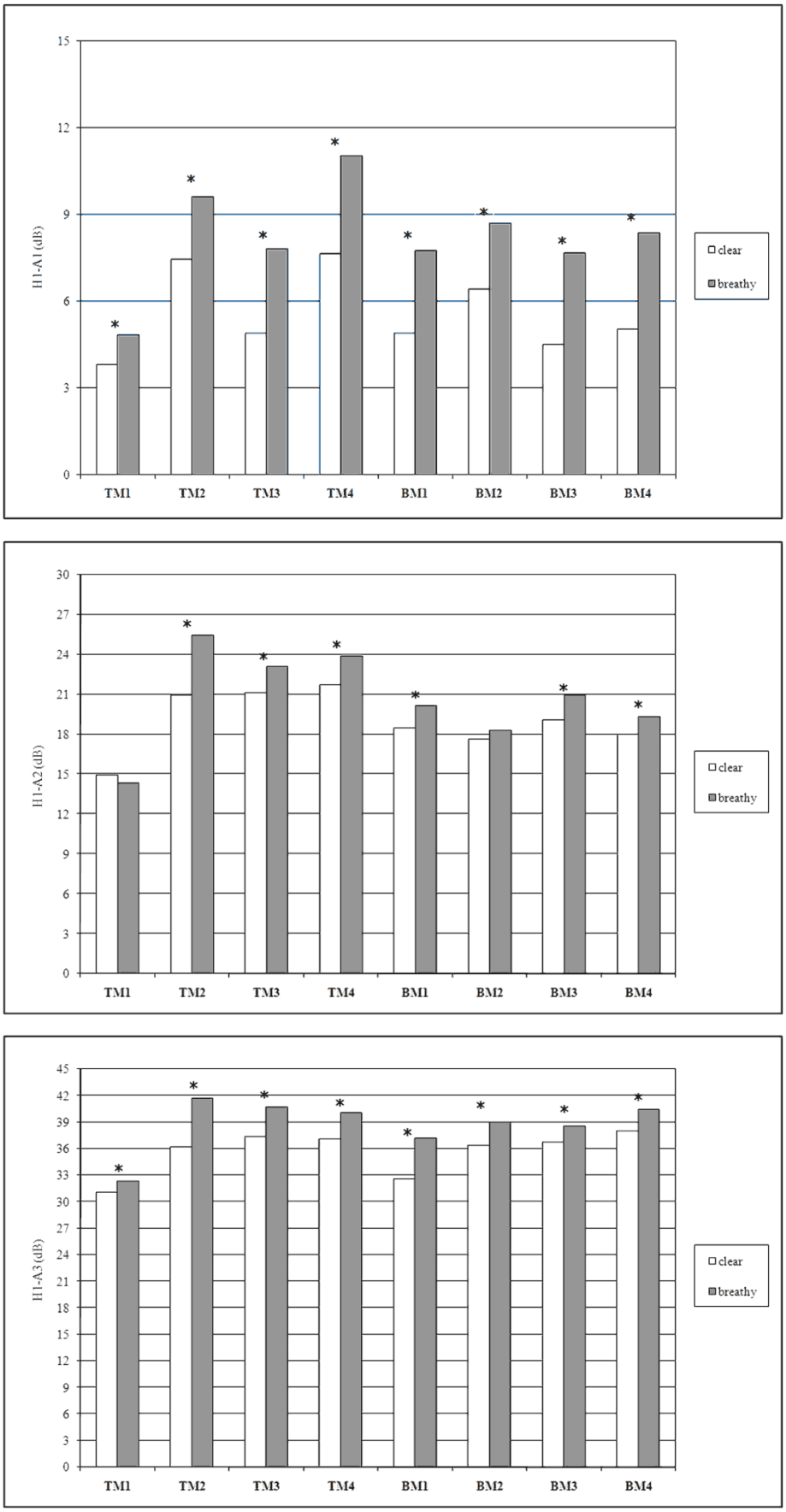


Figure 2.2: Mean values of H1-A1, H1-A2 and H1-A3 (in dB) at 0% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels (An asterisk indicates the values that are significantly different.)

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

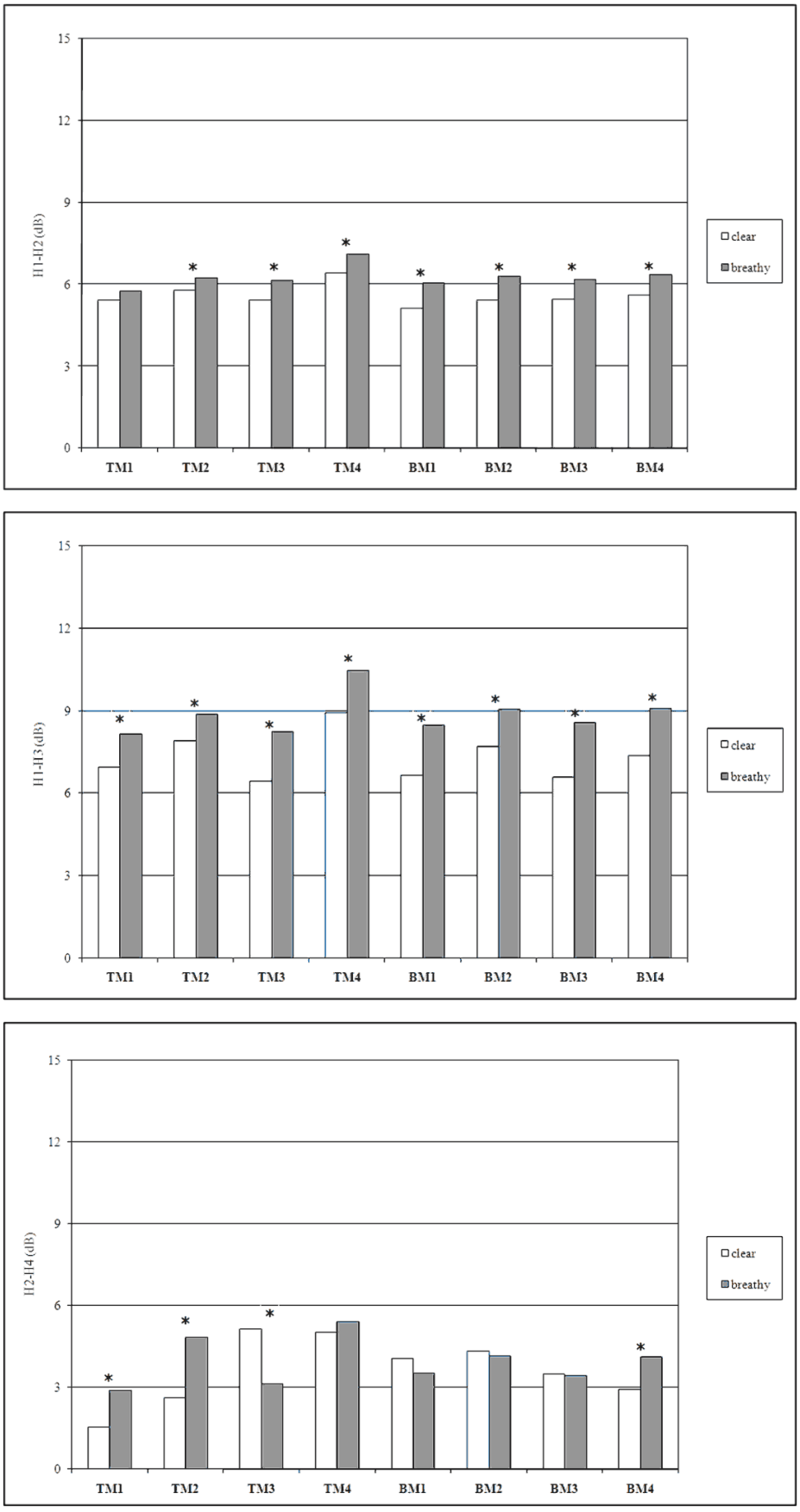


Figure 3.1: Mean values of H1-H2, H1-H3 and H2-H4 (in dB) at 25% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels (An asterisk indicates the values that are significantly different.)

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

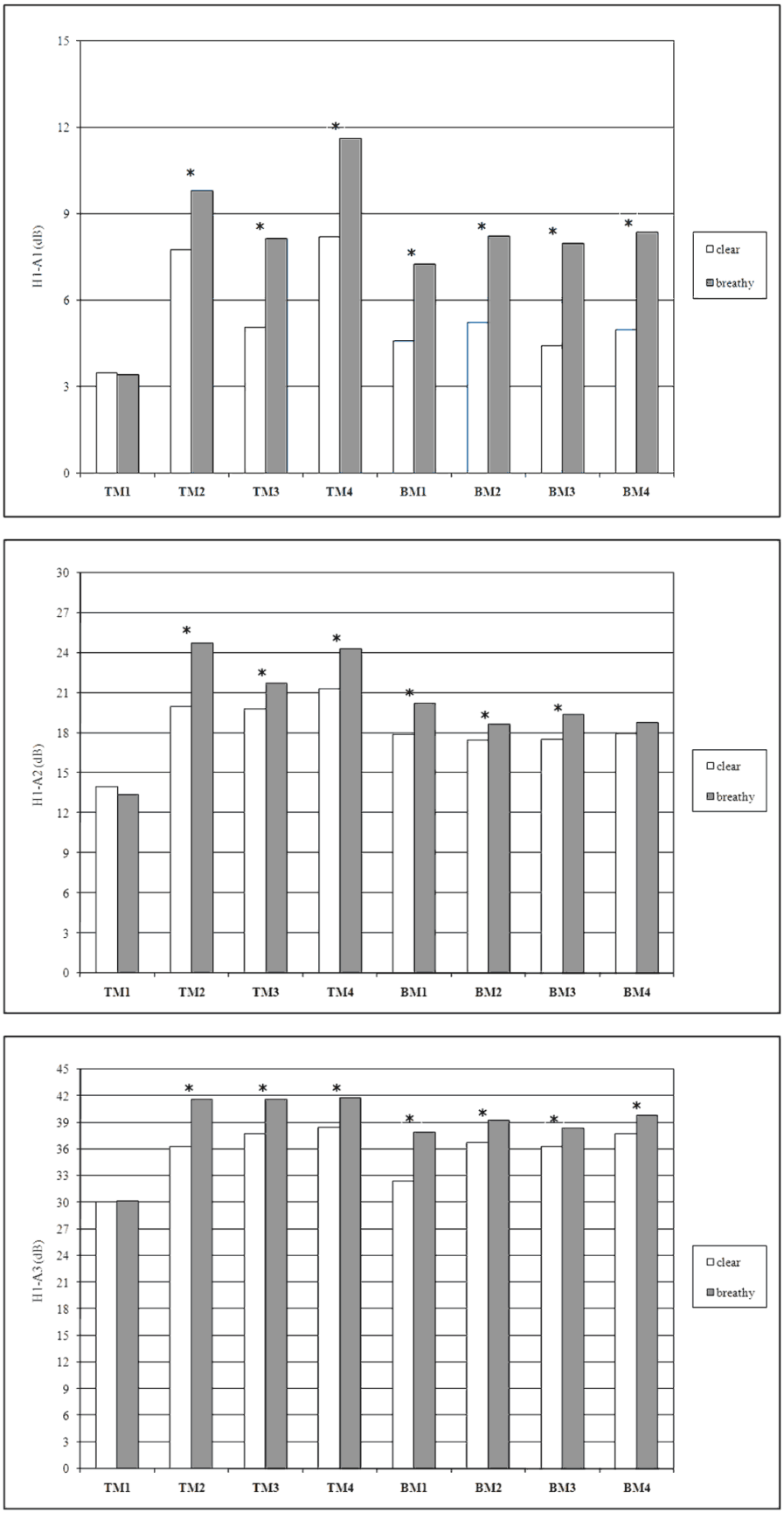


Figure 3.2: Mean values of H1-A1, H1-A2 and H1-A3 (in dB) at 25% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels (An asterisk indicates the values that are significantly different.)

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

Table 1: Mean of F0 values (in Hz) at 5 time points (0%-100%) (An asterisk indicates the values that are significantly different.)

Variety	Time / Vowel	0%	25%	50%	75%	100%
TM1	V	166.31*	178.63*	183.96*	168.71*	146.87
	V̤	122.04*	137.06*	162.82*	167.09*	140.43
TM2	V	167.38*	167.31*	162.99*	151.56*	141.33*
	V̤	153.57*	156.31*	156.21*	146.8*	137.03*
TM3	V	165.35*	170.42*	166.55*	150.53*	135.73*
	V̤	141.33*	148.14*	149.18*	139.16*	128.92*
TM4	V	163.37*	163.42*	156.04*	142.30*	130.30*
	V̤	137.59*	139.48*	135.73*	128.95*	124.50*
BM1	V	184.12*	189.65*	183.70*	162.68*	143.05*
	V̤	144.96*	153.96*	157.96*	151.20*	136.83*
BM2	V	183.55*	185.54*	178.35*	157.54*	141.16*
	V̤	147.42*	151.88*	151.93*	141.49*	128.53*
BM3	V	195.01*	199.79*	194.43*	169.25*	145.94*
	V̤	136.66*	143.72*	153.93*	151.54*	136.42*
BM4	V	178.52*	178.72*	168.41*	148.40*	134.82*
	V̤	144.68*	144.62*	142.98*	136.41*	128.76*

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

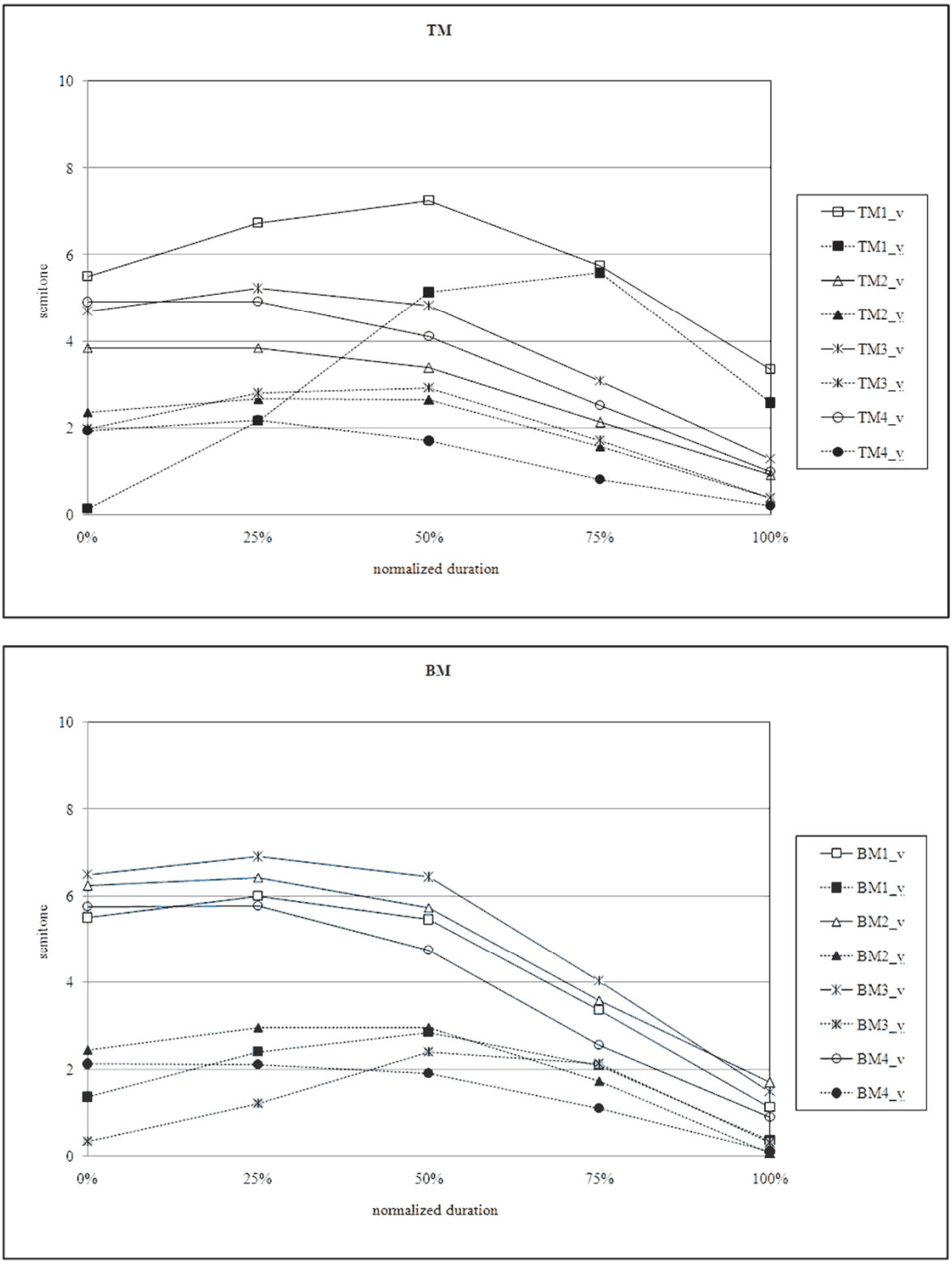


Figure 4: Semitone values at 5 time points of normalized duration from Thai Mon varieties (top) and Burmese Mon varieties (bottom).

Table 2: Mean of F1 and F2 values (in Hz) at 50% in steady state of vowel in TM varieties (An asterisk indicates the values that are significantly different.)

Variety	Vowel		i	e	ɛ	a	ɑ	ɔ	o	u	ɜ
TM1	F1	v	477.48	618.07	864*	1008	959.26	845.50	666	567.50*	567.50
		v̤	471.27	624.04	787*	979	-	852	726	726*	741
	F2	v	2112.68*	1977.50	1958	1558	1260.15	1116	945	964*	1252.50*
		v̤	2055.00*	1928.43	1865	1340	-	1148	974	850*	1244*
TM2	F1	v	478.90*	570.27	844.65	897.69	895.54	827.51	650.38	538.19	676.66
		v̤	453.90*	538.63	756.77	841.21	-	-	550.30	466.73	566.14
	F2	v	1972.47*	1854.48*	1726.11	1532.29	1383.59	1162.82	979.33*	985.78*	1335.15
		v̤	1918.63*	1802.86*	1771.13	1307.52	-	-	919.99*	902.21*	1259.34
TM3	F1	v	448.28	565.03*	872.55	979.89	868.07	761.16	603.12*	485.32*	664.45
		v̤	444.53	532.20*	-	899.00	-	-	580.66*	516.89*	600.82
	F2	v	2048.52*	1949.18*	1875.29	1512.47	1180.12	1051.15	910.36	949.52	1222.90*
		v̤	1979.85*	1827.95*	-	1372.45	-	-	906.78	893.21	1178.97*
TM4	F1	v	391.78	562.84	761.19	966.28*	814.32	726.20	583.43	489.12	600.47
		v̤	400.20	554.08	-	925.58*	-	-	567.16	470.03	593.13
	F2	v	2233.22*	2064.49	1863.34	1530.42	1174.64	1032.25	837.06*	959.66*	1429.59
		v̤	2194.72*	2036.71	-	1432.82	-	-	880.89*	868.13*	1357.41

Table 3: Mean of F1 and F2 values (in Hz) at 50% in steady state of vowel in BM varieties (An asterisk indicates the values that are significantly different.)

Variety	Vowel		i	e	ɛ	a	ɑ	ɔ	o	u	ɜ	ɨ
BM1	F1	v	465.79*	558.14	743.24	879.40*	810.77	736.14*	579.74*	442.14	492.18	643.62
		v̤	411.33*	552.99	748.80	854.46*	-	756.41*	613.29*	437.38	499.84	643.81
	F2	v	2056.71	1956.83	1796.81	1544.97	1160	1108.35*	946	985.39	1545.91*	1321.16
		v̤	1860.88	1935.15	1745.202	1389.76	-	1103.96*	930.21	956.28	1449.40*	1266.64
BM2	F1	v	439.82*	548.51*	756.91	905.13*	820.29	706.83*	581.35*	434.90	453.09	622.18
		v̤	400.70*	556*	762.54	871.01*	-	715.25*	596.87*	444.81	519.28	655.16
	F2	v	1975.74	1958.97	1800.29	1510.72	1171.80	1030.94*	868.58*	1056.11	1427.81	1379.56
		v̤	1989.51	1977.52	1736.63	1361.11	-	1059.01*	962.90*	1035.96	1470.31	1319.46
BM3	F1	v	454.17*	574.99	827.69	990.26	869.28	783.30	608.38	473.11	490.63	678.93*
		v̤	431.57*	570.68	840.38	988.49	-	805.12	625.22	459.69	541.53	665.35*
	F2	v	2189.97	2086.64	1861.49	1670.95	1167.24	1126.55	973.95*	1031.33*	1313.97	1272.63
		v̤	1968.39	2062.28	1871.64	1540.62	-	1118.27	911.32*	1085.55*	1327.21	1277.43
BM4	F1	v	444.70	591.13	706.97	855.53*	796.42	722.72*	598.82	458.16	457.08	638.55
		v̤	441.31	519.06	714.09	899.20*	-	756.35*	589.99	445.64	492.75	632.17
	F2	v	2148.38	1944.10	1725.24*	1518.98	1122.68	1072.13*	940.05	911.25	1361.99*	1258.39*
		v̤	1967.12	1881.31	1798.99*	1453.31	-	1107.02*	944.11	918.89	1481.86*	1310.62*

BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

5. Discussion

5.1 Vowel system

This study found minor differences between the TM and BM vowel systems, in line with Huffman (1987-1988). The number of monophthongs in TM and BM is similar to those of TM discussed in Bauer (1982) and those of BM in Jenny (2005) respectively. In addition, more diphthongs were found in BM varieties. However, this study cannot conclude whether this is an innovation or retention of vowels from old Mon. A fuller explanation would require a diachronic study. Nevertheless, it can be concluded that TM and BM varieties do belong to the same language.

5.2 Acoustic analysis

5.2.1 Relative amplitude

To produce phonation contrast, it is possible that vocal-fold velocity, a posterior glottal opening and ligament closure which depend on a degree of vocal-fold abduction and vocal-fold adduction and volume of air passing through are different in the particular contrast. This is related to the strength of higher frequencies in the spectrum. Even though, it is not absolute that one of which is greater than the others, the values of amplitude differences can distinguish the phonation contrast (Ladefoged, Maddieson and Jackson, 1988). In this study, the larger difference of relative amplitude mostly occurs in breathy vowels than that of clear vowels. Thus, it shows that distinct glottal stricture apparently exists among TM and BM varieties. Phonation type plays an important role in these varieties as well as Nakhon Chum Mon (Luangthongkum, 1988a). In addition, H1-An (H1-A1, H1-A2 and H1-A3), notably H1-A1, seems to indicate phonation contrast in the Mon varieties studied.

5.2.2 Fundamental frequency

The findings show that phonation type interacts with pitch in TM and BM varieties: clear vowels with higher pitch and breathy vowels with lower pitch. According to statistical analysis, F0 values of clear vowels and breathy vowels are significantly different at every time point in most varieties. Pitch is apparently a salient exponent, as found in Lee (1983) and Luangthongkum (1988a); however, it occurs with phonation type, for example in Nakhon Chum Mon (Luangthongkum, 1988a). This may lead both TM and BM varieties to become tonal languages. Moreover, the difference of pitch contours in BM and TM1 is larger than that of TM2, TM3 and TM4 as shown in Figure 4. Pitch may be more important than other cues in BM and TM1 perception. To give a definite answer, a perception study is needed.

5.2.3 Formant frequency

F1 and F2 values between clear vowels and breathy vowels do not show any systematic differences. The vowel quality of most clear vowels is similar to that of breathy vowels. Vowel quality cannot indicate whether breathy vowels are more close or more open, more front or more back or more centralised (Shorto, 1966) than clear vowels. Neither raising the larynx versus lowering the larynx (Thurgood, 2000) nor tongue-root retraction versus tongue-root advancement (Gregerson, 1976) has been found to be a primary exponent of register contrast in TM and BM.

Nonetheless, the limitations of acoustic measurement at 50% in steady state of vowels may not appropriately demonstrate the real characteristics of vowels in BM. Many on-gliding and off-gliding vowels are found in BM, while the vowel characteristics are mostly pure in TM, as shown in Figure 5.

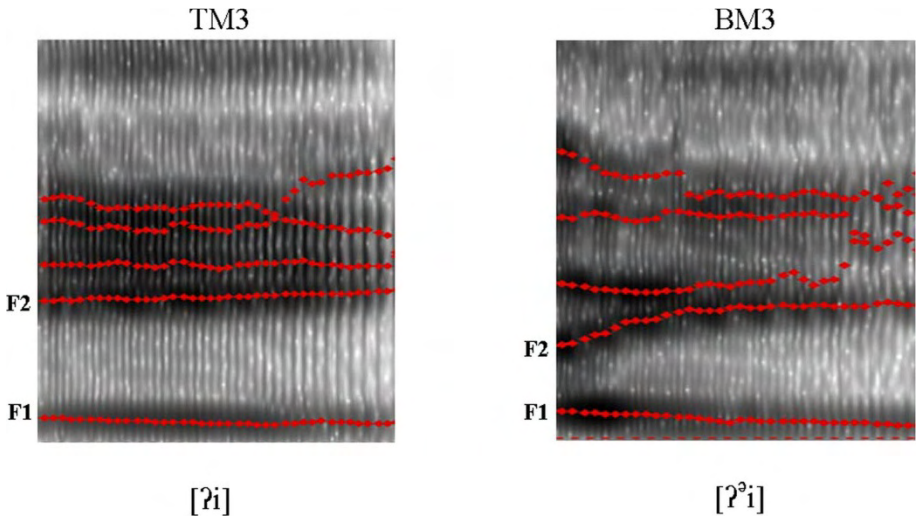


Figure 5: Examples of wide band spectrogram showing F1 and F2 movements in a Thai Mon variety (left) and a Burmese Mon variety (right)

Figure 5 shows F1 and F2 values at the beginning of the vowel for the word /həʔi/ 'cucumber' – for TM they are static, while those of BM are dynamic, especially F2 values. This reflects tongue movement from the centre of vowel area towards the front, for example [əi] occurring in BM varieties.

On-gliding and off-gliding appear with both clear vowels and breathy vowels. This finding differs to the viewpoint of Thurgood (2000) which speculates that clear vowels occur with off-gliding and breathy vowels with on-gliding. This findings discussed in this paper suggest that voice quality and vowel quality correlation do occur in BM, affecting vowel characteristics and possibly increasing the number of vowel phonemes. In this way, BM varieties could become restructured.

5.2.4 Duration

Breathy vowels can be perceived as longer than clear ones although Mon has no vowel length distinction. However, duration of most clear vowels and breathy vowels in this study are not significantly different which differs from the finding of Lee (1983). In these varieties, vowel length may not be an important exponent in register complexes but it may in other varieties contribute to indicate the distinction of clear vowels and breathy vowels.

6. Conclusion

Even though it can be said that Mon is a register language which phonation type combines with pitch patterns, vowel quality may also become a prominent component. Obvious pitch patterns in TM and BM varieties could result from internal and external factors. To illustrate pitch per se is one parameter of register complexes. In addition, language contact with Thai and Burmese, a tonal language, could help enhance salience of pitch. However, vowel quality as another parameter of register complexs could also develop and might be a salient parameter like those found in BM varieties. Consequently, TM varieties alone could become tonal while BM varieties may evolve to either a tonal or restructured language. Nevertheless, the perception test might help exhibit an important cue for native speakers in order to determine the tendency of language change in the future.

Acknowledgement

My gratitude goes to Prof. Dr. Theraphan Luangthongkum, my dissertation advisor. I would like to thank Prof. Dr. Christian Bauer for his valuable suggestions and express my appreciation to Dr. Mathias Jenny for his help. I also would like to thank the Royal Golden Jubilee (RGJ) for a grant (PHD/0172/2551), and Sripatum University for financial support during studying at Chulalongkorn University. This research cannot be done without all the help from my language consultants. I am thankful for their kind co-operation.

References

Abramson, Arthur S., Theraphan Luangthongkum and Patrick W. Nye. 2004. Voice register in Suai (Kuai): an analysis of perceptual and acoustic. *Phonetica* 61: 147-171.

Abramson, Arthur.S., Patrick W. Nye and Theraphan Luangthongkum. 2007. Voice register in Khmu: experiments in production and perception. *Phonetica* 64: 80-104.

Bauer, Christian. 1982. *Morphology and syntax of spoken Mon*. Thesis for the degree of Doctor of Philosophy: University of London.

Bauer, Christian. 2010. *A Guide to Mon Studies*. Berlin: Humboldt University. [2nd edition]

Blankenship, Barbara. 2002.The time course of nonmodal phonation in vowels. *Journal of Phonetics* 30:2.

Brunell, Marc. 2005. *Register in Eastern Cham: phonological, phonetic and sociolinguistic approach*. Doctorial dissertation, Cornell University, Ithaca, New York.

Charoenma, Narumol. 1982. The phonologies of a Lampang Lamet and Wiang Papao Lua. *MKS* 11: 35-45.

DiCanio, Christian. 2009. The acoustic of register in Takhian Chong. *Journal of the international Phonetic Association* 39(2): 162-188.

Diffloth, Gérard.1984. *The Dvaravati – Old – Mon language and Nyah-Kur (Monic Language Studies1)*. Bangkok: Chulalongkorn University Printing House

Ethnologue. 2013. Available from: http://www.ethnologue.com/language/mnw [2013, February 2]

Esposito, Christina M. 2006. *The Effects of Linguistic Experience on the Perception of Phonation*. Ph.D. Dissertation University of California, Los Angeles.

Gregerson, Kenneth. L. 1976. Tongue-root and register in Mon-Khmer. In P.N. Jenner et al. (eds.), *Austroasiatic I*:323-366.

Jenny, Mathias. 2005. *The Verb Systems in Mon*. Zurich: University of Zurich.

Henderson, Eugénie J. A. 1952. The Main Features of Cambodian Pronunciation. *Bulletin of the School of Oriental and African Studies* 14: 149-174.

Huffman, Franklin. 1976. The register problem in fifteen Mon-Khmer languages. *Austroasiatic Studies part I*: 25-37.

Huffman, Franklin. 1985. Vowel permutations in Austroasiatic languages. *Linguistics of the Sino-Tibetan Area*. Pacific Linguistics Series c-no. 87.

Huffman, Franklin. 1987-1988. Burmese Mon, Thai Mon, and NyahKur: a synchronic comparison. *MKS* 16-17: 31-84.

Keating, Patricia., Christina M.Esposito., Marc Garellek M., Sameer ud Dowla, Khan and Jianjing, Kuang. 2010. Phonation contrast across languages. *UCLA Working Papers in Phonetics* 108: 188-202.

Ladefoged, Peter. 1971. *Premilinaries to Linguistic Phonetics*. Chicago: University of Chicago.

Ladefoged, Peter., Ian Maddieson, and Jackson, M. 1988. Investigating phonation types in different languages. In O. Fujimura (ed.), *Vocal Fold Physiology: Voice Productions, Mechanism and Functions* 297-317. New York: Raven Press.

Lee, Thomas. 1983. An acoustical study of the register distinction in Mon. *UCLA Working Papers in Phonetics* 57: 79-96.

Luangthongkum, Theraphan. 1982. Register without tongue-root in NyahKur (Chao Bon). *The 15th International Conference on Sino-Tibetan Languages and Linguistics*, August 1982, Beijing.

Luangthongkum, Theraphan. 1988a. Another look at the register distinction in Mon. *The International Symposium on Language and Linguistics* 22-51. Bangkok: Thammasat University.

Luangthongkum, Theraphan. 1988b. Phonation type in Mon-Khmer languages. In O. Fujimura (ed.), *Vocal Fold Physiology: Voice Productions, Mechanism and Functions* 319-333. New York: Raven Press.

Luangthongkum, Theraphan. 1990. The interaction between pitch and phonation type in Mon: phonetic implications for a theory of tonogenesis. *MKS* 16-17: 11-24.

Mundhenk, Alice Tegenfeldt and Hella Goschnick. 1977. Haroi Phonemes. *Papers in Southeast Asian Linguistics No.4, Chamic Studies*, ed. D. Thomas et al., vol. 4, pp. 1-16. Pacific Linguistics, the Australian National University.

Phillips, Richard., John Phillips and Carolyn Miller. 1976. The Brũ vowel system: alternate analyses. *MKS* 5: 203-217.

Ploykaew, Pornsawan. 2001. The phonology of Samre. *MKS* 31:15-27.

Premsrirat, Suwilai. 1996. Phonological characteristics of So (Thavung), a Vietic language of Thailand. *MKS* 26: 161-178.

Premsrirat, Suwi lai. 2004. Register complex and tonogenesis in Khmu dialects. *MKS* 34: 1-17.

Shorto, Harry.L. 1962. *A Dictionary of Modern Spoken Mon*. London: Oxford University Press.

Shorto, Harry.L. 1966. Mon vowel systems: A problem in phonological statement. In Bazell et al. (ed.), *Memory of J.R. Firth*, 398-409. London: Oxford University Press.

Sukgasame, Preecha. 2005. Phonological variations and changes due to language contact: A case study of consonants in four Kuai-Kui (Suai) dialects. *MKS* 35: 37-54.

Thurgood, Graham. 2000. Voice quality differences and the origin of diphthongs. *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect* (2000): 295-303.

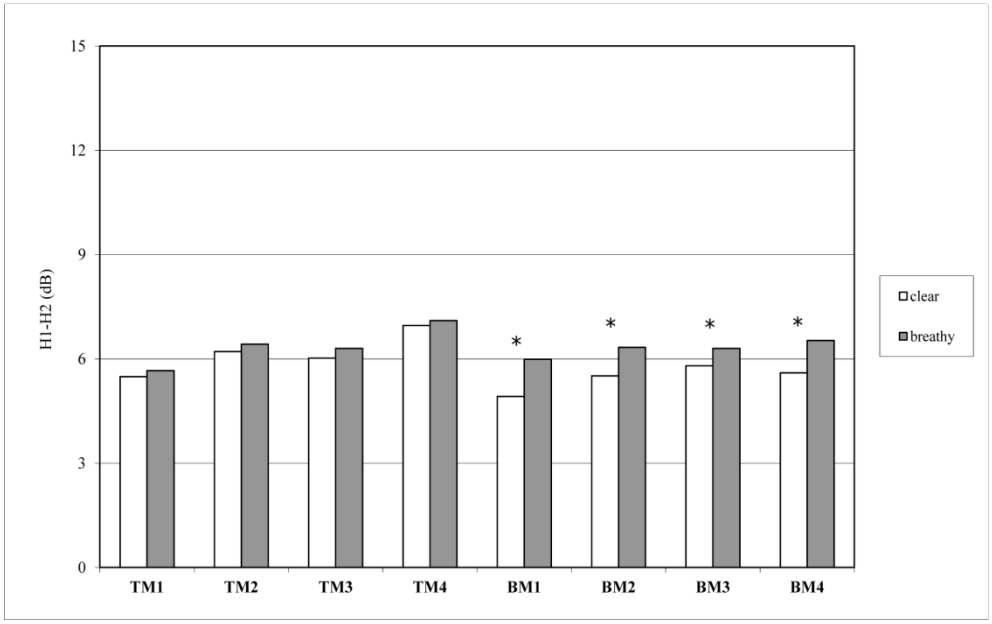
Watkins, Justin. 2002. *The Phonetics of Wa*. Canberra: Pacific Linguistics.

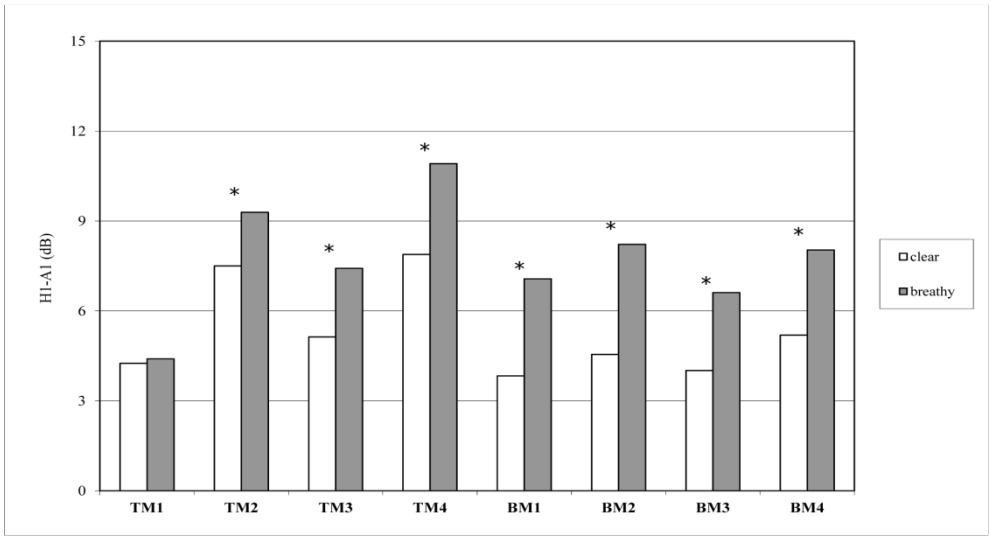
Wayland, Ratree. and Allard Jongman. 2003. Acoustic correlates of breathy and clear vowels: The case of Khmer. *Journal of Phonetics* 31: 181-201.

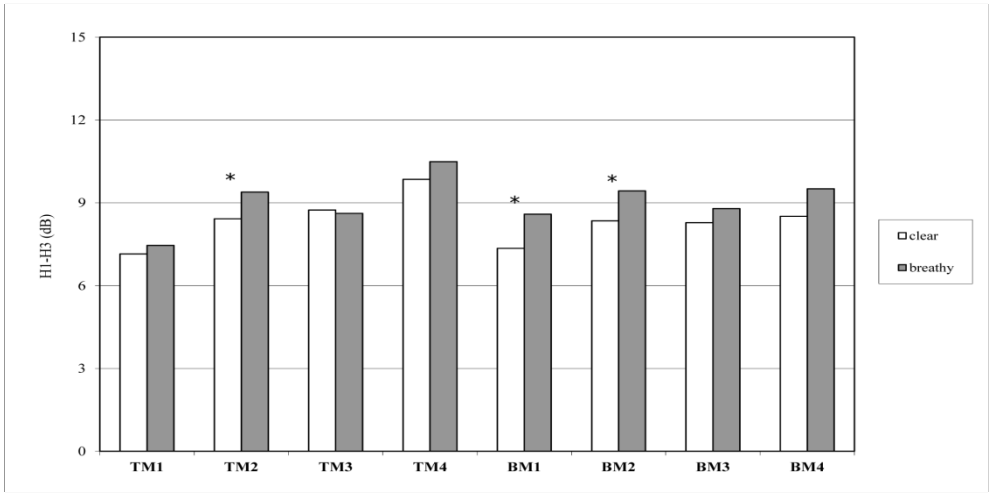
BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

Appendix 1

Mean values of H1-H2, H1-H and, H2-H4 (in dB) at 50% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)

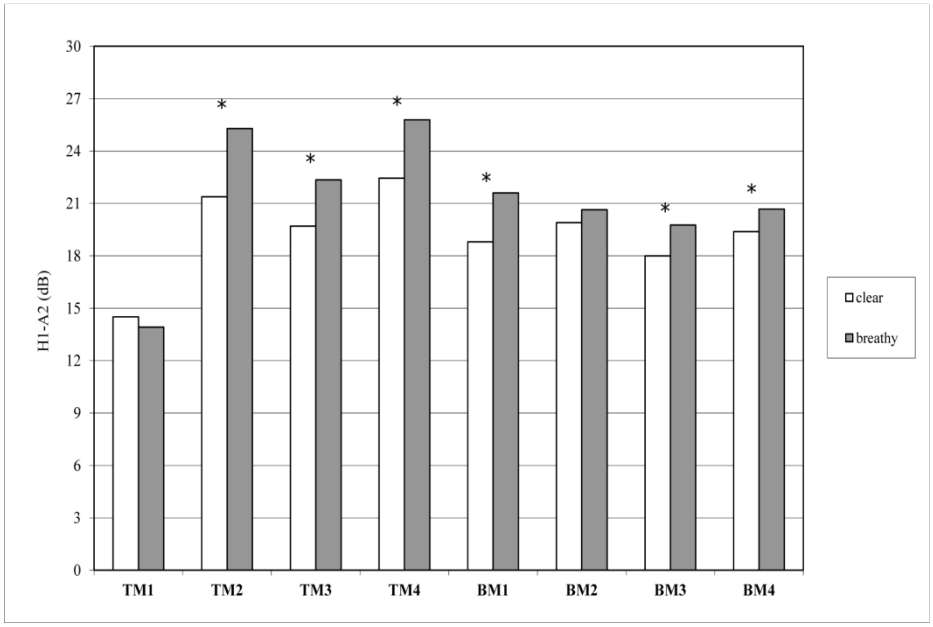


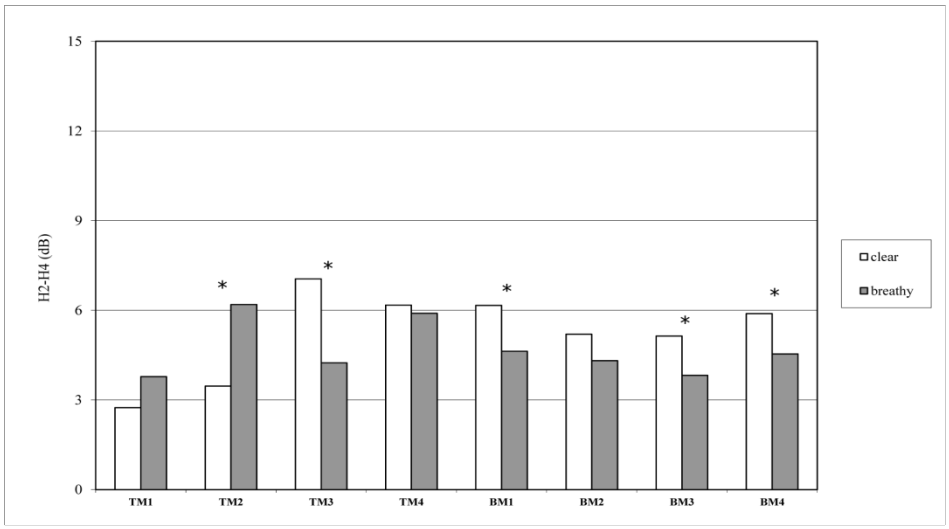


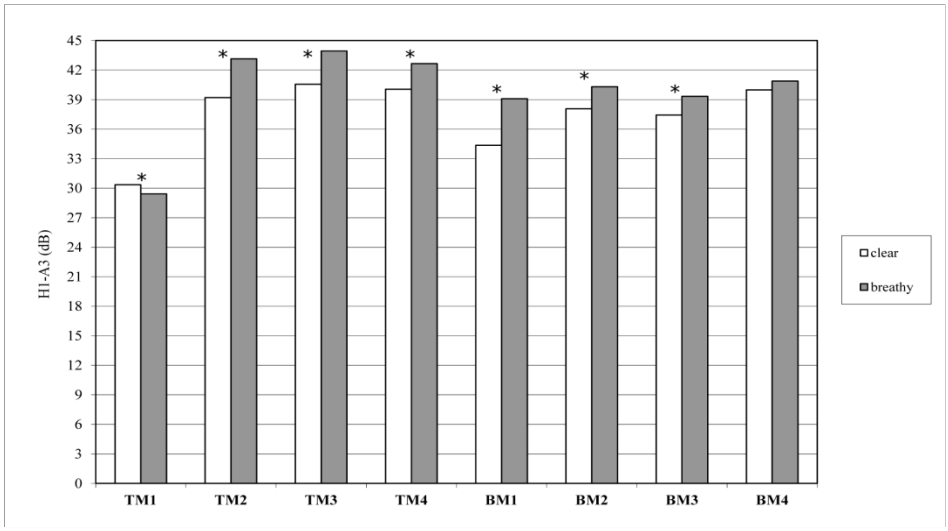


BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

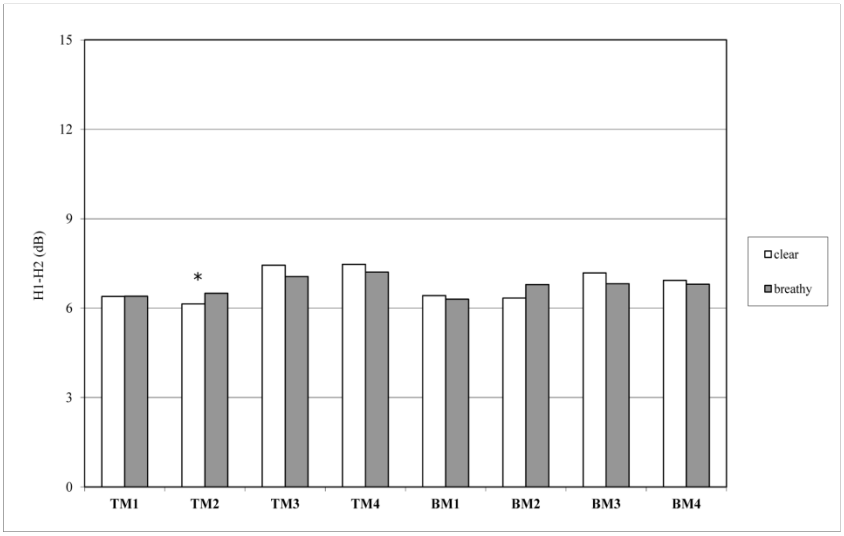
Mean values of H1-A1, H1-A2 and H1-A3 (in dB) at 50% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)

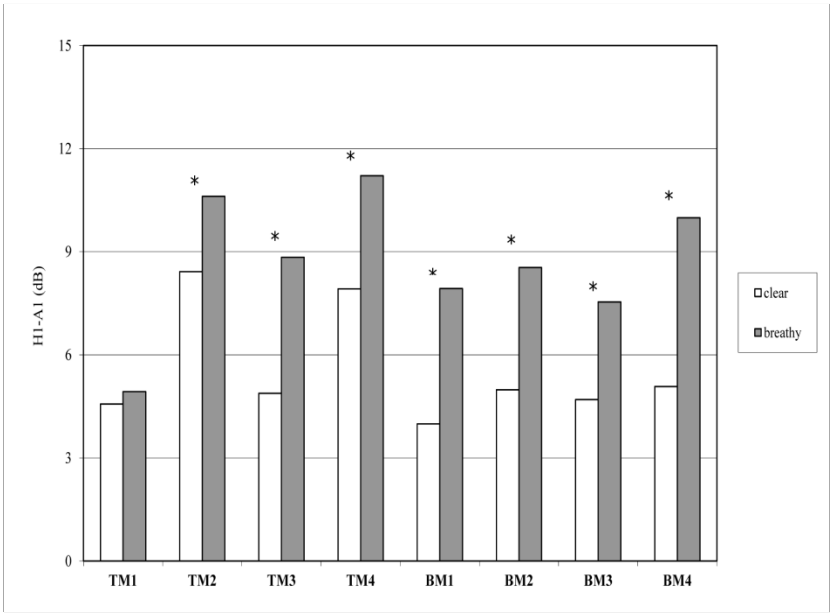


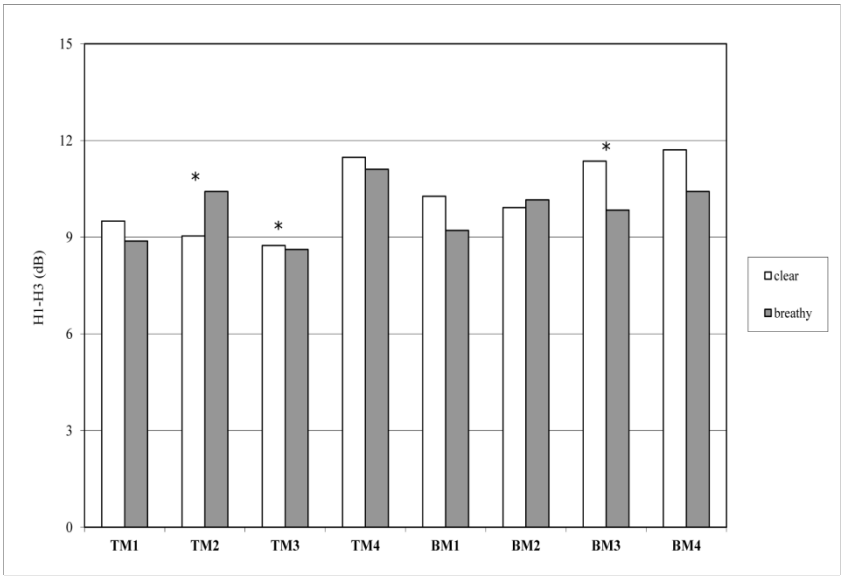




Mean values of H1-H2, H1-H3 and H2-H4 (in dB) at 75% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)

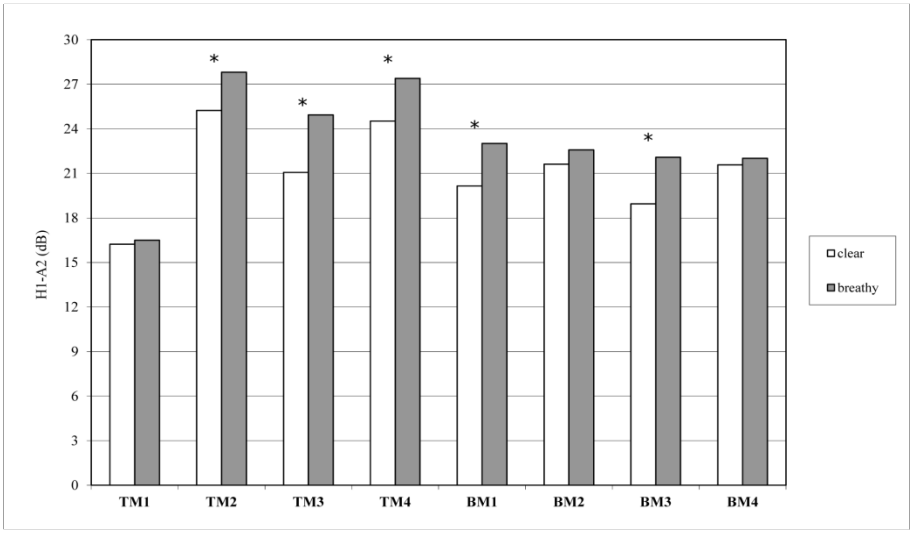


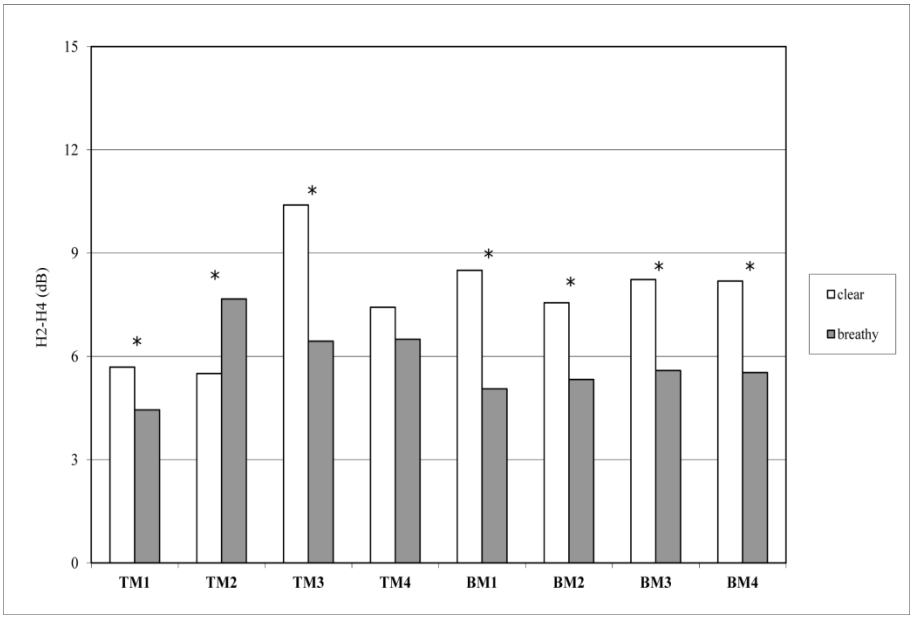


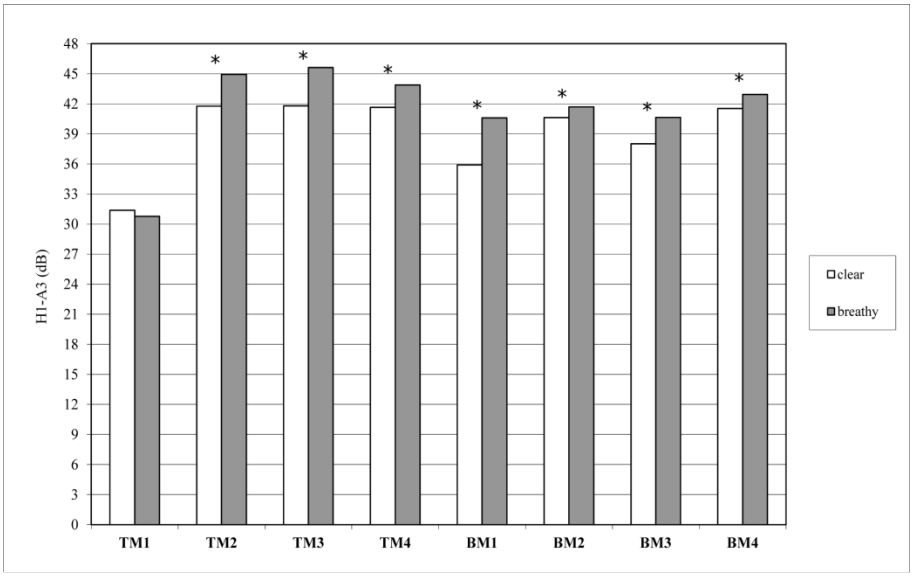


BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

Mean values of H1-A1, H1-A2 and H1-A3 (in dB) at 75% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)

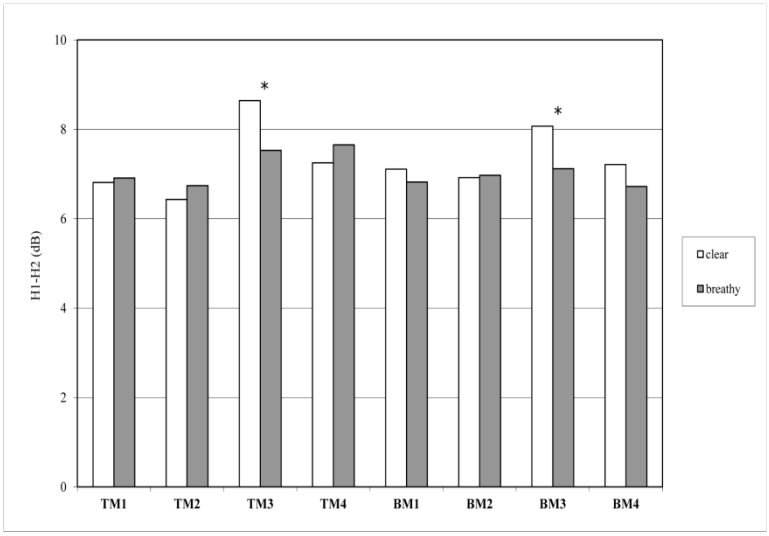


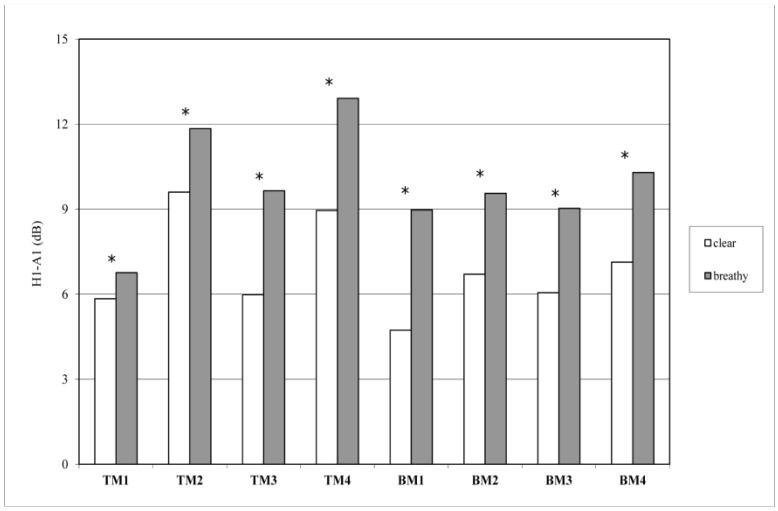


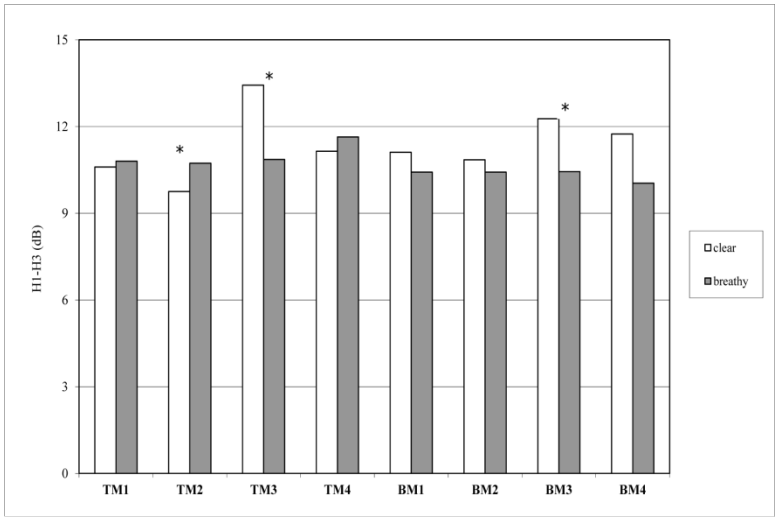


BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

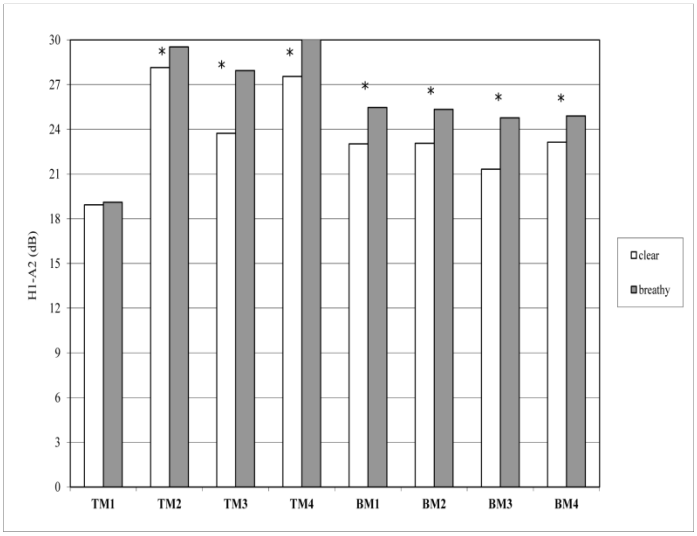
Mean values of H1-H2, H1-H3 and H2-H4 (in dB) 100% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)

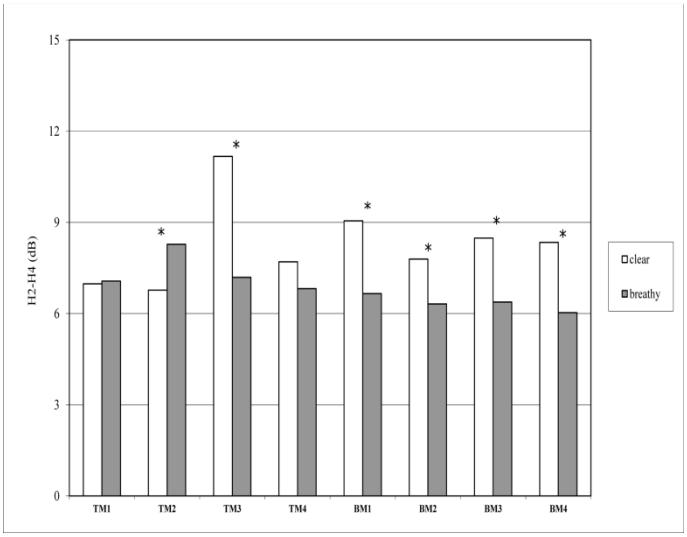


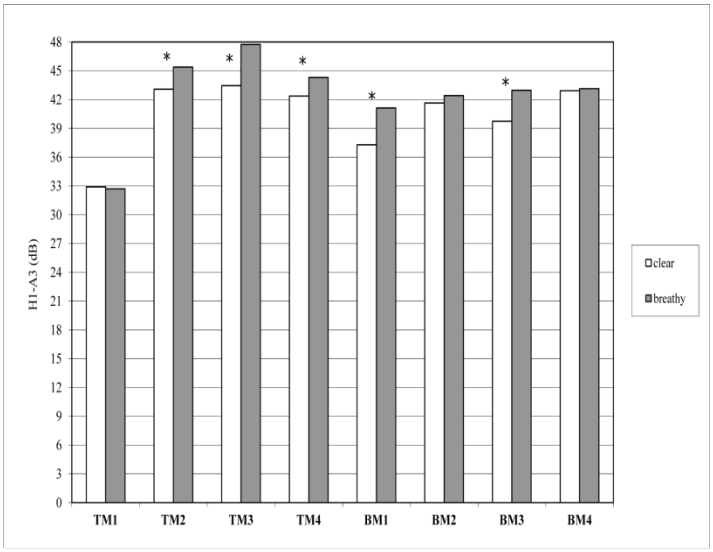




Mean values of H1-A1, H1-A2 and H1-A3 (in dB) 100% from four Thai Mon and four Burmese Mon varieties with clear vs. breathy vowels. (An asterisk indicates the values that are significantly different.)



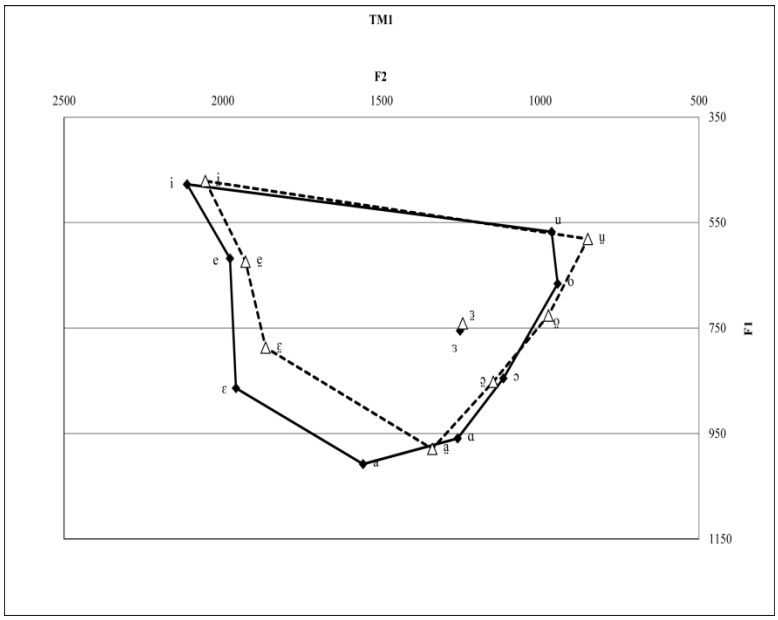


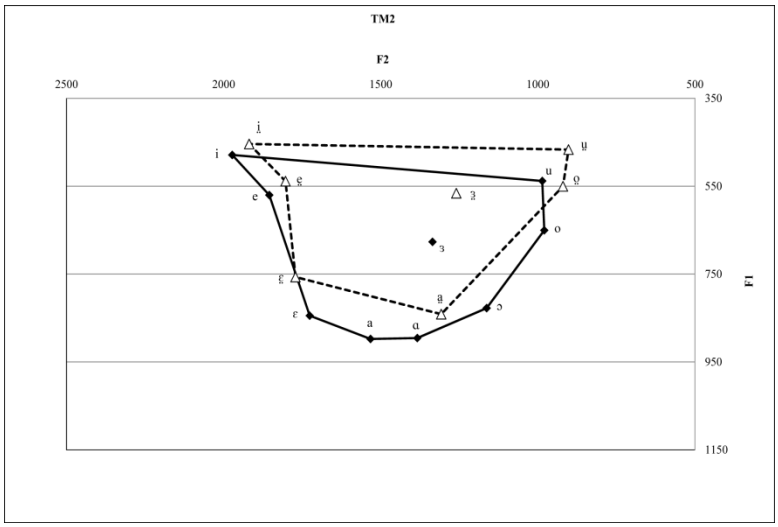


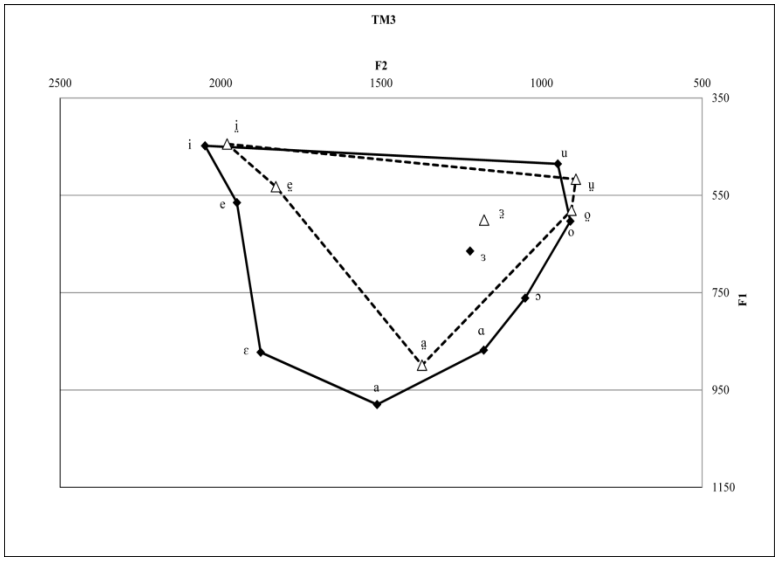
BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

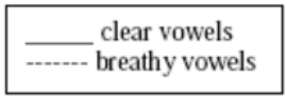
Appendix 2

Vowel space of clear vowels and breathy vowels in four Thai Mon and Burmese Mon varieties

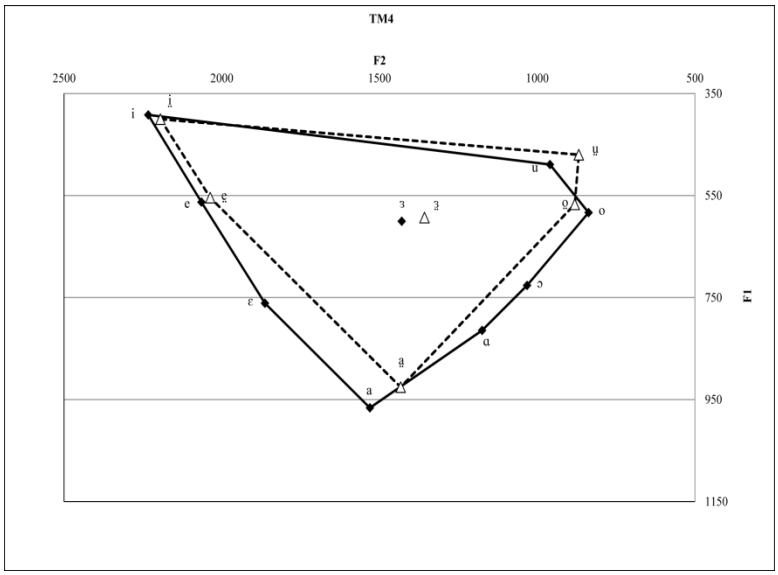


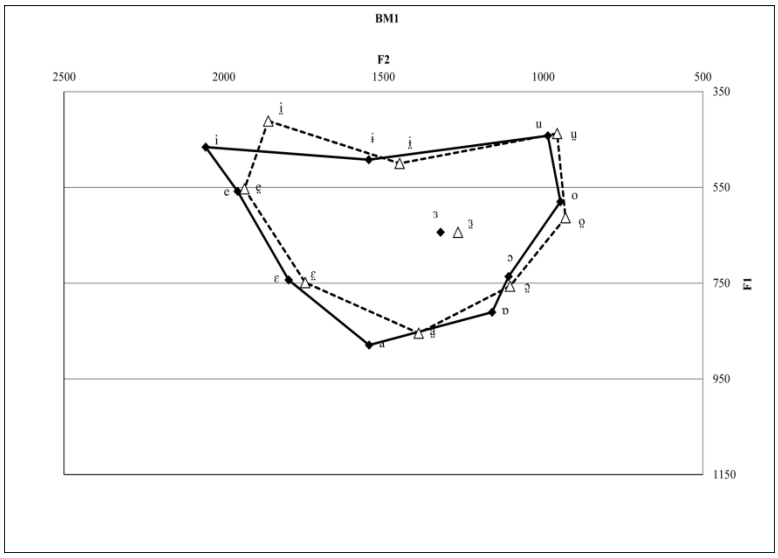


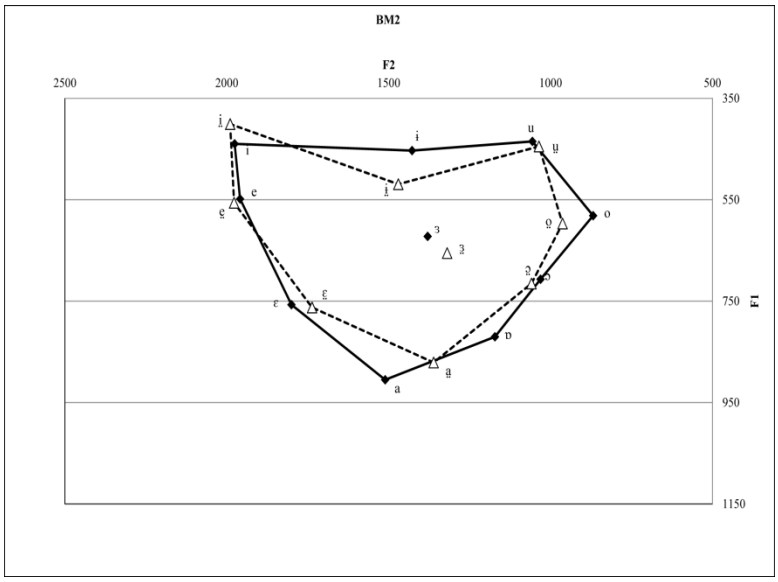


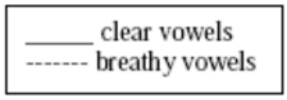


BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

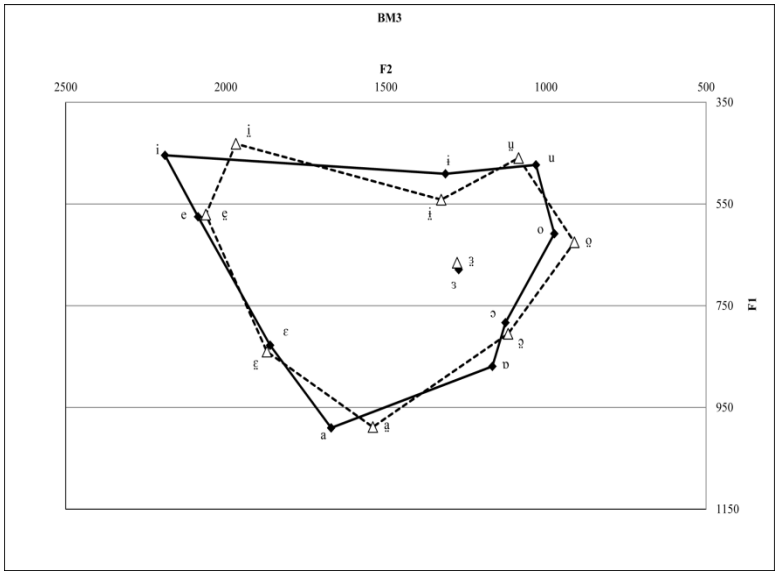


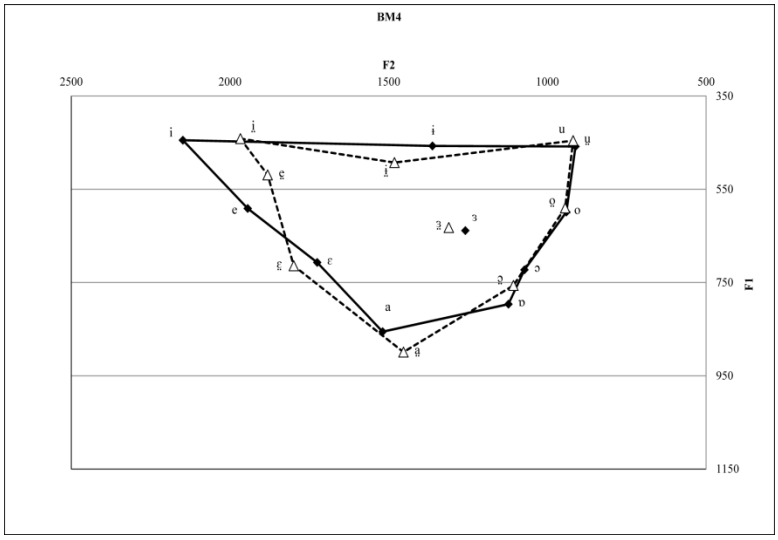


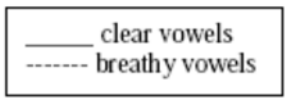




BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80



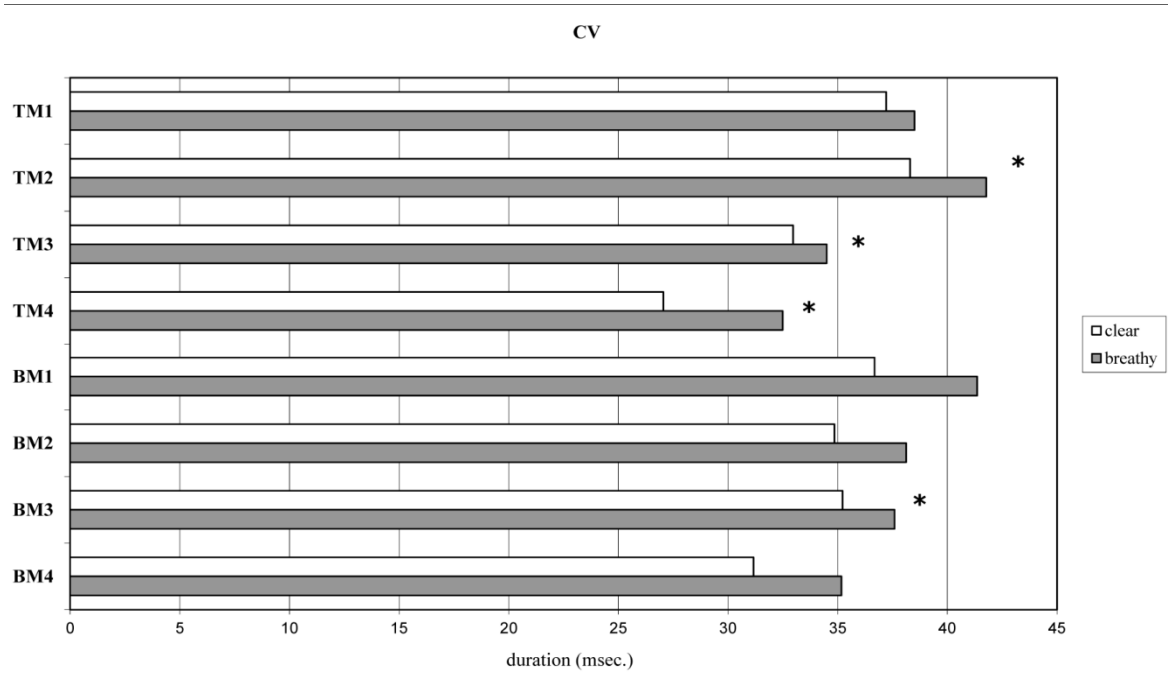


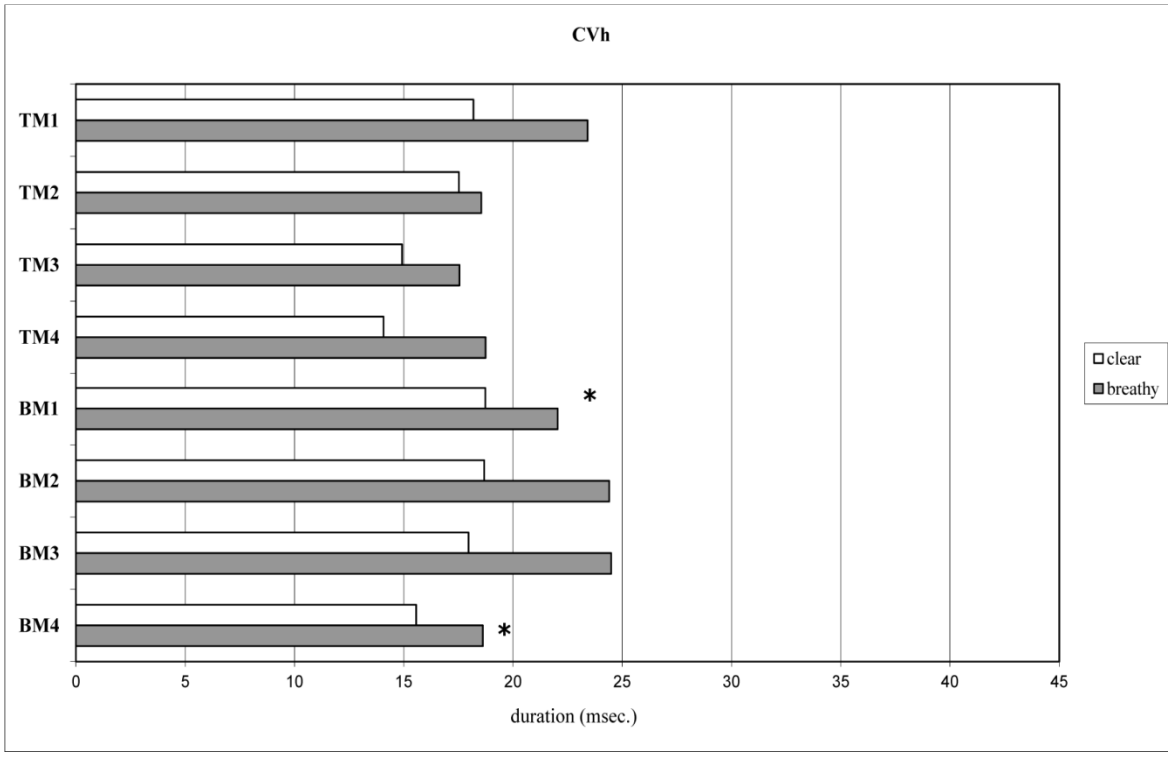


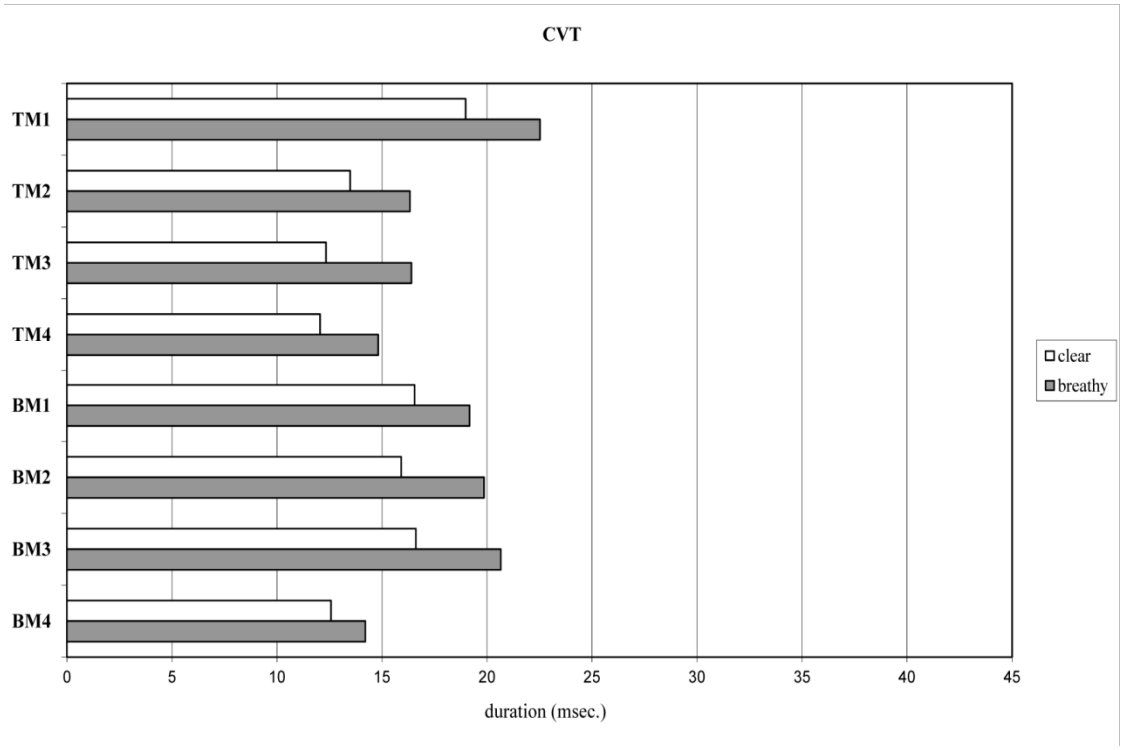
BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

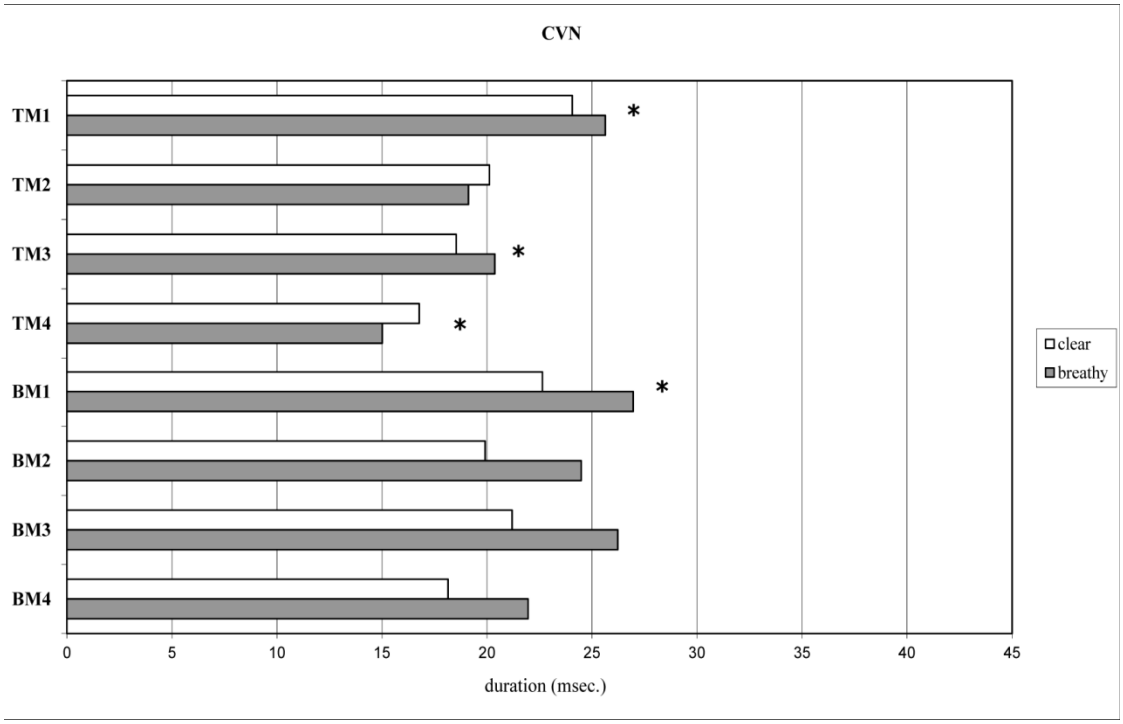
Appendix 3

Duration of clear vowels and breathy vowels in CV, CVh, CVT and CVN syllable types in four Thai Mon varities and four Burmese Mon varieties. (An asterisk indicates the values that are significantly different.)





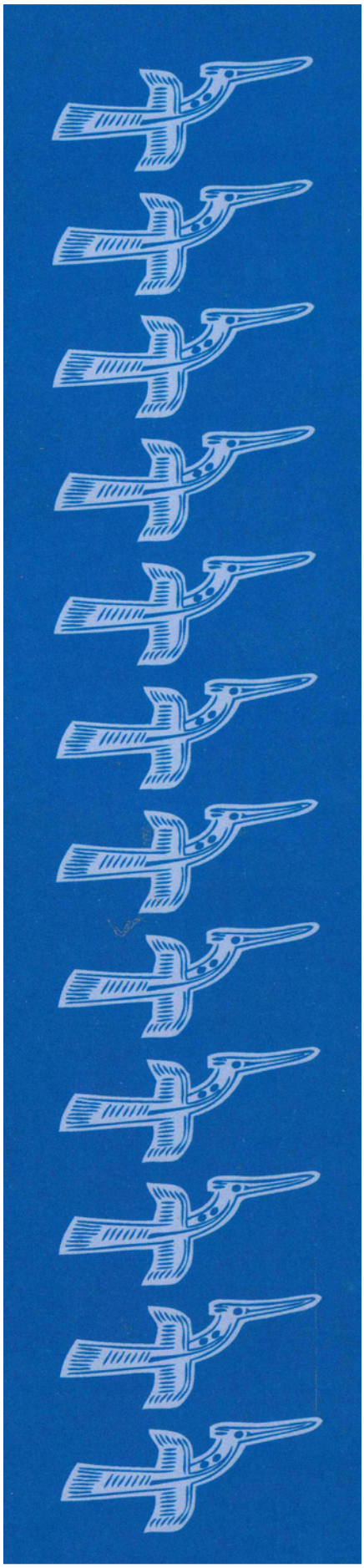




BEHR, Narinthorn Sombatnan. 2013. A comparison between the vowel systems and the acoustic characteristics of vowels in Thai Mon and Burmese Mon: a tendency towards different language types. *Mon-Khmer Studies* 42:54-80

Mon-Khmer Studies

Volume 42



Author: P. K. CHOUDHARY
Title: *Tense, Aspect and Modals in Ho.*
Pages: 81-88
Date received: 5/9/2013
Revised text accepted: 6/11/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

Tense, Aspect and Modals in Ho

P. K. CHOUDHARY
CIIL, Mysore

Abstract
North Munda languages in general and Ho in particular have a grammatical category for tense marking. There are lots of aspect markers in Munda languages. We can draw time of a particular sentence from the aspect part of the sentence. It has progressive marker *tan* in present progressive sentences.Ho has several aspect marker for past time i.e. *keḍ*, *leḍ*. *yen*, *ɟan* and *yeḍ* etc to show different time of the past. There is not any aspect marker for future time in Ho. There are two Modals in Ho, *dai* 'can/may' and *utar* 'must' to perform various purposes of the language. This paper discusses the aspect markers of Ho and also explores its role in the said language.
Keywords: tense, aspect, modals
ISO 639-3 language codes: bnq

1. Introduction

In language, time is expressed with tense that relates events situation with moments of speaking, whereas aspect represents the ways of viewing internal chronological constituency of a situation (Comrie, 1976, p.3). Thus the purpose of tense is to situate event in time line with reference to other event. Aspect on the other hand reflects speaker's internal point of view on a given situation.

There are two instruments to mark tense i.e. grammatical and lexical categories. That is some languages have grammaticalised a category to express time reference; the semantic concept of time reference may be grammaticalised in the language in which we can say the language has tenses. Grammaticalisation refers to integration in to grammatical system of a language, where as lexicalization refers merely to integration in to the lexicon of a language without any reflection on its grammatical structure. Many languages lack the tense i.e. do not have time reference, though probably all languages of the world have means of expressing tense i.e. adverbials that locate situation in time. North Munda languages, as the verb do not change for person and number, have a grammatical category for the purpose.

Tense is primarily a category of the verb of the sentence. There are languages which have only two way tense distinctions. As many languages have a basic two-way division with either opposition between past versus non-past, future versus non-future and perfect versus non-perfect (imperfect). Future versus non-future split occasioned primarily by mood and perfect versus imperfect by aspect.

As far as modals are concerned, many languages expressed it in the verb with verbal morphology; a verb has obligatory mood markers in a language. The shape of the verb also conveys information about person, number of the subject along with modality. Some languages like German have lexical words such as may, can, must, could, should etc to show modality. Like an adverb a verb can express modality lexically. Complementizers and conjunctions may convey modality in some languages.

Ramswami (2007:96) has shown that Ho has Past and Non-Past distinction. He further says that the three past markers are *ɟan*, *keḍ* and *ki*. Deeney (2002:28) has even shown that there is future tense marker in Ho but I have different idea regarding future tense of Ho. In this paper we will see the case of Ho with regards to tense aspect and modals. I claim that Ho has Perfect and Imperfect distinction without having pure Tense marker in the language.

2. Tense

Bernard Comrie (1985) defines tense as the grammaticalisation of location in time. Languages of the world vary in tense systems; that is verbs of some languages do not change for tenses but remains same in all tense. Many languages have grammatical category to mark time

CHOUDHARY, P. K. 2013. Tense, Aspect and Modals in Ho.
Mon-Khmer Studies 42:81-88

reference; lack tense marker. North Munda languages in general (i.e. Santali, Mundari and Ho) and Ho in particular too have temporal adverbials such as *tisiŋ* 'today', *gapa* 'tomorrow', *hola* 'yesterday', *setaʔ* 'morning', *tārāsiŋī* 'afternoon', *nīdē* 'night', *āyūb* 'evening' etc. to mark time reference; there are a lots of such time adverbials available in the language to locate time.

Ho has two tense divisions namely perfect and imperfect. There is no pure tense marker in the language as verbs do not change for any tense moreover aspect differs in perfect and imperfect tenses.

2.1. Imperfect Tense

Imperfect means incomplete or may be progressive. It also covers continuity of action as in the case of progressive sentences. In Ho imperfect also covers incomplete work or the work whose result has yet to come.

1. *aɲ* *seno-^{w}a*
I go-FM
'I go'

2. *aɲ* *seno-tan-a*
I go-PROG-FM
'I am going'.

Other present tense marker like present perfect tense, Ho speakers use perfective aspect marker to express the perfectness of the sentence. There is past continuous marker aspect *–taiken* which has been discussed under aspect section as it is used as past progressive marker as well.

2.2. Perfect Tense

Perfect denotes completion of an action having present relevance. As mentioned above, Ho has many perfect markers. There are distinction between completetive and in-completive aspects in Ho. Sometimes, they also use *–liya,* as past tense marker. For example:

3. *aɲ* *aɲ-te-ɲ* *goʔe-n-**ɟan**-a*
I I-pp-1SG kill-REFL-PRF-FM
'I killed myself'

4. *aɲ* *aɲ-te* *ʈhokan-**len**-a*
I I-pp hit-PRF-FM
'I hit myself'

5. *aɲ* *kuɽi* *nel-**liʸ**-a*
I girl see-PRF-FM
'I saw the girl'

6. *rām* *sīta-ke* *nel-**liɟ**-a*
Ram Sita-ACC see-PRF-FM
'Ram saw Sita'

Other examples are presented under aspect section.

2.3. Future Tense

There is neither a future tense marker nor any aspect for future tense in Ho. For example:

7. *aɲ* *intaʔre-ŋ*[1] *tain-a*
I there –1SG be-FM
'I will be there'

8. *gapā* *aɲ* *bar* *baje* *seno-^{w}a*
tomorrow I two o'clock go-FM
'I shall go at 2'clock, tomorrow.'

[1] In influence of Mid-close vowel 'e' palatal nasal *ɲ* became velar nasal *ŋ*.

3. Aspect

Tense and aspect both are concerned with time but differently. While discussing aspect it is necessary to see the differences in states, events and process. As mentioned above aspect may be inflectional or may have separate grammatical category.

According to Comrie (1976:3), aspects can be defined as 'different ways of viewing the internal sequential constituency of a situation'. The most well studied aspectual contrast is the one between perfect and imperfect. The term 'perfective' contrasts with 'imperfective' and denotes a situation viewed in its entirety, without regard to internal temporal constituency; the term 'perfect' refers to a past situation which has present relevance (1976:12). Ho has a number of aspects, which have been classified by different linguists like Deeney (2002) and Ramswami (2007) in their Ho[2] grammars.

Aspect has been classified as perfect and imperfect. Further imperfect may be subdivided in to two distinct concept of habituality and continuous/durativity. Aspectual differences in a language vary as some has differences in one tense but other has more than one but it is past tense that most shows the aspectual differences in world languages.

3.1. Imperfect Aspect

The progressive denotes continuity; it is same as continuousness as defined traditionally. Let us see the role of ***-tan,*** a present progressive marker in Ho. In present tense, it occurs as progressive marker. For example:

9. *en* *kuɽi* *hiju-**tan**-a*
that girl come-PROG-FM
'That girl is coming'.

10. *aɲ* *seno-**tan**-a*
I go-PROG-FM
'I am going'.

-taiken is past progressive tense marker which also express duration such as an event started in past and have relevance in present time too. It is combination of *tai*[3] 'to be or remain' plus *ken* 'a perfective aspect'.

11. *aɲ* *ninta?* ***tāi-ken**-a-ɲ*
I here be-PRF-FM-1SG
'I used to be here/ I used to live here'.

12. *am* *hola* *konʈā* *seno-* ***tai-ken**-a-m*
you yesterday where go-be-PRF-FM-2SG
'Where were you going yesterday?'

The examples shown below have been taken from Deeney (2002); I feel he has not broken the verb and aspect properly. It should be like shown above in example (11) and (12).

13. *maraŋ-e* ***taiken**- a* *(Deeney,2002:9)*
big-3SG ASP-FM
'He was big'

14. *owa-* *re-ko* ***taiken**-a* *(Deeney,2002:59)*
house-pp-3PL ASP-FM
'They were in the house'.

-akan as durative marker as it has been mentioned by Deeney (2002:41) but see in example (18) it shows something else i.e. *hiju-akan-a*, may be said here without alteration in meaning. Therefore, in my view this is *-kan* which is a continuation marker such as 'keep' and 'continue' in English. As far as first *-a* in *akan* and *-w* between vowels are concerned these are for syllable breaking and glide insertion between two vowels respectively.

---

[2] Kherwarian group have three important languages namely Santali, Mundari and Ho. Out of these three Ho has been considered as dialect of Mundari as it closely resembles with Mundari and Bhumij.

[3] *tai* means be or remain when it combined with *ken* it means in past time someone was doing something.

15. *ape dub-**akan**-a-pe* *(Deeney,2002:41)*
you seat-ASP-FM-2PL
'You keep sitting'

16. *ape horo^{w}-**aka**-pe* *(Deeney,2002:41)*
you(pl) man-ASP-2PL
You keep guarding it

17. *ape horo^{w} **aka**-ko-pe* *(Deeney,2002:41)*
you(pl) man ASP-3PL-2PL
'You keep guarding them'

18. *miɋo ho hiɲu-kan-a*
one man come-ASP-FM
'One person came'.

-kan as perfective but in-completive marker or in other word the action whose result has yet to come. It is used with intransitive verb whereas *-kaɋ* may be used with transitive verb.

19. *aɲ ɟetana-re-o ka-ɲ dub-**kan**-a*
I any-pp-EMP not-1SG sit-ASP-FM
'I am not sitting on anything'.

20. *borzo bazār-te sena-**kan**-a*
Borzo market-pp go-ASP-FM
'Borzo went to the market'

21. *aɲ naʔ-geʔ bage-ātu[4]-**kan**-a-ɲ*
I now-EMP left-ASP-FM-1SG
'I have left just now'.

3.2. Perfect Aspect markers

-li as simple past marker, it express past times action; short of completed action.

22. *rām sītā-ke nel-**liɟ**-a*
Ram Sita-DAT see-PRF-FM
'Ram saw Sita'.

23. *aɲ kuɽi nel-**liˀ**-e*
I girl see-PRF-FM
'I saw the girl'.

-liɟ and ***-li*** shown above may have used for different purposes like animate, inanimate as discussed by (Ramswami, 1992:100) but in above examples it is not seen. Even transitive and intransitive distinction cannot be done. It needs further research to distinguish them.

-le as past perfect marker (*-leɋ*' is used with transitive verb whereas *-len* is used with intransitive verb. I have shown (Choudhary, 2005) that *–ɋ* is transitive marker and *-n* is intransitive marker in the language which always occurs with aspects).

24. *aɲ am- ke miʔɋ uli ama- **leɋ**- m- a*
I you- ACC one mango give- PRF- 2SG-FM
'I had given a mango to you'.

25. *holā ciyā-mente ka-m hiɲu- **len**- a*
yesterday Q-pp not-2SG come-PRF-FM
'Why did you not come yesterday?'

26. *aɲ ciɽhi naʔ-geʔ ōl-**leɋ**-a-ɲ*
I letter now-EMP write-PRF-FM-1SG
'I have just written a letter'.

[4] *bage-atu* means leave this place, here *bage* means to leave and *-atu* is a bound form, it does not have an independent meaning.

-ɟan as Past-perfect marker (it is non-completive marker usually used for natural phenomenon and inanimate nouns)

27. *daʔ leŋa-**ɟan**-a*
water flow-PRF-FM
'Water has flown'.

28. *ini uku-**ɟan**-a*
he lost-PRF-FM
'He has been non-existent'.

29. *aliŋ coʔ-p-o-**ɟan**-a*
we $kiss_1$-RECP-$kiss_2$- PRF-FM
'We kissed each other'.

-ɟan a past perfective marker although there is no such sense in Ho, in example (30) two sentences has been conjoined to express past perfect and progressive together.

30. *[aŋ upun sirma-hobā-**ɟan**-a] nere paiti-hiɟu-**tan**-a*
I four year-stay-PRF-FM here work-come-PROG-FM
'I had/have been working here for last four years'.

31. *ɟān-o sakam ka-ʔ iyū-**ɟen**-a*
any-EMP leaf not-3sg,inanimate fall-PRF-FM
'Nothings leaf had fallen'.

-keɖ/-kaɖ completive marker used with transitive verb and shows that work has been completed.

32. *aŋ uli ɟom-**keɖ**-a*
I mango eat-PRF-FM
'I ate a mango'.

33. *aŋ-nagin am kuɽi-ko kul-**keɖ**-ko-a*
I-DAT you girl-PL send-PRF-PL-FM
'You sent girls to me'

4. Modals and Mood in Ho

Ho has only one modal *dai* 'can/may' used for various purposes. Some others like *cahī* 'needs' and *lagtīya* 'seem' have been borrowed from Hindi for the purpose of expressing such feelings as for English verbs 'needs' or 'seem'; it is used by some speakers of certain Ho speaking areas.

Mood denotes the relationship of the agent with regard to the kind or manner of action or event. An action or event may be real, intended, demanded or desired. Ho has four different moods which are presented below:

4.1. Indicative Mood

Sentences that express a statement are in the indicative mood. Ho does not have any special morpheme to mark it but semantics of sentence tell us about its mood. It may be the case that finite marker *–a,* is used as mood marker for indicative sentences. For example:

34. *rām duruŋ-ai-a*
Ram sleep- 3SG-FM
'Ram sleeps'

35. *ini lad-ai ɟom-a*
he bread-3SG eat-FM
'He eats bread'.

4.2. Imperative Mood

It demands on the part of the listener to do an action. Imperative in Ho can be formed with any verb by dropping finite marker *-a.* There are *-m/me, -ben* and *-pe* to mark imperative mood as they are second person pronominal clitic also used as agreement marker. For example**:**

36. *gel takā aŋ-ke ama-i-**me***
ten rupee 1SG-ACC give-3SG-you/IMP
'Give me ten rupees'

37. *dub-**ben***
sit-you(two)/IMP
'Sit down'

38. *aŋ sidha seno-rikai-**pe***
I first go-do-you(pl)/IMP
'Let me go first'

4.3. Interrogative Mood

The sentences, which require an answer from listener, are in the interrogative mood. There are interrogative words in Ho to form interrogative sentence as some of it has been presented below. To form yes-no question *ci* 'what' has been used. For example:

39. ***oko*** *kuɽi hiɲu-tan-a*
which girl come-PROG-FM
'Which girl is coming?'

40. ***okoni-a*** *eŋā men-a*
Whose mother be-FM
' Whose mother is?'

41. ***cina*** *am iskul-em seno-tan-a*
Q you school-2SG go-PROG-FM
'Are you going to school?'

3.4. Optative mood

Optative mood indicates the attitude of the speaker; it expresses wishes and there is *dai* 'can/may' for it in Ho to play a role. Irrespective of permission or desire it has been used for the purpose. For example:

3.4.1. Permission

The language has a modal *dai* 'may/can' along with question marker *ci* has been used to ask permission.

42. *aŋ sidha sen-**dai**-e-ci-ɲ*
I first go-can-FM-y/n Q mkr -1sg
'May I go first, please?'

43. *aŋ ayur-te seni[5]-**ci**-ɲ-ben/pe*
I first-pp go- y/n Q mkr -1sg -you
'Let me go first'

44. *aŋ sen-**dai**-e-**ci**-ɲ*
I go-can-FM-y/n Q mkr-1sg
'Can I go?'

[5] Here seno became seni; I mean o became i in influence of i of ci.

3.4.2. Desire

There is a word *dai* 'can/may' in the language to express ones desire; it is also used for seeking permission. Along with *dai, y*es-no question marker '*ci*' has been used to ask or express desire or seek permission.

45. *aŋ* *ama-lo^{w}o* *hiju-**dai**-e-**ci***
I you-with come-can-FM-y/n Q mkr
'May I come with you'?

46. *am-ke* *deŋā-me-ŋ* ***ci***
you-DAT help-2SG-1SG y/n Q mkr
'May I help you'?

3.4.3. Obligative

There is a modal *utar-ka* 'must' to form the obligative sentence in Ho. As far as moral duty is concerned as in (49), there is no such morpheme to add in the verb. It may be because the plural marker *–ko* is there for the purpose.

47. *aŋ-ke* *miəʔ* *cithi* *ōl-**utar-ka***
I-DAT one letter write-must-mood
'I must write a letter'

48. *aŋ-ke* *hiju-**utar-ka***
I-DAT come-must- mood
'I must come'

49. *ale* *kōetan-ko* *kōe* *ema-**ko-**^{w}a*
we beggar-PL alm give-PL-FM
'We should give alms to the beggars"

5. Conclusion

It has been mentioned in the literature that there is close relationship between imperfect aspect and present time and perfect aspect and past time in those languages which has no tense markers i.e. the languages whose verb form is same for all tenses, moods, person and number and only differences in aspect. In Ho too, there is no pure tense marker or in other word verb does not change for tense instead a morpheme called aspect can come to convey time reference. As I know languages like Hindi has tense markers i.e. *tə* for present, *ɵə* for past and *gə* for future.

The Munda languages have agglutinating[6] morphology which does not allow modification in verb form for person, number, gender and any other grammatical items i.e. tense, mood markers. Rather, it has very rich morphology in terms of time and direction marking such as adverbs and deixis. It has many time adverbials as some of it has been given above. It has also four ways distinction in deixis to locate place. Thus Ho has way to mark time reference without having tense marker.

Abbreviations:

PRS = present tense	EMP= Emphatic	3SG = Third person singular
2PL= second person plural	3PL= third person plural	2SG = second person singular
PP = post positions	1SG= first person singular	1PL = first person plural
FM = finite marker	PST= past tense	y/n Q mkr: *yes/no* question marker
ACC=accusative	PROG= progressive,	REFL=verbal reflexive
DAT=dative	RECP=reciprocal	PERF=perfective
ASP=aspect	Q=question marker	IMP=imperative

[6] Agglutinating languages can combine simple words without changing their form to make a new word.

Acknowledgements: thanks to the independent reviewer and the editor of the MKSJ for their valuable comments that helped me in improving the data analysis of this paper and also I learn several other things as well.

References

Choudhary, P.K. (2005). Anaphors and Pronouns in Mundari and Nepali. Unpublished M. phil. thesis, University of Delhi, Delhi.

Comrie, B. (1976). *Aspect.* Cambridge University Press, Cambridge, London, New York.

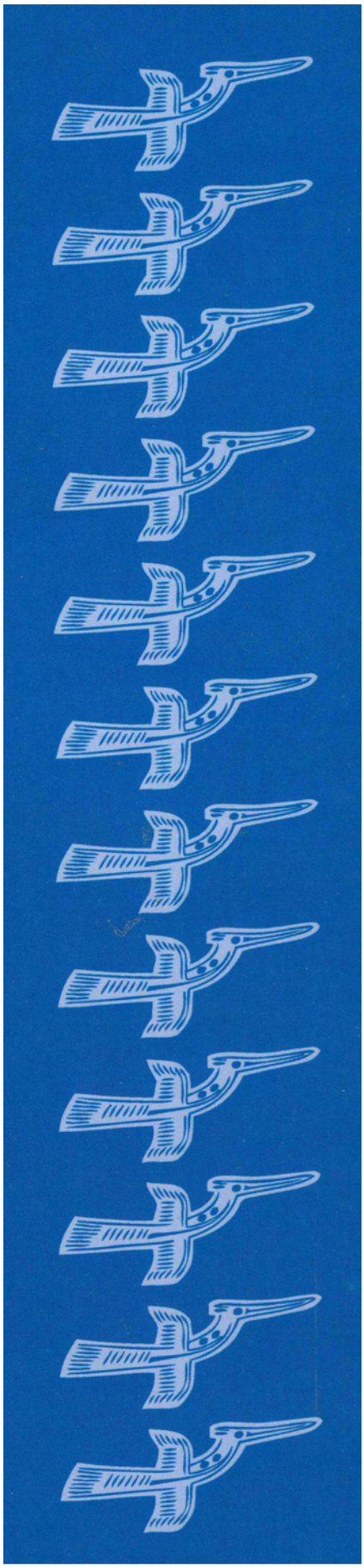
Comrie, B. (1985). *Tense.* Cambridge University Press, Cambridge, London, New York.

Deeney, J. (2002). *Ho Grammar and Vocabulary*. Xavier Ho Publication, Chaibasa.

Ramswami, N. (2007). *Ho Grammar*. CIIL, Mysore.

Mon-Khmer Studies

Volume 42



Authors: Anh-Thư T. NGUYỄN and John C. L. INGRAM
Title: *Perception of prominence patterns in Vietnamese disyllabic words.*
Pages: 89-101
Date received: 22/7/2013
Revised text accepted: 13/11/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

Perception of prominence patterns in Vietnamese disyllabic words

Anh-Thư T. NGUYỄN
Mountain Creek, Sunshine coast, Australia

John C. L. INGRAM
School of Languages & Comparative Cultural Studies, University of Queensland

Abstract
This paper reports an experiment which examined native Vietnamese listeners' perception of prominence pattern on disyllabic compounds, phrases, full reduplications and tone sandhi reduplications. Thirty three Vietnamese listeners listened to disyllabic words and judged which syllable of the two (the first or the second) is more prominent or if both syllables are of equal prominence. The results showed that subjects performed at chance level for most of the disyllabic word types. Both syllables of the full reduplications were perceived as of equal weight while the tone sandhi forms were judged to be more prominent on the second syllable. A follow-up analysis based on the classification of tones into even (lax tones: level and falling) and uneven (tense tones: rising, drop, curve and broken) showed that syllables with uneven tones tended to be heard as more prominent than even tones. The result of this perception experiment, supported by a tone pattern analysis of Vietnamese dictionary entries, suggests that asymmetrical prominence relations in Vietnamese may be realized by means of tones themselves and their tone patterning.
Key words: phonology, prominence, perception
ISO 639-3 language codes: vie

1. Introduction

1.1. Stress, accent and prominence

This paper is the latest in a series of experiments on the status of word-level prominence pattern in Vietnamese. The term 'stress' is used to refer to the perceptual salience at certain places in strings of syllables, but it has several different referents (Kohler, 2008): (a) relative syllable salience in an utterance; this is syllable-, not word-oriented; (b) stress in a word; this is part of the lexical phonology; (c) stressing of words in utterances for various aspects of propositional and expressive meaning, often called 'accent(uation)'. In more recent phonetic research, the terms 'stress' and 'accent' both came to be regarded as being physically manifested, 'stress' being the default realization of a lexically stressed syllable outside focus, as determined by word phonology, 'accent' its realization for contrast under narrow focus, superimposed on its default properties (de Jong, 2004; de Jong and Zawaydeh, 2002). In this paper, 'stress refers to a structural, linguistic property of a word that specifies which syllable in the word is the strongest…the potential docking sites for accent placement' (Sluijter and van Heuven, 1996, p. 2471) in a language that has lexical stress, but it also has physical properties of its own. Vietnamese is a contour tone language that 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).

Stress and accentuation depends crucially on the speaker's ability to make certain syllables more noticeable than others. A syllable which "stands out" in this way is a prominent syllable. There are many ways in which a syllable can be made prominent: experiments have shown that prominence in English is associated with greater length, greater loudness, pitch prominence (i.e. having a pitch level or movement that makes a syllable stand out from its context) and with "full" vowels and diphthongs. Despite the complexity of this set of interrelated factors it seems that the listener simply hears syllables as more prominent or less prominent (Roach, 2002). In Vietnamese, a contour tone language, accentual prominence is shown to be realised by syllable duration, intensity, and full tonal realization (fo height and fo shape) (Do The Dung, 1986; Chaudhary, 1983, Hoang Tue and Hoang Minh, 1975; Gsell, 1980).

NGUYỄN, Anh-Thư and John C. L. INGRAM. 2013. Perception of prominence patterns in Vietnamese disyllabic words. *Mon-Khmer Studies* 42:89-101

A system of culminative word stress is fairly clearly not a phonologically contrastive property of Vietnamese word prosody (Nguyễn, Ingram and Pensalfini, 2008; Nguyễn and Ingram, 2007a). Nevertheless, there are sufficient indications of prominence asymmetries in disyllabic word-level constructions in Vietnamese (compounds of various kinds) to raise the question of the status of rhythmic structures at the level of the foot, which indicates a 'latent' contrast of prominence, that has the potential to become lexicalized. In a recent study, Schiering, Bickel and Hildebrandt (2010) remarked that "Vietnamese provides ample evidence for a genuine stress domain that is preferably disyllabic and maximally trisyllable. Within this domain, stress is realised on the final syllable in the default case. Crucially, this domain is computed irrespective of the morphosyntactic status of its constituent syllables, i.e. stress phonology does not distinguish between a word-level and a phrasal-level of prosodic structure. Metrically, polysyllabic words are thus indistinguishable from other combinations of syllables. Since the most complex structures which are referenced by the rules for iambic rhythm are phrasal, stress may most adequately be attributed to the prosodic domain of the Phonological Phrase."(p.673)

In a series of previous papers we have presented acoustic phonetic evidence for a latent word-level prominence contrast in Vietnamese (Nguyen & Ingram, 2007a, b). In this paper we report perceptual evidence from native listeners that complement the findings of the acoustic phonetic analysis presented in previous studies. Perceptual tests are of course a necessary complement to any acoustic phonetic findings if the phonological significance of phonetic findings is to be established.

We will begin by making some general remarks about the nature of stress at the word and phrasal levels in English and how these two domains intersect in producing contrastive prosodic patterns of compound (word) and phrasal stress. We began by inquiring whether such contrasts applied in the case of a tone language like Vietnamese, and if so, under what speaking conditions.

1.2. Acoustic prominence of compounds/ phrases, reversible coordinative compounds and reduplications.

In English, speakers will generally produce a contrasting stress pattern between compound words and their corresponding phrases, (*hot-dog* vs. *hot dog*), regardless of the speaking task (see Nguyen, Ingram and Pensalini, 2008 for a detailed review). It matters little whether they are asked to name a picture depicting a '*hot-dog'* (*sausage-in-bun*) or a '*hot dog'* (*dog which is hot*), or to pronounce the compound word and its phrase counterpart in such a manner as to disambiguate the two meanings for the listener. Vietnamese speakers, responded quite differently from English speakers when asked to produce comparable Vietnamese compound – phrasal pairs (e.g.: hoa hồng (*rose*: compound vs *pink flower*: phrase)). With respect to the phonetic cues of syllable duration, F0 range, intensity, spectral slope and formant frequencies, no significant difference is detectable between compound-phrase minimal pairs of the type hoa hồng 'rose' vs. hoa hồng 'pink flower' under a normal speech production condition of a picture naming task. Under a maximal contrast condition, where native speakers are encouraged to bring out the difference in meaning between the two forms, the disambiguation strategy that speakers chose most often consists of inserting a phrasal juncture between the two constituents of the phrase. This prosodic juncture has a number of phonetic effects such as a pause after the first syllable and concomitant lengthening, widening of the pitch range and lowering of intensity in the pre-pausal syllable. In a perception experiment, native speakers showed a response bias towards compound forms when asked to distinguish compounds from phrasal structures on the basis of the stimuli obtained from the normal speech production condition of a picture naming task. When confronted with noun phrases lacking a phrasal juncture between the two constituents, listeners performed at chance level when asked to identify the phrasal structures. Performance was slightly better with stimuli from the maximal contrast condition where the two syllables of the noun phrase were separated by a juncture. This shows that listeners do not rely on any phonetic cues for distinguishing compounds and phrases beyond phrasal juncture cues, which were most pronounced under the maximal contrast condition. In line with the findings from the production experiments, we concluded that there is no word-level phonological mechanism for distinguishing disyllabic lexical units from phrases. However, we did not at the time test Vietnamese listeners on their perception of the prominence pattern of the phrasal – compound contrast elicited under the two speaking conditions. It is noted that in Nguyen, Ingram (2007a)'s perception experiment, listeners heard the contextual sentence containing the target item. Two different meanings of the target item were given in the answer sheet: one as a

compound and one as a noun phrase. The subjects' task was to choose the meaning which they judged was expressed by the speaker by circling the letter corresponding to their response in the answer sheet. By contrast, in the perception test presented in this paper, listeners listen to the target items and were asked to judge which syllable is more prominent (clearer, stronger and louder) by circling a number on the answer sheet: 1 for the first syllable, 2 for the second syllable and 3 if both syllables are of equal prominence.

Arguably, a perceptual test is more relevant for the phonological status of any compound – phrasal stress contrast (whether it is made using word or phrasal stress mechanisms) than are any strictly acoustic phonetic differences observed. So, this is one motivation for the perception experiment reported here.

We also conducted other production studies on more carefully phonetically controlled and specialized sets of Vietnamese compounds (Nguyen and Ingram, 2007a, b). These experiments suggested that there was at least a phonetic tendency for the right hand element of a disyllabic compound word to be more prosodically prominent by a number of relevant phonetic measures: greater tonal f0 range, higher intensity, greater duration of the second syllable, and formant measurements indicative of more centralized vowel nuclei (vowel reduction) on the first syllable.

In summary, the results of acoustic comparisons using stimuli that were much better controlled in segmental and tonal features (namely reversible coordinative compounds (e.g., bàn ghế vs. ghế bàn) and reduplications) revealed a clear tendency for second position in disyllabic compounds to show greater acoustic prominence on the second syllable. But was this acoustic right-headed tendency perceptually salient enough to be reliably reported by native listeners? This is the principle requirement for phonological significance and the major question to which the following study was directed.

2. Perceptual experiment

2.1. Stimuli

There were three types of stimuli: (1) disyllabic compounds/ phrases, (2) disyllabic coordinative compounds and (3) disyllabic reduplications.

The *compound/phrase* stimuli were drawn from the items used in a previous production experiment reported in Nguyen and Ingram (2007a). Fifteen pairs of two-syllable compounds and their corresponding phrases consisting of three types of compounds, formed on the basis of their grammatical structures: noun-adjective (NA type), noun-verb (NV type), and noun-noun (NN type) were recorded under two elicitation conditions: the *maximal contrast* condition in which subjects were asked to read minimal sentence-pairs in a natural way, such that listeners could distinguish between the meaning of a compound and its corresponding phrase and a *picture naming* task in which subjects were asked to describe a picture, using a constant carrier sentence. For this perception study, tokens from four Hanoi-dialect female speakers (two speakers from the picture-naming task and two from the minimal pair sentence task) were used. The compound and phrasal items were presented as isolated words or phrases. The test items in isolation were segmented from the contextual sentences using a speech editor (Praat, Boersma, 2001). There were in total 120 tokens: 15 test items x 2 conditions (compound [CP] vs. Noun phrase [NP]) x 4 speakers (2 in picture-naming task vs. 2 in minimal pair sentence task).

The *coordinative compound* stimuli were taken from the items used in a previous production experiment (Nguyen and Ingram, 2007a). Ten minimal pairs of coordinative compounds with reversible syllable positions (e.g., bàn ghế vs. ghế bàn) were elicited in a picture naming task in which subjects were asked to describe a picture, using the target word in a constant carrier sentence. For this perception experiment, tokens from two Southern Saigon-dialect speakers (one male and one female) were used. There were in total 40 tokens: 10 test items x 2 minimal pairs x 2 speakers.

The *reduplication* stimuli were drawn from the items used in a previous acoustic study reported in Nguyen and Ingram (2007b). Full reduplications (copying of full segmental composition and tone: sáng sáng [rising-rising] and tone sandhi reduplications (same segmental composition but with alternate tones: sang sáng [level-rising]) were elicited in an imperative carrier sentence having the same grammatical structure. For this perception study, 15 full reduplications

and 9 tone sandhi reduplications spoken by two Southern Saigon-dialect speakers (one male and one female) were used. There is a constraint on tone harmony in reduplications that the tone of the copying (reduplicant) syllable must be in the same register as that of the base form: a phonological patterning of high tones: level (ngang: high level), rising (sắc: high rising), curve (hỏi: gradual fall-rise) and that of three low tones: falling (huyền: gradual falling), dropping(nặng: low dropping), and broken(ngã: fall glottalised and abrupt rise). In addition, many full reduplications undergo tone sandhi, which is also constrained by the within-register tone harmony. In reduplications with syllables ending in stop consonants, when the first syllable of the word undergoes tone sandhi, the final stop consonant is replaced by its correspondent homorganic nasal. Please see Nguyen and Ingram (2007b) for a complete description of the tone patterns of these reduplications. The target test items were segmented from the carrier sentence using Praat. There were in total 48 tokens: (15 full reduplication + 9 tone sandhi reduplications) x 2 speakers.

The tokens were put in random order each with two immediate repetitions in a block of ten with a gap of about 6 seconds between each item. There were in total 198 tokens: 120 compounds/ phrases + 40 coordinative compounds + 48 reduplications, making up of 20 blocks of stimuli. Please see the appendix for the full list of stimuli.

2.2. Subjects

Thirty three subjects of the Southern dialect (19 females, 14 males) with no known auditory deficiencies, participated in the perception experiment. Since Hanoi dialect is a prescribed national standard in instruction and national broadcasting, listeners from other dialects, particularly the well-educated ones generally have no difficulty understanding Hanoi dialects, but the reverse does not hold. Listeners of this study are all tertiary students and in the age range from 20 to 30 years of age.

2.3. Procedure

In the perception test, subjects listened to the test token twice each and were asked to judge which syllable is more prominent (clearer, stronger and louder) by circling a number on the answer sheet: 1 for the first syllable, 2 for the second syllable and 3 if both syllables are of equal prominence. A sample of the answer sheet is as follow:

	Words	1st syllable	2nd syllable	Equal prominence
1.	Aó dài (long dress)	1	2	3
2.	Hoa hồng (rose)	1	2	3
3.	ghế bàn (furniture)	1	2	3
4.	đỏ đỏ (rather red)	1	2	3
5.	lành lạnh (rather cold)	1	2	3

The experiment was carried out using a laptop computer with loud speakers in a quiet classroom at Can Tho University located in the delta to the southwest of Saigon.

2.4. Analysis

First, agreement among listeners was assessed using a modified form of Cohen's Kappa, which takes into account the amount of agreement that can be expected by chance. Values of *kappa* can range from -1.0 to 1.0, with -1.0 indicating perfect disagreement below chance, 0.0 indicating agreement equal to chance, and 1.0 indicating perfect agreement above chance. A rule of thumb is that a *kappa* of .70 or above indicates adequate interrater agreement. The particular form of Kappa used here is based on Brennan and Prediger (1981). It is suitable for tasks with multiple raters in which the raters are constrained as to how many items they assigned to each category ("fixed marginal"). Calculations were made using the Online Kappa Calculator (Randolph, 2008). Kappa values were determined for each data set (i.e. stimulus types), pooling across all the listeners. The results are reported in table 1.

Then, mixed (fixed and random) effects analysis of variance (ANOVA) models, using the restricted maximum likelihood method (REML) to estimate variance components were used to statistically analyse the data. Tukey post-hoc tests were carried out to determine the significant differences among levels of the main effects when necessary. The use of REML overcomes the

potentially serious deficiency of the ANOVA-based methods which assumed that data are sampled from a random population and normally distributed. REML also avoids bias arising from maximum likelihood estimators in which all fixed effects are known without errors, consequently tend to downwardly bias estimates of variance components. The data analysis was carried out using SAS program. The specific mixed model applied to each set of data is specified at the beginning of each result section corresponding to each data set.

2.5. Results

The results of the three types of stimuli, namely compound [CP]/phrase[PH], coordinative compound and reduplication, were analysed and presented separately.

2.5.1. Agreement among listeners

The kappa statistic was used to assess agreement among listeners. As shown in table 1 below, Kappa values for the three types of stimuli were very low, ranging from 0.05 to 0.24, which are below the proposed cut-off of 0.7 as suggested by Randolph (2008). Therefore, this result suggested that the agreement rate among listeners for the stimulus types is not high. It can be said to be at chance level. Only the agreement rate for the sandhi reduplications is a little better than chance level (= 0.24).

Table 1. The kappa results. MAX: maximal contrast experiment, PIC: picture naming experiment, CP: compound, NP: noun phrase.

Stimulus types		kappa values
CP-NP	MAX-CP	0.14
	MAX-NP	0.17
	PIC-CP	0.05
	PIC-NP	0.08
COORDINATIVE		0.09
REDUPLICATION	full reduplication	0.2
	sandhi reduplication	0.24

2.5.2. Compound/phrase

A four-way mixed effect ANOVA was conducted on the percentage number of responses (the number of listeners over the total listeners). The fixed factors were prominence positions (1st syllable, 2nd syllable and equal prominence), prosodic types(CP vs. PH), elicitation tasks (maximal contrast: MAX vs. picture naming: PIC) and compound types (NA, NN, NV)). The random factors were speakers and items.

Significant effects were obtained for the main factor prominence positions: $F(2, 360)=8.1$, $p<0.001$ only and the interaction effects (prominence positions x compound types: $F(4, 360)=16.31$, $p<0.0001$, prominence positions x elicitation tasks: $F(2, 360)=36.82$, $p<0.0001$, prosodic types x prominence positions x elicitation tasks: $F(2,360)=7.69$), $p<0.0001$, prominence positions x compound types x elicitation tasks: $F(4, 360)=3.6$, $p<0.001$, and particularly the four-way interaction: prominence positions x prosodic types x elicitation tasks x compound types: $F(4, 360)=2.01$), $p<0.01$). The significant main effect of prominence positions needs to be examined in conjunction with the significant four-way interaction (figure 1).

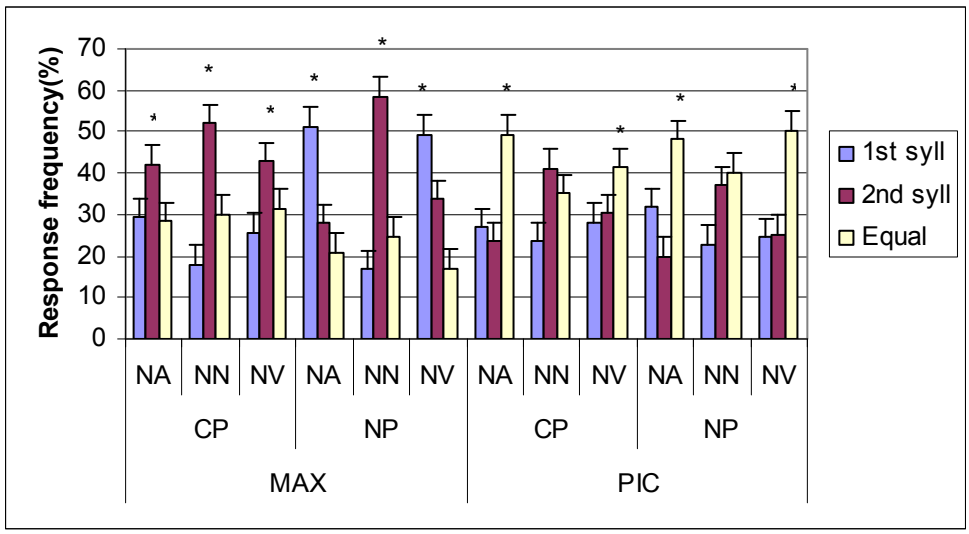


Figure 1. Mean of responses by four-way interaction: prominence positions x prosodic types x elicitation tasks x compound types. (*) = significantly greater than the other two stress positions at $p<0.01$. MAX: maximal contrast task, PIC: picture naming task, CP: compound, NP: noun phrase, NA: noun+adjective, NN: noun+noun, NV: noun+verb.

The four-way interaction showed three main things. First, compounds of all three types (NA, NN, NV) of the MAX elicitation task were generally judged to have more prominence on the second syllables. Second, the noun phrases of the MAX task, which have a juncture between their constituents (NA and NV), were judged to have more prominence on the first syllable, consistent with the acoustic results in Nguyen and Ingram (2007a) that the first syllables were lengthened and had fuller tone realization. Nevertheless, the NN phrases which had no juncture between their constituents were heard to have prominence on the second syllable. Third, both the compounds and phrases from almost all of the compound types the PIC elicitation task were judged to have equal prominence on both syllables. Nevertheless, the agreement rate among listeners for each prominence position of the compounds and phrases of the PIC task is not high (below 50%). It can be said to be at chance level, consistent with the Kappa results and indicating listeners' difficulty in assigning prominence between the syllables of these tokens. This is probably due to the variation in segmental composition and tone of the two constituent syllables of the test tokens.

2.5.3. Coordinative compounds.

A one-way mixed effect ANOVA was conducted on the percentage number of responses (the number of listeners over the total listeners). The fixed factor was prominence positions (1st syllable, 2nd syllable and equal prominence). The random factors were speakers and items. The results showed a significant main effect of prominence positions: $F(2, 120)=6.6$, $p<0.001$.

Examination of the main effect prominence positions (figure 2) showed that more listeners judged the coordinative compounds to have prominence on the second syllables, however, the agreement rate per prominence position is not high and seems to be at chance level (36-38-26), consistent with the Kappa results. This is probably due to the difference in segmental composition and tone of the two constituent syllables of the test tokens that made it hard for the listeners to judge the prominence between the two syllables. Probably reduplicative words which provide segmental feature control may enable a sensitive evaluation of asymmetries of prosodic prominence between adjacent syllables of disyllabic words. Particularly, the existence of a tone sandhi subclass of reduplicated forms enabled us to investigate the role of tone sandhi as a reduction phenomenon occurring on prosodically weak positions. The results on reduplication is presented in the following section.

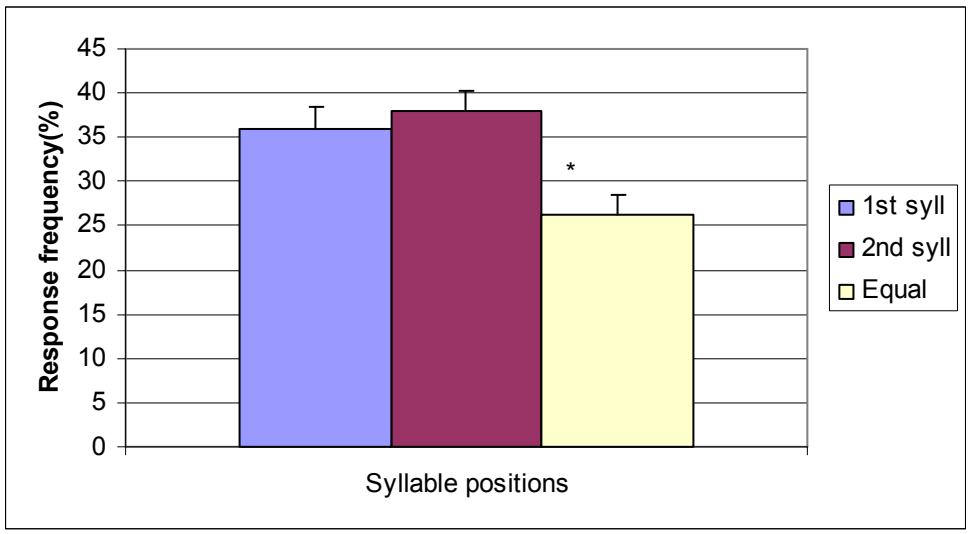


Figure 2: Mean of responses of the three prominence positions. (*) significant at p<0.01.

2.5.4. Reduplication

The full reduplication and tone sandhi stimuli were analyzed separately. A two-way fixed effect ANOVA (3 prominence positions [1st syllable, 2nd syllable and equal prominence] x 2 speakers) was conducted on the percentage number of responses (the number of listeners over the total listeners) of each of the two data sets (full reduplication and tone sandhi). The results of the ANOVA of the full reduplication data set showed a significant effect of prominence positions: $F(2, 90)=337$, $p<0.0001$ and the interaction of prominence positions x speakers: $F(2, 90)=12.63$, $p<0.0001$, while there is no significant effect for speakers: $F(1, 90)=0$, $p=1$ ns. Similar significant effects were found for the ANOVA of the tone sandhi data set (prominence positions: $F(2, 54)=233.9$, $p<0.0001$, interaction of prominence positions x speakers: $F(2, 54)=3.98$, $p<0.05$ and speakers: $F(1, 54)=0$, $p=0.9$ ns.).

Examination of the interaction effect of prominence position x speakers of the full reduplication data set (figure 3) showed that the full reduplications were significantly judged to have equal prominence on both syllables (68% for speaker 1 and 57% for speaker 2).

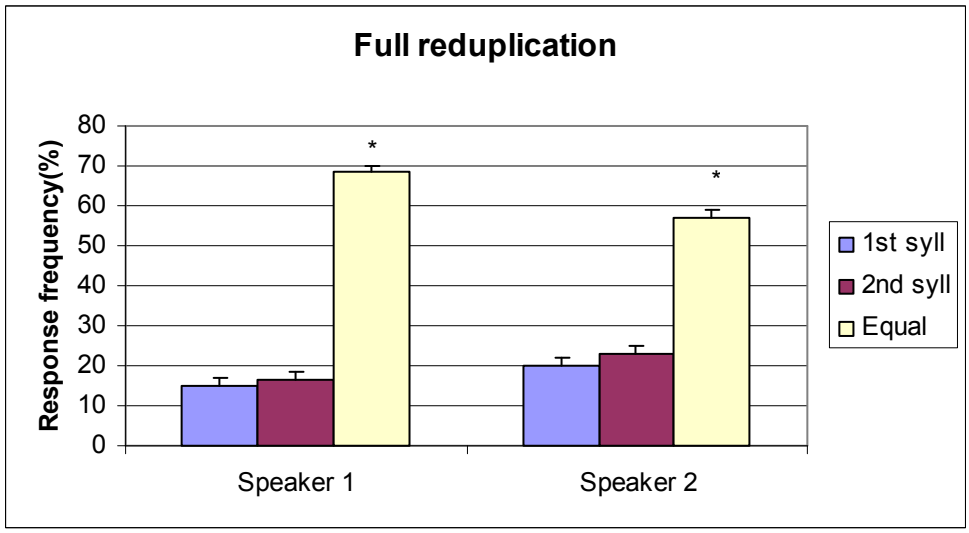


Figure 3: Mean of responses by two-way interaction: prominence positions x speakers. (*) significantly greater than the other two prominence positions at p<0.01.

Examination of the interaction effect of prominence position x speakers of the tone sandhi reduplication data set (figure 4) showed that the tone sandhi forms were significantly judged to have more prominence on the second syllable (64% for speaker 1 and 67% for speaker 2).

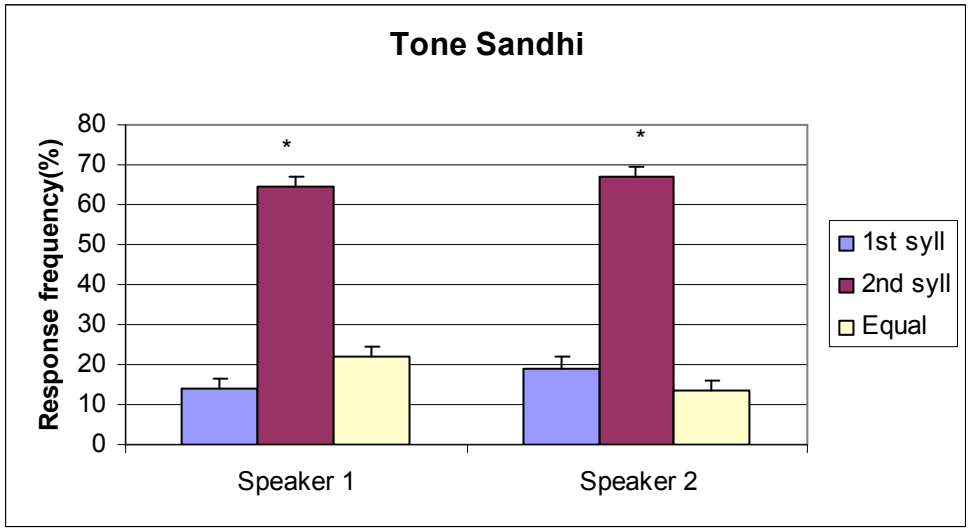


Figure 4: Mean of responses by two-way interaction: prominence positions x speakers. (*) significantly greater than the other two prominence positions at $p<0.01$.

2.6. Discussion

The perceptual result of the compound/phrase and coordinative compound tokens is less robust. Only the perceptual patterns of the compounds and phrases of the maximal contrast task were somewhat consistent with the acoustic cues as presented in Nguyen and Ingram (2007a). That is, the first syllable of the NA and NV phrases which were accompanied by a juncture were heard to be more prominent and compounds of all three types of compound types (NA, NN, NV) of the MAX task were judged to have more prominence on the second syllables, consistent with previous researchers' observations: *weak-strong* for compounds 'with weak stress on their first base/syllable' (Thompson 1965, Trần Hương Mai 1969, Ngô 1984) and *strong-weak* for noun phrases (Thompson 1965: 121). Nevertheless, this contrastive prominence pattern between compounds and phrases only applies under the maximal contrast condition. The judgment of the compounds and phrases of the picture naming task and that of the coordinative compound were at chance level. This is probably due to the variation in segmental compositions and tones of the two constituent syllables of the test tokens and indicates that only reduplicative words which provide segmental feature control enable a sensitive evaluation of asymmetries of prosodic prominence between adjacent syllables of disyllabic words in Vietnamese tonal language.

The acoustic parameters examined in previous study (Nguyen and Ingram, 2007b) suggest that the second syllable of a Vietnamese reduplication is more acoustically prominent than the first syllable. This is realized by longer duration, fuller vowel, lower spectral tilt, larger tone range and more fully realized tone shape, suggesting that if there is a prominence pattern in these Vietnamese disyllabic reduplications, it will be right-headed. The results of this perception study showed that the full reduplications were judged to have equal prominence on both syllables, implying that prosodic asymmetry at the level of the disyllabic word 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. Nevertheless, the tone sandhi reduplications were consistently judged to be right headed i.e., the reduplicated syllable with a tone sandhi was heard to be less prominent than the base syllable, consistent with the acoustic results that tone sandhi is accompanied by vowel reduction and less articulatory effort (spectral tilt) (Nguyen and Ingram, 2007b). This is consistent with the classification of Vietnamese tones on the basis of laxness and tenseness: level and falling (the sandhi tones) as laxness and other tones (curve, rising, drop and broken) as tenseness (Thompson 1965: 40-41; Đoàn Thiện Thuật 1977: 142) and consistent with traditional classification of tones in verse and folk rhymes: level and falling as *even* (bằng), and other tones as *uneven* (trắc).

2.7. Tone patterning: a follow-up analysis.

In light of the result of the tone sandhi data set in which syllables with the sandhi tones (level and falling- the even tone) were heard to be less prominent than the same syllables with uneven tones, consistent with the classification of Vietnamese tones on the basic of laxness and tenseness (Thompson 1965; Đoàn Thiện Thuật 1977) and traditional classification of tones in verse and folk rhymes: even and uneven, the perceptual results of coordinative compound and compound/phrases elicited in the picture naming task were reanalyzed using tone pattern as a variable. Syllables of these two stimulus types were labelled with either even or uneven on the basis of their tones: level and falling as even and rising, drop, curve and broken as uneven. For example,

1. aó dài: rising –level: uneven-even
2. cỏ hoa: curve-level: uneven-even
3. bàn ghế: falling-rising: even-uneven
4. người làm: falling-falling: even-even
5. cà chua: falling-level: even-even

Then a mixed effect two-way ANOVAs were conducted on the percentage number of responses of the coordinative compound data set. It is predicted that syllables with uneven tones would be heard to be more prominent than the even ones. The fixed effects were prominence positions (1st syllable., 2nd syllable and equal prominence) and tone patterns (even-uneven vs. uneven-even). The random factors were speakers and items. The results showed a significant main effect of prominence positions ($F(2, 120)=7.8$, $p<0.001$) and an interaction effect: prominence positions x tone patterns only ($F(2, 120)=14.9$, $p<0.0001$). Examination of the interaction effect (figure 5) showed that when the uneven tone was in the first position, the first syllables were heard to be more prominent. By contrast, when the uneven tone was in the second position, the second syllables were judged to be stronger, consistent with the prediction. There are two words which were excluded from the analysis because there is not enough tokens in their category for the statistical test: vở sách (broken-rising: uneven-uneven) and xuồng ghe (falling-level: even-even). An item analysis showed that the word vở sách was heard to be more prominent on the second syllable while the word xuồng ghe had more prominent rating on the first syllable.

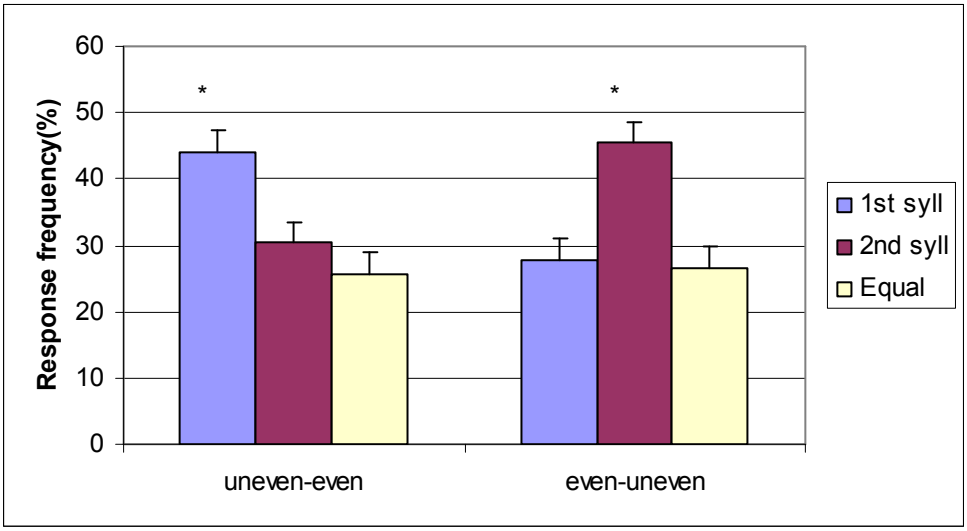


Figure 5: Mean of responses by two-way interaction: prominence positions x tone patterns. (*): significantly greater than the other two prominence positions at $p<0.01$.

The results of the three-way mixed effect ANOVA (3 prominence positions [1st syllable, 2nd syllable and equal prominence] x 2 prosodic types [CP vs. PH] x 3 tone patterns [even-uneven, uneven-even, even-even]) on the compounds/phrases of the picture naming data set showed a significant main effect of prominence positions($F(2, 180)=46$, $p<0.0001$) and an interaction effect: prominence positions x tone patterns only ($F(6, 180)=12.66$, $p<0.0001$). Examination of the interaction effect showed that there is a marginal tendency for the uneven tones to be perceived as

more prominent than the even ones though the compound/phrase tokens in this data set tend to have more equal prominence rating, consistent with the analysis in section 2.5.2.

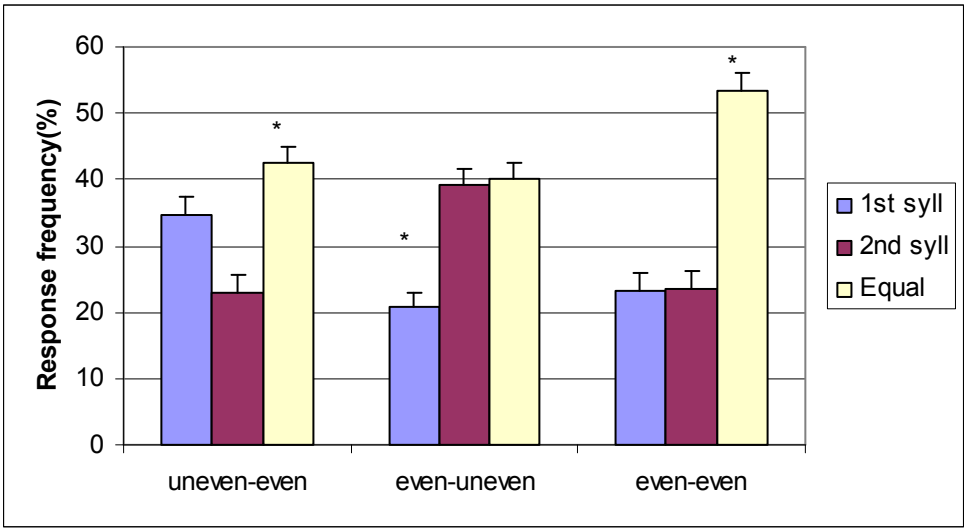


Figure 6: Mean of responses by two-way interaction: prominence positions x tone patterns. (*): significantly greater than the other two stress positions at $p<0.01$.

In general, the perceptual result of the follow-up analysis by tone pattern is consistent with the result of tone sandhi reduplication: syllables with the uneven tones were heard to be more prominent than those with even or sandhi tones, supporting the classification of Vietnamese tones on the basis of laxness and tenseness (Thompson 1965; Đoàn Thiện Thuật 1977) and the traditional classification of tones in Vietnamese verse and folk rhymes: even and uneven. These results together with the finding that even though full reduplications have asymmetrical phonetic cues they were heard to have equal weights on both syllables because they are of the same tones suggest that asymmetrical prominence relations in Vietnamese – a tone language- may be realized by means of the tones themselves and their tone patterning.

To further investigate the tone patterning in the lexicon inventory in Vietnamese language, the tone patterns in disyllable word/compound entries in a Vietnamese dictionary (Nguyễn, Hồ and Nguyễn , 2005) was counted. The result showed that the tonal patterns of the two even/ lax tones, namely Level (ngang) and Falling (huyen), and other uneven/ tense tones accounted for 49% of all the disyllabic entries. As shown in the figure 7 below, the pattern of Level+ other tones accounted for 32% and the pattern of Falling + other tones accounted for 17% while the rest of the tone patterns accounted for only 51%. It is also found that among 41850 entries in the dictionary, 65% were disyllabic, only 20% were monosyllabic and 15% were loanwords and idioms. This indicates that there is a tendency for Vietnamese words to have two syllables (disyllabic).

In summary, the perceptual result of the follow-up analysis by tone pattern is well-supported by the distribution of tone patterns in lexicon inventory. This is consistent with a relation between prominence and tone in register tone languages: stressed position attracts high tone, and high tone attracts stress (Liberman 1975, Selkirk 1984, 1995, Goldsmith 1987); low tone attracts non-prominence and non-prominent position attracts low tone (de Lacy 1999). Particularly, it has been recently found that in Mandarin Chinese (Qu, 2009), there is a relation between high register and prominence, between low register and non-prominence; between rising contour and prominence and between level tone and non-prominence.

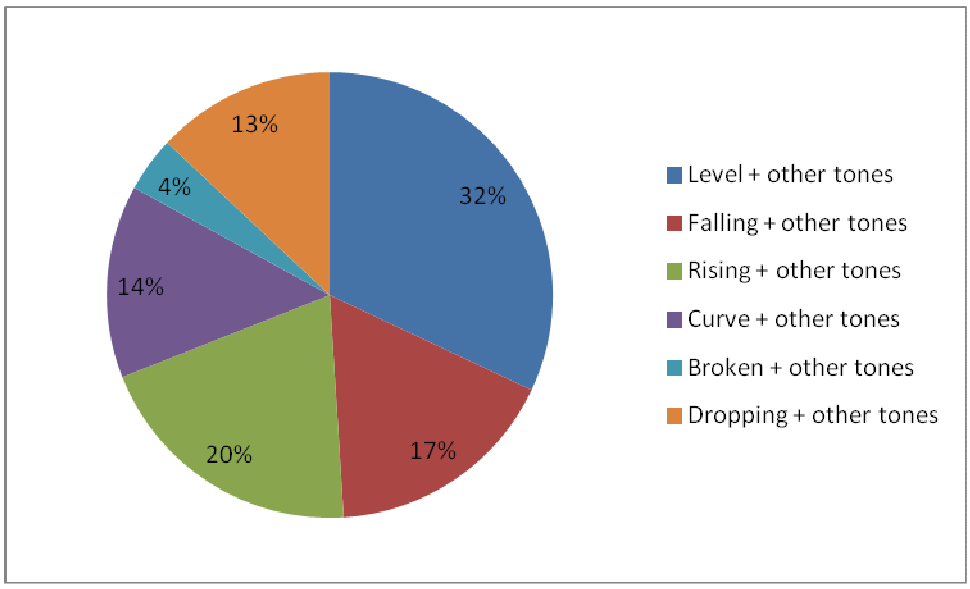


Figure 7: Distribution of tone patterns in Vietnamese disyllabic dictionary entries

3. Conclusion

The results of the perception experiment show that subjects performed at chance level for most of the disyllabic word types and only the tone sandhi forms (with constant segmental makeup but with an alternation of tones) were judged to have more prominence on the second syllable. Particularly, the results showed that the full reduplications were judged to have equal prominence on both syllables. This perception result does not support the acoustic asymmetry of disyllabic words found in previous studies (Nguyen and Ingram, 2007a b). In other words, the acoustic right-headed tendency was not perceptually salient enough to be reliably reported by native listeners. This contradicts findings showing that listeners in other languages, such as English, can hear stress contrasts regardless of syllable segmental composition (Fry 1958) and implies that prosodic asymmetry at the level of the disyllabic word 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.

The perceptual result of the follow-up analysis by tone pattern is consistent with the result of tone sandhi reduplication: syllables with the uneven tones were heard to be more prominent than those with even or sandhi tones, supporting the classification of Vietnamese tones on the basis of laxness and tenseness (Thompson 1965; Đoàn Thiện Thuật 1977) and the traditional classification of tones in Vietnamese verse and folk rhymes: even and uneven. These results are supported by a tone pattern analysis of Vietnamese dictionary entries, suggesting that asymmetrical prominence relations in Vietnamese may be realized by means of tones themselves and their tone patterning, This is consistent with a relation between prominence and tone in register tone languages: (Liberman 1975, Selkirk 1984, 1995, Goldsmith 1987; de Lacy 1999) and recent analysis in Mandarin Chinese (Qu, 2009). Nevertheless, this tone and prominence interaction in Vietnamese needs to be further investigated in future studies.

References

Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glot International* 5:9/10, 341–345.

Brennan, R. & Prediger, D. 1981. Coefficient Kappa: Some uses, misuses and alternatives. *Educational and Psychological Measurement* 41.687-699.

Chaudhary, C. C. 1983. Word stress in Vietnamese: A preliminary investigation. *Indian Linguistics* 44.1-10.

Chen, Matthew Y. 2000. *Tone sandhi: Patterns across Chinese dialects*. Cambridge: Cambridge University Press.

De Lacy, Paul. 1999. Tone and prominent. Ms, University of Massachusetts, Amherst,

de Jong, K. 2004. Stress, lexical focus, and segmental focus in English: patterns of variation in vowel duration. *Journal of Phonetics*, 32.493–516.

de Jong, K.; Zawaydeh, B. 2002. Comparing stress, lexical focus, and segmental focus: patterns of variation in Arabic vowel duration. *Journal of Phonetics* 30.53–75.

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, Paris.

Đoàn, Thiện Thuật. 1977. *Ngữ âm tiếng Việt (Vietnamese phonology)*. Hà Nội, Nhà xuất bản Giáo dục và trung học chuyên nghiệp.

Fry, D. B. 1958. Experiments in the perception of stress. *Language and Speech* 1.126–152.

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.

Goldsmith, J. 1987. "Tone and accent, and getting the two together." *Berkely Linguistics Society* 13.88–104.

Hoàng, T. and Hoàng M. 1975. *Remarques Sur la Structure Phonologique du Vietnamien. Essais Linguistiques*, Etudes Vietnamiennes 40, Hanoi.

Kohler, K.J. 2008. The Perception of Prominence Patterns. *Phonetica* 65.257–269

Liberman, M. 1975. *The intonational system of English*. Cambridge: MIT Press.

Lienard, J. & Di Benedetto, M. 1999. Effect of vocal effort on spectral properties of vowels. *Journal of the Acoustical Society of America*, 106.411-423

Ngô, Thanh Nhàn 1984. The syllabeme and pattern of word formation in Vietnamese. Ph.D dissertation. New York University.

Nguyễn, Kim Thản, Hồ , Hải Thuy, and Nguyễn Đức Dương. *2005. Tự Điển Tiếng Việt (Vietnamese Dictionary)*. Saigon Culture Publishing House.

Nguyễn, Anh-Thư T. & John C. L. Ingram. 2007a. Acoustic and perceptual cues for compound - phrasal contrasts in Vietnamese. *Journal of the Acoustical Society of America* 112(3).1746–1757.

Nguyễn, Anh-Thư T. & John C. L. Ingram. 2007b. Stress and tone Sandhi in Vietnamese reduplications. *Mon-Khmer Studies*, 37.15-40.

Nguyễn, Anh-Thư T., Ingram, J., and Pensalfini, R. 2008. Prosodic transfer in Vietnamese acquisition of English contrastive stress patterns. *Journal of Phonetics*, 36.158-190

Qu, Chen. 2009. Prosodic licensing and Mandarin tone 2 sandhi - Interaction between tone and Stress. *Toronto Working Papers in Linguistics* (TWPL), Volume 34

Randolph, J. 2008. Online Kappa Calculator. http://justus.randolph.name/kappa (date last view September 13th, 2013)

Roach, P. 2002. A Little Encyclopaedia of Phonetics: http://www.personal.rdg.ac.uk/ ~llsroach/encyc.pdf (originally Roach, P. 1992. *Introducing Phonetics*. London: Penguin)

Rose, Philip.1990. Acoustics and phonology of complex tone sandhi. *Phonetica* 47.1–35.

Schiering, R., B. Bickel, & K. Hildebrandt. 2010. The prosodic word is not universal, but emergent. *Journal of Linguistics* 46.657-709

Selkirk, E. O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge: MIT Press.

Selkirk, E. O. 1995. "The prosodic structure of function words." In Beckman, J. Dickey, L. W. & Urvanczyk, S. (eds.) *Papers in Optimality Theory*. Amherst, MA: GLSA, pp. 439–470.

Sluijter, A.M.C.; van Heuven, V.J. 1996. Spectral balance as an acoustic correlate of linguistic stress. *Journal of the Acoustical Society of America* 100.2471–2485.

Thompson, Laurence C. 1965. *A Vietnamese reference grammar* Honolulu: University of Hawaii Press.

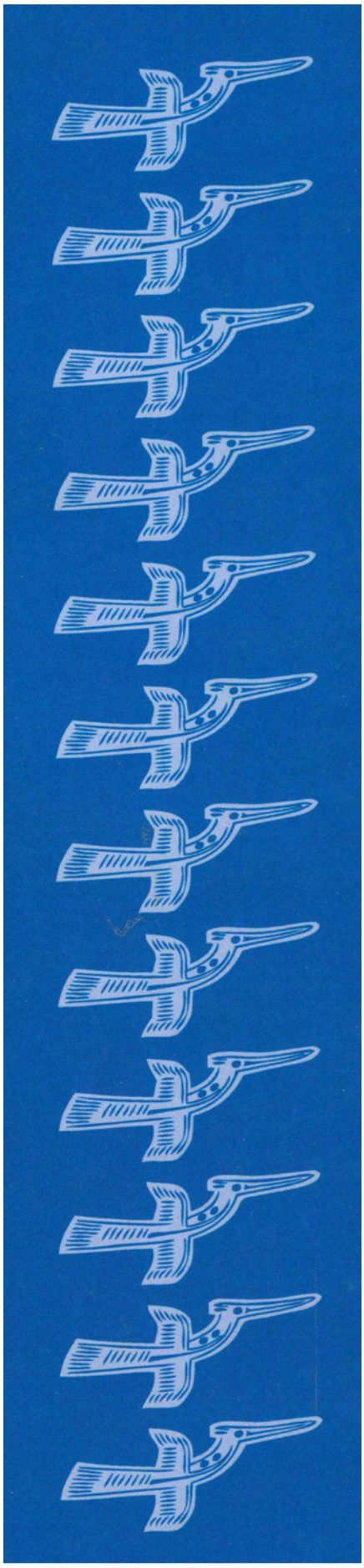
Trần, Hương Mai. 1969. Stress, tone and intonation in South Vietnamese. Ph.D dissertation, Australian National University.

Appendix: List of stimuli

Compounds/phrases	Reversible coordinative compounds	Reduplications
Noun+Verb 6. Người ở/ Người ở 7. Người làm/ Người làm 8. Thợ rèn/ Thợ rèn 9. Gà đẻ/ Gà đẻ 10. bò cày/ bò cày *Noun+ Adjective* 11. Aó dài/ Aó dài 12. Hoa hồng/ Hoa hồng 13. Cá mập/ Cá mập 14. Cà chua/ Cà chua 15. Bánh dày/ Bánh dày *Noun+Noun* 16. Bạn trai/ Bạn trai 17. Bạn gái/ Bạn gái 18. Bàn giấy/ Bàn giấy 19. Nhà đá/ Nhà đá 20. Chân vịt/Chân vịt	1. áo quần / quần áo 2. bàn ghế / ghế bàn 3. chồng vợ/ vợ chồng 4. cỏ hoa/ hoa cỏ 5. cửa nhà/ nhà cửa 6. đậu mè/ mè đậu 7. gái trai/ trai gái 8. gà vịt/ vịt gà 9. ghe xuồng/ xuồng ghe 10. sách vở/vở sách	*Full reduplications* 1 thơm thơm 2 đen đen 3 xanh xanh 4 bằng bằng 5 hồng hồng 6 vàng vàng 7 nóng nóng 8 sáng sáng 9 sát sát 10 đẹp đẹp 11 rộng rộng 12 lạnh lạnh 13 thẳng thẳng 14 đỏ đỏ 15 nhỏ nhỏ *Sandhi reduplications* 16 đèm đẹp 17 đo đỏ 18 lành lạnh 19 nong nóng 20 nho nhỏ 21 sang sáng 22 san sát 23 rồng rộng 24 thăng thẳng

Mon-Khmer Studies

Volume 42



Authors: Peter NORQUEST
Title: A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic Evidence
Pages: 102-126
Date received: 14/10/2013
Revised Text Accepted: 5/12/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A Revised Inventory of Proto Austronesian Consonants: Kra-Dai and Austroasiatic Evidence[1]

Peter Norquest
The University of Arizona

Abstract
This paper examines the Kra-Dai and Austroasiatic evidence for a revised Proto Austronesian consonant inventory that includes three new phonemic distinctions (*f, *g, and *l̩) and an expanded domain for two others (*t̪ and *c). Corroborative evidence is found in Kra-Dai, strengthening the hypothesis of a genetic relationship (Austro-Tai) between Kra-Dai and Austronesian. Evidence from Austroasiatic is weaker, but still suggestive of a non-accidental relationship with Austro-Tai, through either genetic relatedness or contact.
Keywords: Austronesian, Kra-Dai, Austroasiatic, reconstruction, phylogeny
ISO 639-3 language codes: nia, otd, bhp, hvn, nfa, lic, onb

0.0 Introduction

In Norquest & Downey (2013, forthcoming) it is argued that a set of phonological distinctions have been preserved in certain subgroups and languages throughout the greater Austronesian-speaking world. Besides the Formosan languages, these also include: the North Sarawak, Sabahan, and Northwest Barito groups on Borneo; the languages of the Philippines in the northwest; Nias in the southwest; Oceanic in the northeast; and various languages of Nusa Tenggara in the southeast.

This evidence allows for the reconstruction of several additional PAn consonants, as well as a more specific phonetic interpretation of existing phonemes (*C, *j, *z, *N, *S, *R, and *g). The present analysis is supported to various degrees by evidence from Kra-Dai families, allowing refinements of reconstruction in both phyla and ultimately supporting the hypothesis of a genetic relationship between the two.

Although there is some debate about the total number of reconstructible Proto-Austronesian consonants as well as their phonetic interpretation (see Wolff 2010 for discussion), the consensus inventory of PAn consonants cited in most references on Austronesian phonology is the following (Adelaar & Himmelmann 2005:5):

Table 1: The Proto Austronesian Consonant Inventory

p	t	C	c	k	q	ʔ
b	d	j	z	g		
	s		S			h
m	n		ɲ	ŋ		
	l		N			
w	r		y		R	

Of the phonemes above, *j, *z, *S, and *R are firmly established, but their phonetic interpretation is still debated. Wolff argues that *C and *ɲ are allophones of *t and *N respectively, and do not need to be reconstructed for PAn (Wolff 2010: 32). *g, *r, and *c are of comparatively low-frequency and therefore more controversial (for discussion of each of these, see Blust 2009);

[1] I wish to acknowledge help from NSF, award 1030031, for supporting much of the research which made this presentation possible. Thanks to Robert Blust for making his Austronesian Comparative Dictionary (Blust 1995) available online, without which this research would have been much more difficult. Thanks to Weera Ostapirat, Joe Pittayaporn and Paul Sidwell for fruitful discussion about this material at the SEALS22 conference, as well as independent discussion with Andrew Hsiu and Sean Downey; thanks also to the audience at ICSTLL-46 for constructive comments on an initial presentation of this paper. Finally, thanks to Doug Cooper and the staff at the Center for Research on Computational Linguistics for ongoing work in developing tools to further this research. All mistakes are my own.

NORQUEST, Peter. 2013. A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42:102-126

*c also has a restricted geographic distribution in western Indonesia, and is limited largely to frozen monosyllabic morphemes and their prefixes (Norquest & Downey 2013).

This paper follows the proposal originally explicated in its full form by Benedict (1975), in which he argues that Austronesian and Kra-Dai (Tai-Kadai) are genetically related (Austro-Tai[2]). While a definitive argument for this hypothesis remains premature until a full reconstruction of Kra-Dai has been completed, it can nevertheless be stated with confidence that there is a non-accidental relationship between these two language phyla which must be due to either genetic relationship or contact and that the former appears to be the more likely explanation.

The outline of this paper is the following: section 1 presents the evidence for a group of new phomemic distinctions in Proto Austronesian (PAn) which in turn informs the revision of the Proto Austronesian forms given below (Revised Austronesian, or RAn). Section 2 comprises the main body of this paper in which RAn forms are compared with putative cognates from two of the four branches of Kra-Dai: Hlai and Be-Tai. Section 2.1 discusses the distinction between *p and *f, 2.3 retroflex consonants, 2.4 palatal consonants, and 2.5 uvular consonants; section 3 examines the more tentative question of what evidence – if any – is to be found in Austroasiatic; section 4 concludes.

1.0 Additional phonemic distinctions in Austronesian

Correspondences between certain key languages and subgroups are presented in table 2 (for additional witnesses, see Norquest & Downey 2013, forthcoming). The languages used in these tables are the following: Nias is a language of the Barrier Islands off of the northwest coast of Sumatra; Dohoi is a member of the northwest Barito group of languages on Borneo; Bimanese, Proto Sumba, and the closely related Proto Hawu-Dhao are all languages of western Nusa Tenggara; and Proto Western Oceanic (PWOc) is a subgroup of Proto Oceanic (POc) (see Ross 1988):

Table 2: Five additional phonemic distinctions in PAn

PAn	RAn	Nias[3]	Dohoi[4]	Bima	PSumba	PHD	PWOc
*p	*p	f-	-p-	p	*p	*p	*p
*p	*f	[β-]	-hp-	f	*p	*0	*β
*t	*t	t-	-t-	t	*t	*t	*t
*C	*ʈ	[d-]	-ht-	d	*t	*ɖ	*t
*s	*s	[z-]	-s-	s	*s	*s	*z
*s	*c	s-	-s-	c	*ç	*c	*s
*k	*k	k-	-k-	k	*k	*k	*k
*k	*g	[g-], -ʔ-	-hk-	h	*ɣ	*0-, *-ʔ-	*ɣ
*g	*ɢ	(g)	(g)	g	*g	*ɠ	*g
*l	*l	l	-ɾ-	l	*l	*l	*l
*l	*ɭ	l	-ɾ-	r	*l	*ɾ	*l

Table 2 provides the correspondences for five additional distinctions that we use as evidence to reconstruct an additional three novel PAn phonemes (*f, *g, and *ɭ) and expand the scope of two more (*ʈ, which has until now only been distinguished in certain Formosan languages, and *c, the evidence for which has been restricted to a handful of WMP languages); it also includes reflexes of the voiced uvular stop *ɢ (traditional PAn *g) for comparison. Revised Austronesian (RAn) reconstructions are placed to the right of traditional Proto Austronesian ones. For more on these five phonemes, see Norquest & Downey (2013):

[2] Benedict originally spelled this 'Austro-Thai'; it is referred to here as 'Austro-Tai' in keeping with more recent convention, reserving the spelling 'Thai' for a political designation and using 'Tai' as a linguistic designation.

[3] The distinctions in the Nias initials have gone unrecognized in the past because they depend on environment. Lase (2011: xxiv-xxv) describes these as "initial mutations", where the initial of a word undergoes a change in the middle or at the end of the sentence (i.e. in an intervocalic position within a phrase). "Mutated" forms are given in square brackets.

[4] Distinctions in Dohoi occur only in intervocalic position.

The revised PAn consonant inventory proposed here is thus the following (traditional phonemes with revised phonetic interpretations are placed in parentheses):

Table 3: The revised Proto Austronesian consonant inventory

p	t	ʈ (C)	c	k	q	ʔ
b	d	ɖ (j)	ɟ (z)	g	ɢ (g)	
f	s		ç (S)			h
m	n		ɲ	ŋ		
	l	ɭ	ʎ (N)			
w	r		j (y)		ʁ (R)	

Examples of traditional *p, *t, *s, and *k are given in (4) in initial position, and in (5) in medial position. Traditional Proto Austronesian (PAn) forms are given on the far left, with Revised Austronesian (RAn) forms following; when a PAn form lacks Formosan reflexes (i.e. when it can be reconstructed to the level of Proto Malayo-Polynesian (PMP)), it is placed in brackets (PMP *h is projected back to PAn *S, of which it is the regular reflex). Forms with unexpected reflexes in all examples are placed in parentheses. Since the literature has tended to focus on exceptional rather than regular correspondences, it is more difficult to find examples with traditional phonemes than it is those with the novel phonemes proposed in this paper:

(1) Examples of initial *p, *t, *s, and *k

Gloss	PAn	RAn	Nias	Bima	PSumba	PHD	PWOc
seven	*pitu	*piʈu	fitu	pidu	*pitu	*piɖu	---
three	*telu	*təlu	təlu	tolu	*t[ə]lu	*təlu	---
elbow	*siku	*sigu	[z]iʔu	(cihu)	*siɣu	*siʔu	---
scratch	*kaRaw	*kaʁaw	---	kao	*kaʔu	*kao	---

(2) Examples of medial *p, *t, *s, and *k

Gloss	PAn	RAn	Dohoi	Bima	PSumba	PHD	PWOc
four	*Sepat	*çə[p]ac	(ohpa)	upa	*pat-ə	*əpa	(*βati)
calf	*beties	*bətiəs	boti	wisi	*βici	---	---
sea	*tasik	*ʈasik	---	dasi	*tasik	*ɖasi	---
open	[*bukas]	*ɓukas	---	---	*ɓukas	*mboka	---

Examples revised to show PAn *f, *ʈ, *c, and *g in initial and medial positions are given in (3) and (4) below:

(3) Examples of initial *f, *ʈ, *c, and *g

Gloss	PAn	RAn	Nias	Bima	PSumba	PHD	PWOc
turtle	*peŋu	*fəŋu	[β]ənu	fonu	---	*əŋu	*βoŋu
feces	*CaqiS	*ʈaqiç	[d]ai	(taʔi)	*tai	*ɖei	---
nine	*siwa	*ciwa	siwa	ciwi	*çiwa	*ceo	---
tree, wood	*kaSiw	*gaçiw	[g]eu	hadʒu	*ɣaju	*aʃu	*ɣaju

(4) Examples of medial *f, *ʈ, *c, and *g

Gloss	PAn	RAn	Dohoi	Bima	PSumba	PHD	PWOc
dream	[*S-in-ipi]	*ç-in-ifi	nuhpi	nifi	*nipi	*nii	*m-niβi
die	*m-aCay	*m-aʈaj	mahtoi	made	*mate	*maɖe	---
one	*isa	*ica	ihco	ica	*içа	*əci	---
1sg	*i-aku	*i-agu	ahku	n-ahu	*jauwa	*ɟaʔa	*[i]au

NORQUEST, Peter. 2013. A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42:102-126

Outside of Nusa Tenggara, evidence for the distinction between *l and *ɭ has been preserved best in Proto Philippines (Paz 1981). Examples of alveolar laterals are given in (5) and retroflex laterals in (6):

(5) Examples of *l

Gloss	PAn	RAn	PPhilippines	Bima	PHD
buy	*beli	*bəli	*bəlíʔ	weli	*βəli
five	*lima	*lima	*limáʔ	lima	*ləmi

(6) Examples of *ɭ

Gloss	PAn	RAn	PPhilippines	Bima	PHD
moon	*bulaN	*buɭaʌ	*búɭan	wura	*βəru
eight	*walu	*waɭu	*waɭúʔ	waru	*aru

2.0 Kra-Dai Comparisons

Having provided the evidence for the above distinctions in PAn, we can now turn our attention toward Kra-Dai, with an eye toward establishing whether or not evidence from the two phyla is mutually corroborative. Two of the four primary Kra-Dai branches have been used for comparison with our Revised Austronesian forms: Be-Tai and Hlai. The branches cited include Proto North Tai (PNT) and Proto Central-Southwest Tai (PC/SWT) (data for both is taken from Pittayaporn 2009), Proto Be (PBe, Norquest ms), and Proto Hlai (PHlai, Norquest 2007). Proto Kam-Sui (including Biao and Lakkja) and Proto-Kra reconstruction is currently ongoing, and these families are not included in the present paper. It is possible that Be-Tai and Hlai form a higher node within the Kra-Dai phylum; if this is the case, it should be understood that 'Kra-Dai' below technically refers to this subgrouping (Tai-Hlai), although the majority of what is discussed appears to apply to the Kam-Sui and Kra branches as well.

The comparisons below have been selected because they appear reasonable on both semantic and phonological grounds. Although a few of them are novel, most of them have been suggested elsewhere in the literature. In the majority of cases, Kra-Dai forms correspond with the final syllable in an Austronesian word. This is, of course, technically ambiguous in the case of reduplicants and roots, and the only possibility in the case of monosyllabic words.

(7) Examples of RAn reduplicants, roots, and monosyllables

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
mouth	*baqbaq	---	*paak	*paak	*paak
chest, liver	*dəbdəb	---	*t[ə]p	*tap	*tap
hold in fist	*ɢəmɢəm	---	---	*kam	*kam
slap	*-pik, *-bik	*p^{h}i:k	---	---	---
fall	*-tuq	*t^{h}ok	*tɔk	*tok	*tok
bite	*-kət	---	---	*kat	*kat
eat, feed	*gaŋ	*k^{h}ən	*kən	*kɯŋ	*kin
to skin, peel	*-ʎiʈ	*hli:t	---	---	---

Note that the regular outcome for plain voiced initials in all Kra-Dai branches was devoicing. Although it appears to have occurred very early in both Proto Be-Tai and Proto Hlai, it may not be reconstructible for Proto Kra-Dai. The merger of the velar and uvular stops (both voiceless and voiced) in initial and final positions appears to have occurred very early as well:

In a minority of cases, Kra-Dai forms correspond with the penultimate syllable of the Austronesian forms. These presumably correlate with words with original penultimate stress, the final unstressed -V(C) ending having undergone erosion and eventual deletion (segments preserved in Kra-Dai are placed in angle brackets):

(8) Examples of RAn forms with penultimate stress

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
blind	*<ɓuɻ>a	---	---	*ɓo:t	*ɓo:t
flood	*<baç>aq	*ɓa:ɦ	---	---	*ɓa:h
ten	*<fuɭ>uq	*fu:t	---	---	---
astringent	*a<fəɭ>əd	---	---	*fɯət	*fa:t

The majority of Kra-Dai forms correlate with the final syllables of Austronesian words. In cases where the medial consonant is a voiceless stop, the entire first syllable is lost:

(9) Medial voiceless stops

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
ancestor, grandfather	*apu	*p^{h}u:ʔ	---	*pawh	*pu:h
cut	*[q]ətəs	*t^{h}ət	---	*tac	*tat
headlouse	*guɻu	*tʃ^{hw}u:	---	*hraw	*t^{h}raw
hold in cupped hands	*ra(ŋ)kup	*k^{h}op	*kup	---	*ko:p

The one exception to this is cases of *m-ɻ sequences:

(10) *m-ɻ sequences

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
eye	*maɻa	*tʃ^{h}a:	*ta:	*p-ta:	*p-t^{h}ra:
die	*m-aɻaj	---	*ta:j	*p-ta:j	*p-t^{h}ra:j

The case of medial voiced stops is more complicated than that of voiceless stops, since their realization in Kra-Dai depended on the original preceding vowel (for full discussion, see Norquest & Downey forthcoming). In the majority of languages, when the preceding vowel was schwa, the following voiced stop was phonetically lengthened and became an implosive depending on place of articulation.

Table 4: Reflexes of voiced medial consonants after non-schwa and schwa vowels

PKD	PHlai	PBe	PNT	PC/SWT	PKD	PHlai	PBe	PNT	PC/SWT
*-Vb-	*ʊ	*C-b	*C-b	*ɓ	*-əb-	*ɓ	*C-b	*ɓ	*ɓ
*-Vd-	*ɾ	*C-r	*C-d	*ɗ	*-əd-	*ɗ	*C-r	*ɗ	*ɗ
*-Vɖ-	*ɾ	*C-r	*ɖ	*ɗ	*-əɖ-	*ɗ	*C-r	*ɗ	*ɗ
*-Vɟ-	*hj	(*z)	*j	*j	*-əɟ-	*tç	*C-j	*ʔj	*ʔj
*-Vg-	*ɦ	*g	*ɣ	*ɣ	*-əg-	*k	?	?	?

Initial syllables were also lost when a medial consonant was an implosive:

(11) Medial implosives

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
spring, well	*təbuʀ	---	---	*ɓo:h	*ɓo:h
soak	*ədəm	*ɗə:mʔ	---	---	---
borrow	*çəɟam	---	---	---	*ʔjɯ:m

As described above, there was a tendency for anterior medial voiced stops to undergo secondary implosion, particularly in PC/SWT and most NTai languages with the exception of Saek. They generally underwent lenition in PHlai and PBe (with the exception of *-b- in Qiongshan). Posterior medial voiced stops underwent lenition in all branches:

(12) Medial voiced stops

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
shoulder	*qabaʁa	*ʊa:ɦ	*C-biaʔ	*C-ba:h	*ɓa:h
live, raw	*qudip	*Cuɾi:p	*C-rep	*C-dip	*ɗip
sun, star	*qaɖaw	*ɾa:w	---	*ɖa:w	*ɗa:w
fence, field dike	*paɢər	*Ciɦə:n[5]	---	*ɣal	*ɣan

There are some forms in which the first syllable was lost before the medial stops underwent implosion or lenition. In these cases, they developed in the same way as initial stops:

(13) Medial voiced stops: early loss of first syllable

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
father	*a<ba>	*pʰa:ʔ	---	---	---
sharp	*ʈa<ɟəm>	*tçʰə:m	---	---	---
dirt on skin	*da<gi>	*kʰi:	---	---	---

Initial syllables are also lost before medial fricatives. However, the fricative themselves underwent optional intervocalic voicing, depending apparently on the timing of the loss of the initial syllable. Examples of the voiced fricatives are given in (14) and of voiceless fricatives in (15). It is possible that the PHlai forms in (14) were voiced in Pre-Hlai, since all initial voiced obstruents had devoiced by the time of Proto Hlai (see Norquest 2007 for details):

(14) Medial fricatives: intervocalic voicing

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
fire	*çafuj	*fi:	*wi:	*vi:	*vaj
tooth	*-ifən	*fjən	(*sen)	*van	*van
wash (clothes)[6]	*ba[s]əq	*sɯ:k	*dak	*zak	*zak

(15) Medial fricatives: early loss of first syllable

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
water tortoise	*qaʈi<[f]a>	---	---	*fɯə	*fa:
blow the nose	*ə<səŋ>	---	---	*saŋh	*saŋh
sour	*-[s]əm	---	---	*somʔ	*somʔ
2sg, 2pl	*i<çu>	*səw	*su	*su:	*su:

In many words with medial liquids, the preceding consonant has been preserved through the formation of an initial consonant cluster:

(16) Medial liquids

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
buy, exchange	*bəli	*p-ləj	---	---	---
spotted with white	*[ɓ]əlaŋ	---	---	*ɗa:ŋh	*ɓla:ŋh
dark (red)	*tiləm	---	---	*klamh	*klamh
saliva	*ŋalaj	*hlə:j	*ma[:]j	*mla:j	*mla:j
hear, listen (PF)	*tumaʎa	*p-lɯ:	---	---	---
head	*quʟu	*Curəwʔ	---	*krawʔ	---
taro	*biʁaq	*ra:k	*ʃa:k	*prɯək	*pʰrɯək
ribcage	*taɢəʁaŋ	*kʰa:ŋʔ	---	---	*kʰra:ŋʔ

[5] The first vowel in this comparison doesn't match; it may not be valid unless *-ɢər is a root.

[6] The Austronesian evidence suggest that there may have been two original roots, *-səq and *-cəq, for 'wash clothes' and 'wet' respectively. See also Proto Be *ɟak 'wet'.

In other cases, the initial syllable was lost completely, leaving the medial liquid as the onset:

(17) Medial liquids: early loss of first syllable

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
snake	*çu<laʁ>	*lja:ɦ	---	---	---
fear(ful)	*ta<law>	---	---	*hla:w	*hla:w
three, two	*tə<lu>	*hlu:	---	---	---
child	*a<ʎak>	*hlɯ:k	*lɯk	*lɯk	*lu:k
eight	*wa<ɭu>	*ru:	---	---	---
indigo (grass)	*ʈa<ʁum>	---	---	*hro:mʔ	*[h/s]o:mʔ
dry, withered (PF)	*qa<ʁiw>	---	*ʃ[o]:	*hriəw	*hiəw
cut, reap (PF)	*kə<ʁət>	*rət	---	---	---

*2.1 Kra-Dai *p and *f*

The Kra-Dai correspondences for initial and medial *p and *f are given below:

Table 5: Correspondences for initial and medial *p and *f

RAn	PHlai	PBe	PNT	PC/SWT
*p-	*p^{h}	*p	*p	*p
*-p-	*p^{h}	*p	*p	*p
*f-	*f	(*f)	*f	*f
*-f-	*f	*w	*v	*v

With the exception of the secondary aspiration which all voiceless stops underwent in Proto Hlai, *p remained unchanged in Kra-Dai:

(18) Examples of *p

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
ancestor, grandfather	*apu	*apu	*p^{h}u:ʔ	---	*pawh	*pu:h
slap	*-pik, *-bik	*-pik, *-bik	*p^{h}i:k	---	---	---
to cover	[*lipud]	*lipud	*p^{h}ɯt	---	---	---

*f remained unchanged in initial position, but underwent intervocalic voicing in medial position (this depended on the timing of the loss of the previous syllable, as discussed above).

(19) Examples of *f

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
ten	*puluq	*fuɭuq	*fu:t	---	---	---
astringent	[*apeled]	*afəɭəd	---	---	*fɯət[7]	*fa:t
water tortoise	*qaCipa	*qaʈi[f]a	---	---	*fɯə	*fa:
fire	*Sapuy	*çafuj	*fi:	*wi:	*vi:	*vaj
tooth	*-ipen	*-ifən	*fjən	(*sen)	*van	*van

Note that *-ɭ in final position often became *-t in Kra-Dai, perhaps through the fortition *-ɭ > *-ɖ > *-ʈ > *-t (but see 'body hair' in (28) for an exception)[8].

[7] The PTai reflexes imply a low vowel in the nucleus; for a possible explanation see section 2.3.3 below.

[8] 'Flea' (RAn *qatiməla, PHlai *hmə:t, PBe *C-mat, PTai *hmat), is an example of the opposite kind, where the expected PKD final is *-l. One possible solution to this is to posit metathesis of *l and *t in Austronesian in which the original form was *qali-məta with the animal prefix *qali-; in this case, Kra-Dai would have preserved the original form.

2.3 Kra-Dai retroflex consonants

2.3.1 Kra-Dai *t and *ʈ

Reflexes of *t and *ʈ in initial and medial position are given below:

Table 6: Correspondences for initial and medial *t and *ʈ

RAn	PHlai	PBe	PNT	PC/SWT
*t-	*t^h	*t	*t	*t
*-t-	*t^h	*t	*t	*t
*ʈ-	*$tʃ^h$	*t	*hr	*t^hr
*-ʈ-	*$tʃ^h$	*t	*t	*t^hr

RAn *t remained unchanged in Kra-Dai, with the exception of 'land leech' in PTai:

(20) Examples of *t

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
hit, play	*-teg	*-təɢ	---	---	*tɯk	*tɤk
pound	*-tug	*-tuɢ	---	---	*to:k	*to:k
fall	*-tuq	*-tuq	*t^hok	*tɔk	*tok	*tok
cut	[*[q]etes]	*[q]ətəs	*t^hət	---	*tac	*tat
cut, reap (PF)	*ketun	*kətun	*t^hu[n/ŋ]	---	---	---
seven (PF)	*pitu	*pitu	*t^hu:	---	---	---
fart	*qetut	*qətut	*t^hu:[t/c]	*tot	(*hroc)[9]	*tot
land leech	*qaNi-matek	*qaʃi-matək	*t^ha:k	*ta:k	*da:k	*da:k

The PKD vocalism in 'land leech' probably results from contamination with RAn *qaʃi-məʈaq 'river leech'. The voiced initial in Proto Tai may be due to nasal voicing assimilation: *mətak > *nta:k > *da:k.

The comparisons involving RAn and PKD *ʈ are well-known:

(21) Examples of *ʈ

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
headlouse	*kuCu	*guʈu	*$tʃ^{hw}$u:	---	*hraw	*t^hraw
eye	*maCa	*maʈa	*$tʃ^h$a:	*ta:	*p-ta:	*p-t^hra:
die	*m-aCay	*m-aʈaj	---	*ta:j	*p-ta:j	*p-t^hra:j

The Kra-Dai forms only preserve the final syllable of 'headlouse', although the labiovelar glide in the PHlai form may be a vestige of the first vowel. As with many *m-initial forms (for more examples see (13) above), the initial *m- in 'eye' and 'die' appears to have denasalized to *b- and then undergone regular left-edge devoicing in PTai:

*maʈa > *bəʈa: > *p-ʈa: 'eye'

2.3.2 Kra-Dai *l and *ɭ

While RAn *l is reflected in Kra-Dai as a lateral, *ɭ has rhotic reflexes:

[9] This is a reflex of *ʈ instead of the expected *t.

Table 7: Correspondences for initial and medial *l and *ɭ

RAn	PHlai	PBe	PNT	PC/SWT
*l-	*hl	*l	*l	*l
*-l-	*C-l, *lj	*l	*(h)l	*(h)l
*ɭ-	*r	*r	*r	*r
*-ɭ-	*r	*ʃ	*r	*^{h}r

Examples of *l are given below in (22), and of *ɭ in (23):

(22) Examples of *l

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
snake	*SulaR	*çulaʁ	*lja:ɦ	---	---	---
swallow	*-len	*-lən	---	---	---	*klɯ:n
fear(ful)	*talaw	*talaw	---	---	*hla:w	*hla:w
sink (into mud)	*-lem	*-ləm	---	---	*hlomh	*hlomh
roll, roll up	*luluN	*luluʌ	*C-lun	---	---	---
forget (PF)	*alim	*alim	---	---	*lum	*lɯ:m
dark (red)	*-lem	*-ləm	---	---	*klamh	*klamh
three, two	*telu	*təlu	*hlu:	---	---	---
saliva	[*ŋalay]	*ŋalaj	*hlə:j	*ma[:]j	*mla:j	*mla:j

(23) Examples of *ɭ

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
sesame	*leŋa	*ɭəŋa	*hŋɯ:	---	*ŋra:	*ŋa:
centipede	*qalu-Sipan	*qaɭu-çifan	(*ri:p)	*rep	(*sip)	*k^{h}rep
head	*qulu	*quɭu	*Curəwʔ	---	*krawʔ	---
to plant	[*mula]	*muɭa	*Cura:	*ʃa:	---	---
sell	*saliw	*sa[ɭ]iw	*ri:wʔ	---	---	---
eight	*walu	*waɭu	*ru:	---	---	---

2.3.3 The effects of retroflex consonants on vocalic nuclei

Retroflex consonants appear to have had a centering effect on a following *u; this appears to be consistent in P(Be-)Tai but more sporadic in PHlai:

(24) Effects of initial retroflexes on vowels after high vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
headllouse	*guɭu	*ʧ^{hw}u:	---	*hraw	*t^{h}raw
two, four	*ɖuça	*ʧ^{h}əwʔ	---	---	---
nose, face	*uɖuŋ	*ɗəŋ	*C-raŋ	*ɗaŋ	*ɗaŋ
head	*quɭu	*Curəwʔ	---	*krawʔ	---
eight	*waɭu	*ru:	---	---	---

There appears to be a correlation between final retroflex consonants and the lengthening (and lowering at least in the case of *ə) of the preceding vocalic nuclei (PNT ‘body hair’ is an exception):

(25) Effects of final retroflexes on vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
to skin, peel	*-ʎiʈ	*hli:t	---	---	---
tough, sticky	*-kəʈ	*k^{h}a:t	---	---	---
blind	*[ɓuʈ]a	---	---	*ɓo:t	*ɓo:t
body hair	*[bu[]u	---	---	*pul	---
ten	*[fu[]uq	*fu:t	---	---	---
astringent	*a[fə[]əd	---	---	*fɯət	*fa:t

2.4 Kra-Dai palatal consonants

2.4.1 Kra-Dai *c, *s, and *ç

Since we distinguish RAn *c and *s, and since RAn *s and *ç appear to have merged in Kra-Dai as *s, all three correspondence sets are presented in the table below:

Table 8: Correspondences for initial and medial *c, *s, and *ç

RAn	PHlai	PBe	PNT	PC/SWT
*c-	*tç^{h}	(*c)	(*c)	(*c)
*-c-	*tç^{h}	(*c)	(*c)	(*c)
*s-	*s	(*s)	*s	*s
*-s-	*s	*d	*z	*z
*ç-	*s	*s	*s	*s
*-ç-	(*s)	(*d)	(*z)	(*z)

There is presently only one concrete example of *c; two others are possible:

(26) Reflexes of *c

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
one	*esa	*əca	*tç^{h}ɯ:	---	---	---
snail	[*sisi]	*[c]i[c]i	*tç^{h}i:	---	---	---
sip	*sepsep	*[c]əp[c]əp	*tç^{h}ɯp	---	---	---

As stated above, the palatal fricative *ç appears to have merged with *s in Kra-Dai. Reflexes of *s are given in (27), followed by the sole example of *ç in (28)[10]:

(27) Reflexes of *s

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
blow the nose	[*eseŋ]	*əsəŋ	---	---	*saŋh	*saŋh
sour	[*-sem]	*-[s]əm	---	---	*somʔ	*somʔ
insert, thread a needle	*-suk	*-[s]uk	*sok	---	---	---
wash (clothes)	*baseq	*ba[s]əq	*sɯ:k	*dak	*zak	*zak

(28) Reflexes of *ç

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
2sg, 2pl	*iSu	*içu	*səw	*su	*su:	*su:

2.4.2 The palatal lateral

The palatal lateral *ʎ merged with *l in Kra-Dai (see also section 2.3.2 above):

[10] See also RAn *ɖuça ‘two’, Proto Kra *sa ‘id’ (Ostapirat 1999).

Table 9: Correspondences for initial and medial *ʎ

RAn	PHlai	PBe	PNT	PC/SWT
*ʎ-	*hl	*l	*l	*l
*-ʎ-	*C-l	*l	*l	*l

Examples are given below:

(29) Examples of *ʎ

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
child	*aNak	*aʎak	*hlɯ:k	*lɯk	*lɯk	*lu:k
hear, listen (PF)	*tumaNa	*tumaʎa	*p-lɯ:	---	---	---
fish scale	*quSeNap	*quçəʎaf	*C-lə:p	---	(*kle[p/c])	(*klit)
to skin, peel	*-NiC	*-ʎiʈ	*hli:t	---	---	---
bee, wasp	*waNu	*waʎu	*p-lu:	---	---	---

Note that the frequently cited comparison between RAn *daʎum '(fresh)water' and Kra-Dai 'water' (PTai *C-namʔ, PBe *na:mʔ, PHlai *C-nəm) is incorrect; the correct comparison should be with Proto Formosan *tənəm 'sea'.

2.4.3 The effects of palatal consonants on vocalic nuclei

Palatal consonants appear to have raised the low and mid central vowels *a and *ə, whether they preceded them (30) or followed them (31):

(30) Effects of initial palatals on vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
one	*ə[ca]	*tɕ^{h}ɯ:	---	---	---
sip	*[c]əp[[c]əp]	*tɕ^{h}ɯp	---	---	---
borrow	*çə[ɟam]	---	---	---	*ʔjɯ:m
child	*a[ʎak]	*hlɯ:k	*lɯk	*lɯk	*lu:k
hear, listen (PF)	*tu[maʎa]	*p-lɯ:	---	---	---
fish scale	*quçəʎaf	*C-lə:p	---	(*kle[p/c])	(*klit)
moon	*buɻaʎ[11]	(*C-ɲa:n)	---	*ɓlɯən	*ɓlɯən

The RAn word *gaɲ 'eat, feed' is reconstructed with a palatal final based on both the traditional PAn reconstruction *kaen, in which the schwa (represented by *–e-*) is taken as evidence for the centralizing effect of the following *-ɲ, as well as Proto Oceanic *kani, in which the final vowel is analyzed as excrescent from the preceding palatal nasal[12]:

(31) Effects of final palatals on vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
eat, feed	*gaɲ	*k^{h}ən	*kən	*kɯɲ	*kin
old (living things)	*tuqaç	---	---	*ke:h	*ke:h

2.5 Kra-Dai velars and uvulars

2.5.1 Kra-Dai velar and uvular stops

Due to the initial devoicing described above, Kra-Dai *k and *g have merged in initial position, but are distinguished in medial position:

[11] 'Moon' is included here under the hypothesis that the laterals have metathesized to *buʎaɻ.
[12] For another example of this excrescent vowel, see RAn *çəpac 'four', POc *pati 'id'.

NORQUEST, Peter. 2013. A revised inventory of Proto Austronesian consonants:
Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42:102-126

Table 10: Correspondences for initial and medial *k and *g

RAn	PHlai	PBe	PNT	PC/SWT
*k-	*k^{h}	*k	*k	*k
*-k-	*k^{h}	*k	*k	*k
*g-	*k^{h}	*k	*k	*k
*-g-	*ɦ, *k	*g	*ɣ	*ɣ

Examples of *k and *g are given below:

(32) Examples of *k

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
bite	*-ket	*-kət	---	---	*kat	*kat
catch	[*cikep]	*cikəp	---	---	---	*cap[13]
hold in cupped hands	[*ra(ŋ)kup]	*ra(ŋ)kup	*k^{h}op	*kup	---	*ko:p
cover	*-kup, *-kub	*-kup, *-kub	*k^{h}op	---	---	---
choke	[*cekel]	*cəkəl	*k^{h}ə:nʔ	---	---	---
tough, sticky	*-keC	*-kəʈ	*k^{h}a:t	---	---	---

(33) Examples of *g

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
eat, feed	*kaen	*gaɲ	*k^{h}ən	*kən	*kɯɲ	*kin
1sg	*(a)ku	*(a)gu	*ɦu:	---	*ku:	*ku:
elbow	[*siku]	*sigu	*Ciɦu:ŋɦ	---	---	---
dirty sweat	*daki	*dagi	(*k^{h}i:)	---	*ɣi:	*glaj
thick, viscous	[*buket]	*bu[g]ət	---	*gat	---	---
vagina (F)	*puki	*fugi	---	---	---	*hi:

RAn *k and *q nearly merged in initial position, although they are apparently distinguished in PC/SWT through aspiration in at least some instances. They can be distinguished easily in medial position: although there is variation in the PC/SWT reflexes of *-q- below, they are generally the same as those of intervocalic *-g-, and it appears that *-q- in Kra-Dai underwent intervocalic voicing to *-ɢ- before merging with *-g-.

Table 10: Correspondences for initial and medial *q and *ɢ

RAn	PHlai	PBe	PNT	PC/SWT
*q-	*k^{h}	*k	*k	*k$^{(h)}$
*-q-	*ɦ	*g	*ɣ	*ɣ
*g-	*k^{h}	*k	*k	*k
*-g-	*ɦ, *k	*g	*ɣ	*ɣ

Examples of initial and medial *q and *ɢ are given below:

[13] This form assumes vocalic transfer of the high vowel *i and subsequent palatalization of *k: *cikəp v *kjəp v *cəp v *cɑp.

(34) Examples of *q

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
old (living things)	*tuqaS	*tuqaç	---	---	*ke:h	*ke:h
scratch (mark)	*kur(e)qit	*kur(ə)qit	(ɦɯ:t)	---	---	*k^{h}i:t
thigh, leg	*paqa	*paqa	*ɦa:	(*wa:)	(*k^{w}a:)	*xa:
excrement	*Caqi	*ʈaqi	*ɦa:jʔ	*ga:jʔ	*ɣajʔ	(*k^{h}i:ʔ)
carrying pole	*pasaqan	*pasaqan	---	---	*ɣa:n	*ga:n
to throw	[*buqaŋ]	*buqaŋ	---	---	---	*ɣ^{w}aaŋh

(35) Examples of *ɢ

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
ribs	*tageRaŋ	*taɢəʀaŋ	*k^{h}a:ŋʔ	---	---	*k^{h}ra:ŋʔ
hold in fist	*gemgem	*ɢəmɢəm	---	---	*kam	*kam
high-pitched sound, whinny	[*egik]	*əɢik	*ki:k	---	---	---
fence, field dike	[*pager]	*paɢər	*Ciɦə:n	---	*ɣal	*ɣan

2.5.2 Kra-Dai *ʀ

The most common reflexes of RAn *ʀ in Kra-Dai are given below:

Table 11: Correspondences for initial and medial *ʀ

RAn	PHlai	PBe	PNT	PC/SWT
*ʀ-	*r	(*h)	*hr	*h
*-ʀ-	*r	*ʃ	*r	*^{h}r

*ʀ often appears as a rhotic in initial and medial position:

(36) Examples of *ʀ

Gloss	PAn	RAn	PHlai	PBe	PNT	PC/SWT
dry, withered (PF)	*qaRiw	*qaʀiw	---	*ʃ[o]:	*hriəw	*hiəw
indigo (grass)	*CaRum	*ʈaʀum	---	---	*hro:mʔ	*[h/s]o:mʔ
taro	*biRaq	*biʀaq	*ra:k	*ʃa:k	*prɯək	*p^{h}rɯək
give	[*beRay]	*bəʀaj	---	*ʃe:	---	---
cut, reap (PF)	*keRet	*kəʀət	*rət	---	---	---
wait	*taRah	*taʀah	---	---	*ʧ^{h}a:ʔ	*t^{h}a:ʔ
ribcage	*tageRaŋ	*taɢəʀaŋ	*k^{h}a:ŋʔ	---	---	*k^{h}ra:ŋʔ

Although the Tai reflexes of 'wait' don't provide direct evidence for a rhotic, it can be inferred based on a similar comparison with PHlai for the form 'ask': PC/SWT *t^{h}aa:m, PNT *ʧ^{h}a:m, PHl *ra:m < *təra:m < *təʀam.

In final position, *ʀ appears to be one source of Kra-Dai tone category B:

(37) Examples of final *ʀ

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
spurt (from mouth)	*buʀəç	*p^{h}uɦ	---	---	---
spring, well	*təbuʀ	---	---	*ɓo:h	*ɓo:h
shoulder	*qabaʀa	*ʋa:ɦ	*C-biaʔ	*C-ba:h	*ɓa:h
snake	*çulaʀ	*lja:ɦ	---	---	---

2.5.3 The effects of uvular consonants on vocalic nuclei

Uvular consonants tended to have a lowering effect on high vowels; they promoted breaking of *i and the lowering of *u:

(38) Effects of uvulars on high vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
excrement	*ʈaqi	*ɦa:jʔ	*ga:jʔ	*ɣajʔ	(*k^{h}i:ʔ)
dry, withered (PF)	*qaʀiw	---	*ʃ[o]:	*hriəw	*hiəw
indigo (grass)	*ʈaʀum	---	---	*hro:mʔ	*[h/s]o:mʔ
pound	*-tuɢ	---	---	*to:k	*to:k
fall	*-tuq	*t^{h}ok	*tɔk	*tok	*tok
spring, well	*təbuʀ	---	---	*ɓo:h	*ɓo:h
spurt (from mouth)	*buʀəç	*p^{h}uɦ	---	---	---

Uvulars did not appear to affect non-high vowels:

(39) Effects of uvulars on non-high vowels

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
hold in fist	*ɢəmɢəm	---	---	*kam	*kam
fence, field dike	*paɢər	*Ciɦə:n	---	*ɣal	*ɣan
cut, reap (PF)	*kəʀət	*rət	---	---	---
wash (clothes)	*ba[s]əq	*sɯ:k	*dak	*zak	*zak
hit, play	*-təɢ	---	---	*tɯk	*tɤk
thigh, leg	*paqa	*ɦa:	(*wa:)	(*k^{w}a:)	*xa:
carrying pole	*pasaqan	---	---	*ɣa:n	*ga:n
to throw	*buqaŋ	---	---	---	*ɣ^{w}a:h
taro	*biʀaq	*ra:k	*ʃa:k	*prɯək	*p^{h}rɯək
give	*bəʀaj	---	*ʃe:	---	---
wait	*taʀah	---	---	*tʃ^{h}a:ʔ	*t^{h}a:ʔ
ribcage	*taɢəʀaŋ	*k^{h}a:ŋʔ	---	---	*k^{h}ra:ŋʔ
mouth	*baqbaq	---	*pa:k	*pa:k	*pa:k
shoulder	*qabaʀa	*ʋa:ɦ	*C-biaʔ	*C-ba:h	*ɓa:h
snake	*çulaʀ	*lja:ɦ[14]	---	---	---

The two PTai exceptions above are ‘hit, play’ and ‘taro’. The centralization of the vowel in the latter was due to the influence of the preceding *i. Other examples of this include the following:

(40) Centralization of PTai *a: after *i

Gloss	PHlai-Tai	PHlai	PBe	PNT	PC/SWT
gadfly	*Cila:k	*lja:k	---	*hlɯək	*hlɯək
yellow	*Cila:ŋ	*lja:ŋ	*C-la:ŋ	---	*hlɯəŋ
bedbug	*Ci[l]a:t	---	*rjat	*rɯət	*l̥ɯət

There are three examples in the present dataset of RAn high vowels corresponding with Kra-Dai mid vowels with no influence from uvular consonants:

[14] The initial *lj- in this PHlai form might be explained via coloring of the first vowel by the initial *ç- and subsequent vocalic transfer: *çulaʀ > *çilaʀ > *çəljɑ:ʀ > *lja:ɦ.

(41) Kra-Dai mid vowels not conditioned by uvular consonants

Gloss	RAn	PHlai	PBe	PNT	PC/SWT
hold in cupped hands	*ra(ŋ)kup	*k^{h}op	*kup	---	*ko:p
small, child	*kə[d]i(k)	*ɗiʔ	---	*ɗek	*ɗek
fart	*qətut	*t^{h}u:[t/c]	*tot	(*hroc)	*tot

There doesn't seem to be a straightforward explanation for these if one assumes an original four vowel system in the parent of Austronesian and Kra-Dai. However, if a six vowel system is assumed (which includes *e and *o), then some Kra-Dai mid vowels may be assumed to be original while *e and *o raised to *i and *u in Pre-Austronesian. More research, including a full reconstruction of Proto Kam-Sui, Proto Kra and ultimately Proto Kra-Dai, is required for a more satisfactory solution.

3.0 Austroasiatic Evidence

Given the Austronesian-Kra-Dai distinctions evidenced above, and the support for Benedict's Austro-Tai (AT) hypothesis which they imply, the question arises as to whether any of the same distinctions occur in Austroasiatic in support of Schmidt's Austric hypothesis (Schmidt 1906) later defended by Shorto (1976). In order to explore this question, a set of Austro-Tai[15]-Austroasiatic comparisons have been assembled using the lexical database at the SEAlang Mon-Khmer Languages Project (http://sealang.net/monkhmer/dictionary/).

Since reconstruction of Proto Austroasiatic is an ongoing project, Proto Mon-Khmer (Shorto)[16] data is cited along with several branch-level reconstructions: Proto Vietic (Ferlus), Proto Monic (Diffloth), and Proto Bahnaric, Proto Katuic, Proto Khmuic, and Proto Palaungic (all reconstructed by Sidwell). Although many of these comparisons have been suggested elsewhere in the literature, some are suggested here for the first time. Not all of them may end up proving valid, but as with the Austronesian-Kra-Dai comparisons above, an attempt has been made to control for both semantic and phonological plausibility. In some cases, there are also potential comparisons with Old and Middle Chinese (Baxter & Sagart 2011); these are footnoted where appropriate.

*3.1 Austro-Tai *p and *f*

There does not appear to be any clear indication in Austroasiatic for the *p and *f distinction:

Table 12: Reflexes of An-KD *p and *f in Austroasiatic

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*p	*p	*p	---	---	---	*p	---
*f	*p	*p	*p	*p	*p	*p	*p

Examples are given below. For practical reasons, they are split between comparisons with Austronesian and Kra-Dai respectively; there are some for the same etyma:

(42) RAn *p

Gloss	RAn	PMK	PPal
fathom	*dəpa	---	(*təp)
pinch	*qapit	*pi[:]t	*piət
fold	*ləpət	*[l]pət	*pat
knife	*pisaw	*pi[:]s	---

[15] For the sake of convenience Austronesian-Kra-Dai etyma and reflexes will be labeled 'Austro-Tai (AT)' when being compared with Austroasiatic (Aa) from this point on.

[16] I make one revision to Shorto's PMK vocalism, in that I replace his diphthong *ai with *ɛ:.

(43) KD *p

Gloss	PBe	PNT	PC/SWT	PMK	PViet
shell, bark	*pa:wʔ	---	---	---	*-pɔh
to peel[17]	---	*po:k	*po:k	*pɔ:k	---

(44) RAn *f

Gloss	RAn	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
heart	*fu[ç]uq	*pu:s	---	---	---	---	---	---
bait	*faniŋ	---	---	*pran	---	---	---	*pran
paddy	*faḍaj	*prleʔ	---	---	---	*prleʔ	---	---
lime	*qafuʀ	*kmpur	*kpu:r	---	*kmbo:r	---	---	---
tortoise	*qaṭi[f]a	*$t_1$paʔ	---	*tpa:	*tpa:	*tmpaʔ	---	---
sweep	*caf[i/u]h	*$t_2$pu:s	---	*po:s	*po:s	*pəs	*pi:s	---
tooth	*-ifən	---	---	*p[ɨ/ə]ŋ	---	---	*piəŋ	---
centipede	*qa⌊uçifan	*kʔi[:]p	---	*kʔe:$_3$p, *kɟe:$_3$p	*kahe:p	*kʔi:p	*sʔi[i]p	---

(45) KD *f

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PBahn	PKat	PKhm
tortoise	---	---	*fuə[ʔ]	*fa:	*t1paʔ	*tpa:	*tpa:	*tmpaʔ

Note that there may be a sporadic tendency for Austro-Tai *-f- to correspond with Austroasiatic *-mp- in medial position:

(46) AT *-f- and Aa *-mp- in medial position

Gloss	RAn	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PKhm
lime	*qafuʀ	---	---	*kmpur	*kpu:r	---	*kmbo:r	---
tortoise	*qaṭi[f]a	*fuə[ʔ]	*fa:	*$t_1$paʔ	---	*tpa:	*tpa:	*tmpaʔ

3.2 Austro-Tai retroflex consonants

3.2.1 Austro-Tai *t and *ʈ

There is no distinction in Austroasiatic between *t and *ʈ. Potential etymological connections between Austro-Tai and Austroasiatic have the same reflexes for both series:

Table 13: Reflexes of AT *t and *ʈ in AA

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*t	*$t_{[1]}$	*t	*t	*t	*t	*t	*t
*ʈ	*$t_{[1]}$	---	*t	*t	*t	*t	*t

(47) RAn *t

Gloss	RAn	PMK	PBahn	PKhm	PMon
hit, play	*-təɢ	*[ɟ]tə:k	---	*tɛk	*tak
pierce	*təbək	*t_1[ə/a]p	---	*tap	---
burst	*bətuʔ	*bt$_1$u[ə]h	*ptoh	---	---

[17] MC 剝*pæwk

(48) KD *t

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKhm	PPal	PMon
tie; wrap	*tʰuːk	*tuk	---	---	*tɩu[ː]k	---	---	*tuk	*tuk	---
hit	*tʰa[ː]jɦ	---	---	---	*tɩeh	---	---	---	---	---
boil	---	---	*tumh	*tomʔ	*tɩu[ə]m	---	---	---	---	*tɔːm
pound	---	---	*tam	*tam	*təm	---	---	---	---	---
warn	---	---	*tɯən	*tɯən	*tɩ[e][r]	---	---	---	---	---
hit, play	---	---	*tɯk	*tɤk	*[ɻ]təːk	---	---	*tɛk		*tak
wasp	---	---	*toːh	*toːh	---	*ptɔː	---	---	---	---
erect	---	---	*taŋʔ	*taŋʔ	---	---	*taŋ	---	---	---
wake up	---	---	(*hrɯnh)	*tɯːnh	---	---	---	---	---	*[k]təːr

(49) RAn *ʈ

Gloss	RAn	PMK	PBahn	PKat	PKhm	PMon
eye	*maʈa	*mat	*mat	*mat	---	*mat
tortoise	*qaʈi[f]a	*tɩpaʔ	*tpaː	*tpaː	*tmpaʔ	---
rope	*ʈa[ɭ]iç	*tɩrliʔ	---	---	---	---
that	*-ʈu	*təʔ	*tɔː	---	---	---
vomit	*u[ʈ]aq	*[s]taʔ	---	*kta:[ʔ]	---	*taːʔ
arrive	*ʈəkas	---	---	*tək	---	---

(50) KD *ʈ

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PBahn	PKat	PKhm	PPal	PMon
eye	*ʈʃʰaː	*taː	*p-taː	*p-tʰraː	*mat	*mat	*mat	*mat	---	*mat
grandfather	*ʈʃʰaːʔ	---	*taː	*taː[18]	---	---	---	*taʔ	*taːʔ	---
buy; sell	*ʈʃʰəc	---	---	---	*tɩac	*tac	*tac	*tac	*təc	---
hot	*ʈʃʰwəwʔ	---	---	---	*ktuːʔ	*toʔ	*ʔatoʔ	---	---	*kmtaw
expose to sun	*ʈʃʰiːŋʔ	---	---	---	---	*tiːŋ	*tiːŋ	---	---	---

3.2.2 Austro-Tai *ɖ

There are not enough potential cognates with Austro-Tai *ɖ to establish correspondences in Austroasiatic with any certainty; the few that have been found are given below:

(51) RAn *ɖ

Gloss	RAn	PMK	PViet	PKhm
foam	*buɖaq	---	*bɔːt	---
paddy	*faɖaj	*prleʔ	---	*prleʔ

(52) KD *ɖ

Gloss	PHlai	PNT	PC/SWT	PMK	PKat	PPal
wart	(*suːc)	*ɖuːt	(*soːt)	*ktɩuut	*ktɔːt	toːt

3.2.3 Austro-Tai *l and *ɭ

Unlike the retroflex stops, there may be tentative evidence for a distinction between *l and *ɭ in Austroasiatic:

[18] One explanation for the conflicting correspondence between PHlai *ʈʃʰɑːʔ (< *ʈɑːʔ) and PTai *taː may be that the PTai form is a backloan from an Austroasiatic source.

Table 14: Reflexes of AT *l and *l̥ in Aa

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*l	*l	*l	*l	*l	*l	*(h)l	*l
*l̥	*rl	---	*l	*r[l]	*l	---	*l

Examples of Austro-Tai *l are given below:

(53) RAn *l

Gloss	RAn	PMK	PBahn	PKat
swallow[19]	*tilən	*lu[:]n	*luən	*ləːn
roll	*luluʎ	*luːn	---	---
fold	*ləpət	*[l]pət	---	---
sink[20]	*-ləm	*lə[:]m	---	---
buy	*bəli	---	---	*bləj

(54) KD *l

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
swallow	---	*C-l[ua]n	*klɯːn	*klɯːn	*lu[:]n	---	*luən	*ləːn	---	---	---
banana	---	---	*kluəjʔ	*kluəjʔ	*$t_1$luəjʔ	---	---	---	*tlɔːj	*klɔːj	---
drum	*C-ləŋ	*C-loŋ	*kloːŋ	*kloːŋ	---	---	---	---	---	*kruŋ	---
roll	*C-lun	---	---	---	*luːn	---	---	---	---	---	---
blind	---	*C-lak	---	---	*klak	---	---	---	*lək	---	---
sickle[21]	*C-liːm	*liːm	---	*liːm	---	*liɛm	---	---	---	---	---
water leech	*ljiŋ	---	*pliŋ	*pliŋ	---	*pliːŋ	---	---	---	---	---
release		---	---	*ploːjh	---	*prəːjʔ	*ləːj	---	---	---	---
buy, exchange	*p-ləj	---	---	---	---	---	---	*bləj	---	---	---
spread out	*p-laːɦ	---	---	---	---	*praːs	---	---	---	*plaːs	*laːs
to cover	*p-lom	---	---	---	---	---	---	*-lum	---	---	*[d/g]rlum
destroy	---	---	---	*mlaːŋʔ	*la[:]ŋ	---	---	---	---	---	---
look, see	*ljuːjʔ	---	---	*leː	---	---	---	*taloːj	---	---	---
lick	*ljiːmɦ	*liːmʔ	---	---	*[c]li[ə]mʔ	*-lɛːm	---	---	---	---	---
take off	*ljaːwɦ	---	---	---	---	---	*[p/b]loh	*luoh	---	---	---
sink, drown	*hlom	---	*hlomh	*hlomh	*lə[:]m	---	---	---	---	---	---
deaf	*hləːk	---	---	---	---	---	*klɨk	---	---	---	---
clear	*hlɯːŋʔ	(*daːŋʔ)	---	---	---	*klɔːŋ	*sla[:]ŋ	*-laŋ		*plaːŋ	---
iron[22]	---	---	---	*hlek	---	---	---	---	---	*hlek	---
stake (n)	---	---	*hlak	*hlak	---	---	---	---	---	*hlaʔ	---
sharp-pointed	---	---	*hleːm	*hleːm	*[r]l[ɛː]m	---	---	---	---	*lam	---
betel	---	---	*bluː	*bluː	*ml[əw]	*bluː	*b(ə)luː		*bluː	*bləw	*[s]abluːʔ
slip and fall	---	---	*blaːt	*blaːt	---	*blaːt	---	---	---		---
crawl	---	*[r]ɯːn	*l̥ɯən	*glaːn	---	*-liəŋ	---	---	---	*glan	---
rinse, wash	---	---	*l̥ɯəŋʔ	*laːŋʔ	*laːŋʔ	---	---	---	---		---
swim, float	---	---	---	*loːj	*luːj[ʔ]	---	---	*loːj	*luːj	*hloːj	---
surround	---	---	---	*loːm	---	---	---		*rɔːm		---

Examples of Austro-Tai *l̥ are given below. It is possible that it is distinguished from *l in medial position:

[19] OC 吞 *l̥ʰən
[20] OC 沈 *l̥rəm

[21] MC 鐮 *ljem
[22] OC 鐵 *l̥ʰik

(55) AT *l̥

Gloss	RAn	PHlai	PNT	PC/SWT	PMK	PBahn	PKat	PKhm	PMon
rope	*ʈal̥iç	---	---	---	*$t_1$rliʔ	---	---	---	---
sesame	*l̥əŋa	*hŋɯ:	*ŋra:	*ŋa:	---	*lŋa:	---	*lŋaʔ	*lŋaw
to plant	*mul̥a	---	---	---	---	*ɟmu:l	---	*ɟmɔ:l	---
forest	*çal̥as	---	---	---	---	---	*ʔariː:s	---	---
run	---	*Curu:ɦ	---	---	---	---	*trluh	---	---

This is strengthened by the addition of the form ‘paddy’ to this set, under the assumption that medial *ɖ̥ lenited to *l̥:

(56) AT *ɖ̥/*l̥: Aa *-rl-

Gloss	RAn	PHlai	PMK	PKat	PKhm
paddy	*faɖ̥aj	---	*prleʔ	---	*prleʔ
rope	*ʈal̥iç	---	*$t_1$rliʔ	---	---
run	---	*Curu:ɦ	---	*trluh	---

There is also one example of a mixed correspondence between Proto Tai and Austroasiatic:

(57) KD *l: Aa *r

Gloss	PNT	PC/SWT	PMK	PBahn	PKhm	PMon
coconut	*bla:wʔ	*bla:w[h/ʔ]	*bra:w	*bra:w	*b[l/r]a:w	*bra:w

3.3 Austro-Tai palatal consonants

3.3.1 Austro-Tai *c, *s and *ç

Although the number of comparable forms is fairly small, there does seem to have been a distinction between AT *c, *s and *ç in Austroasiatic:

Table 15: Reflexes of AT *c, *s and *ç in Aa

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*c	*c	*c	*c	*c	*c	*c	---
*s	*s	*s	*s	---	*s	*s	*s
*ç	*s, *ʔ, *c	---	*s, *ʔ	*s, *h, *ʔ	*ʔ	*s, *ʔ	*s

While there are no known Austronesian witnesses which could distinguish between *c and *s in the form ‘dog’, if related, the Austroasiatic evidence suggests the former:

(58) AT *c

Gloss	RAn	PBe	PC/SWT	PMK	PViet	PBahn	PKat	PKhm	PPal
dog	*a[c]u	---	---	*cɔʔ	*ʔacɔ:ʔ	*cɔ:	*ʔacɔ:	*cɔʔ	*cɔʔ
to tear	---	*ʃek	*c^hi:k	*cri:k	---	---	---	---	---

(59) AT *s

Gloss	RAn	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKhm	PPal	PMon
knife	*pisaw	---	---	---	*pi[:]s	---	---	---	---	---
heart	*fu[s]uq	---	---	---	*puːs	---	---	---	---	---
heart[23]	---	*sem	*sim	*sim	---	*seːmʔ	---	---	---	---
string	---	---	*saːj	*saːj	---	---	*k-sɛː	*seʔ	*siʔ	---
high, tall	---	---	---	*suːŋ[24]	*slu[:]ŋ	---	---	---	---	*sloːŋ

(60) RAn *ç

Gloss	RAn	PMK	PBahn	PKat	PKhm	PPal	PMon
hair[25]	*buçək	*su[:]k	*sɔk	*sok	---	*suk	*sɔːk
scale	*quçəʎap	*krcaːp	---	---	---	---	---
tie	*çigət	*[c]kat	*kət	---	---	---	---
centipede	*qaɭuçifan	*kʔi[:]p	*kʔeː$_3$p, *kɟeː$_3$p	*kaheːp	*kʔiːp	*sʔi[:]p	---
forest	*çaɭas	---	---	*ʔariːs	---	---	---
rope	*ʈaɭiç	*t$_1$rliʔ	---	---	---	---	---

(61) KD *ç

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PBahn	PKat	PKhm	PPal
centipede	(*ɾiːp)	*rep	(*sip)	*k^{h}rep	*kʔi[:]p	*kʔeː$_3$p, *kɟeː$_3$p	*kaheːp	*kʔiːp	*sʔi[:]p

The fact that there are multiple reflexes is suggestive of a borrowing scenario, where Austroasiatic *s and *c represent earlier forms (either inherited or borrowed), and those with glottal reflexes (*h and *ʔ) are later borrowings. However, since the comparisons above are rather tentative at this point, further analysis is difficult.

3.3.2 Austro-Tai *l and *ʎ

Austro-Tai *l and *ʎ appear to have merged:

Table 16: Reflexes of AT *l and *ʎ in Aa

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*l	*l	*l	*l	*l	*l	*(h)l	*l
*ʎ	*l	---	---	*l	---	---	---

(62) AT *ʎ

Gloss	RAn	PHlai	PMK	PKat
peel	*-ʎiʈ	*hliːt	---	*-liət
swim	*ʎaŋuj	---	*[l]ŋuj	---
roll	*luluʎ	---	*luːn	---

The one exception is in the PMK form ‘roll’, which has a nasal final reminiscent of the shift which occurred in Proto Malayo-Polynesian from *ʎ to *n.

[23] MC 心*sim

[24] The form for ‘high, tall’, if connected to the Aa forms, is a rare example of prefix preemption in Tai.

[25] This comparison of the Austronesian and Austroasiatic forms for ‘hair’ rests on the assumption of vocalic metathesis in one of the two phyla.

3.4 Austro-Tai velars and uvulars

3.4.1 Velar and uvular stops

The Austroasiatic reflexes of the Austro-Tai velar and uvular stops are reminiscent of Kra-Dai, in that the uvular series has merged with the velar series, and voiced stops have devoiced. Unlike Kra-Dai, however, medial uvular stops have not undergone lenition:

Table 17: Reflexes of AT *k~*q and *g~*ɢ in Aa

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*k, *q	*k	*k	---	---	*k	*k	*k
*g, *ɢ	*k	*k	*k	*k	---	*k	---

(63) RAn *k and *q

Gloss	RAn	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
bite	*-kət	*ki:t	*ka:t	---	---	*kɛ:t	---	*kɯt
scrape	*karut	*ku:t	---	---	---	*ku:t	---	---
cover	*-kup	*cku[:]p	---	---	---	---	---	---
ginger	*laqia	---	---	*kja:$_1$	---	---	---	---
centipede	*qaluçifan	*kʔi[:]p	---	*kʔe:$_3$p, *kɟe:$_3$p	*kahe:p	*kʔi:p	*sʔi[i]p	---
lime	*qafuʀ	*kmpur	*kpu:r	---	*kmbo:r	---	---	---
scale	*quçəʌap	*krca:p	---	---	---	---	---	---
lead (n)	*timəʀaq	*tra:k	---	---	---	---	*tra:k	---

(64) KD *k and *q

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
cover[26]	*k^{h}op	---	---	---	*cku[:]p	---	---	---	---	---	---
throat	*k^{h}ə:k	---	---	---	*kɔ[:]k	---	---	---	*kɔ:k	*ko:k	---
sword[27]	(*kɯ:mɦ)	*ki:mʔ	---	*ki:mh	---	*tkiəm	---	---	---	---	---
work[28]	---	*koŋ	*hoŋ	*koŋ	*kuŋ		---	---	---	---	---
bite	---	---	*kat	*kat	*ki:t	*ka:t	---	---	*kɛ:t	---	*kɯt
mustard green[29]	---	---	*ka[:]c	*ka:t	---	*ka:s	---	---	---	---	---
scrape	---	---	---	*xu:t	*ku:t	---	---	---	*ku:t	---	---
rice	---	---	*ɣawʔ	*k^{h}awʔ	*rk[aw]ʔ	---	---	---	---	---	---
bite	---	---	*ɣap	*k^{h}op	*ka[:]p	---	*kap	*kap	---	---	---

(65) RAn *g and *ɢ

Gloss	RAn	PMK	PBahn	PKat	PKhm	PPal	PMon
tie	*çigət	*[c]kat	*kət	---	---	---	---
elbow	*sigu	*kj[o]ŋ	---	---	---	---	---
1sg	*agu	---	---	*kəw	---	---	---
hold in closed hand	*ɢəm	*k[u]əm	---	---	---	*kəm	---
hit, play	*-təɢ	*[ɟ]tə:k	---	---	*tɛk	---	*tak

[26] OC 蓋 *k^{ʕ}ap-s
[27] MC 劍 *kjæmH
[28] OC 工 *k^{ʕ}oŋ
[29] OC 芥 *k^{ʕ}ret-s

(66) KD *g and *ɢ

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PPal
edge	*ki:ŋ	---	---	---	*ɟki:[ŋ]	*ɟki:[ŋ]	*ɟkiə[ŋ]	---	---
tight	*kɯŋ	*kəŋ	---	---	---	---	---	---	*kəŋ
bury	*kom	---	---	---	---	*kəmʔ	---	---	---
1sg	*ɦu:	---	*ku:	*ku:	---	---	---	*kəw	---
hold	---	---	*kam	*kam	*k[u]əm	---	---	---	*kəm
elbow	*Ciɦu:ŋɦ	---	---	---	*kj[o]ŋ	---	---	---	---
branch	*Cuɦinɦ	---	*ki:ŋh	*kiŋh	*kiəŋ	---	---	---	---
chin	*ɦa:ŋ	*ŋa:ŋ	*ɣa:ŋ	*ga:ŋ	*ka:ŋʔ	*ka:ŋʔ	*ka:ŋ	---	---
penis	---	---	---	*gwaj	*kləjʔ	---	---	*klɛ:	*kleʔ
throat[30]	---	*go:	*ɣo:	*ɣo:	*kɔʔ	*koh	*kɔ:ʔ	---	---
dove	*k^{h}u:	*gu:	---	---	---	*tku:	---	---	---
tie	---	*gat	---	---	*[c]kat	---	*kət	---	---

There are a few forms in which reflexes of Kra-Dai *g correspond with Austroasiatic *g:

(67) KD *g and Austroasiatic *g

Gloss	PHlai	PBe	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PPal
step	*ɦa:mɦ	---	---	---	*ga:m	---	---	---	---
hold in jaws	---	---	*ga:p	*ga:p	---	*ga:p	---	---	---
handle	---	---	*gal	*gan	---	---	---	*gir	---
tie	---	*gat	---	---	*[]gat	---	(*kət)	---	---
3sg	---	*gə	---	---	---	---	---	---	*gə:ʔ

3.4.2 The uvular rhotic

Austro-Tai *ʀ (or *-ʀ clusters) nearly always correspond with Austroasiatic *r:

Table 18: Reflexes of AT *ʀ in Aa

AT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
*ʀ	*r	*r	*r	*r	*r	*r	*r

(68) RAn *ʀ

Gloss	RAn	PMK	PBahn	PKat	PKhm	PPal
hibiscus	*baʀu	*cba:r	---	---	---	---
lead (n)	*timəʀaq	*tra:k	---	---	---	*tra:k
to moan	*-ʀəŋ	---	*-riŋ	---	---	---
left	*wiʀi	---	---	*ʔawiar	*wiʔ	*w[e/ɛ]ʔ

(69) KD *r

Gloss	PHlai	PBe[31]	PNT	PC/SWT	PMK	PViet	PBahn	PKat	PKhm	PPal	PMon
swim	---	*raj	---	---	---	---	*rɛ:	---	---	---	---
edge	---	---	*rim	*rim	*ri[:]m	---	---	---	---	*ri:m	*gnrəm
track	---	---	*ri:	*ro:j	*ru[ə]j	---	---	---	---	---	---
receive	---	---	*rap	*rap	---	---	---	*ra[:]p	---	---	---
love	---	(*də:k)	---	*rak	*r[a]k	---	---	---	*rak	*rak	---
fight	---	---	---	*rop	*rup	---	---	---	*rup	---	---
hundred	---	---	*ro:jʔ	*ro:jʔ	---	---	---	---	---	*rɔ:j	---
to lead	*rujʔ	---	---	---	---	---	*-rɔ:j	---	---	---	---
prepare	*rəp	---	---	---	---	---	*-rap	[ta]rap	---	---	---
rhinoceros	---	---	*re:t	*re:t	*sri:t	---	---	---	---	---	---

[30] OC 喉*g^{ʕ}o

[31] PBe *ʃ- is the regular reflex of Kra-Dai *Cr- clusters.

squirrel	---	---	*ro:k	*ro:k	*prɔ:k	*prɔ:k	*prɔ:k	*prɔ:ʔ	*prɔ:k	---	---
cricket	---	---	---	*[r]it	*[tɻ/c]ri:t	---	---	*ʔariet	---	*ri:t	---
gnat	---	---	---	*ri:nʔ	---	---	---	---	*trəɲ	---	---
river	---	---	---	*ro:ŋh	*ru[:]ŋ	*kro:ŋ	*krɔ:ŋ	rɔ:ŋ	---	*rəŋ	*kro:ŋ
shrink	---	*ʃot	*hrot	*t^{h}rot	---	---	*-ro:t	---	*hwuət	*ru:t	---
disappear	---	---	*hruəj	*t^{h}ra:j	---	---	---	---	---	*hra:j	---
tear	---	*ʃek	---	*c^{h}i:k	*cri:k	---	---	---	---	---	---
to bark	---	*ʃa:wʔ	*hrawh	*hawh	---	*krəh	---	*kroh	---	*rəh	---
cage[32]	---	---	---	*hroŋh	*tru[:]ŋ	(*kdɔ:ŋʔ)	*krɔ:ŋ	*taruŋ	---	---	---
forest	*rəŋ	---	---	---	---	*krəŋ	---	---	---	---	---
pepper	---	---	---	*brik	---	---	---	*brik	---	---	---
machete	---	---	*ɟra:ʔ	*bra:ʔ	---	---	*bra:	*bra:	*bra:	---	*mra:ʔ
banyan	*ri:	---	*raj	*raj	---	*ɟri:	*ɟri:	*ɟari:	---	---	*ɟrəj

3.5 Summary of the Austroasiatic evidence

Unlike the case of Austronesian and Kra-Dai, in which evidence for the proposed consonants *f, *ʈ, *ɭ, *c and *g is mutually corroborative and indicative of a genetic relationship, the comparisons above between Austro-Tai and Austroasiatic are more complex. The Austroasiatic data tenuously support the distinctions between *c, *s, and *ç; this is also true for *-f- and *-ɭ- in medial position, although reflexes of medial *-f- as Austroasiatic *-mp- are sporadic.

Although it is probable that at least a few of the comparisons above will eventually be shown to be chance resemblances, the cumulative weight of the data, as in the case of the Austronesian-Kra-Dai comparisons, suggests a non-accidental relationship between Austro-Tai and Austroasiatic. However, at the present stage of research, it is not as easy to establish whether these correspondences should be attributed to genetic relationship or to contact.

If Austro-Tai and Austroasiatic are in fact genetically related, this would explain the widespread distribution of some of the comparanda listed above. On the other hand, a contact scenario would explain the sporadic reflexes of some posited proto phonemes (most notably AT *ç) as well as the apparent merger of some categories such as *t and *ʈ; two possible explanations for this would be that they had either (a) already merged in the donor language at the time of borrowing or (b) that they were assimilated to the closest Austroasiatic category (for example, if Austroasiatic lacked an *f, then it would be natural to borrow Austro-Tai *f as *p).

6.0 Conclusion

The revision of the Proto Austronesian consonant inventory described in section one opens up new avenues in the comparison of Austronesian and Kra-Dai etyma. Besides creating a more symmetrical and phonetically plausible system in Proto Austronesian, it also offers the following insights:

1. Although cases of intervocalic lenition are common within the Kra-Dai phylum, not all cases of multiple reflexes need to be explained this way. Etyma which have been argued to result from lenition of *p and *k, for example, can now be traced back to an original *f or lenition of *g (parallel with the lenition of the other voiced stops).

2. Vocalic transfer is also common, but some instances of aberrant vocalism can now be explained via the place of articulation of the flanking consonants. Palatal consonants, for example, tend to raise low vowels, and uvular consonants lower high vowels, while retroflex consonants are correlated with centralization, at least in the case of the vowel *u.

3. While implosives appear to be reconstructible at the highest levels of both Austronesian and Kra-Dai, there is evidence that plain voiced stops became implosive secondarily in certain environments. This helps to explain cases of plain-implosive voiced stop mismatch in some correspondence sets.

[32] OC 籠 *k.r^{ʕ}oŋ[ʔ]

4. Finally, one source of the Kra-Dai tone category B (*-h) is final *-ʀ.

An exploratory comparison of the proposed parent of Austronesian and Kra-Dai (Austro-Tai) and Austroasiatic has also been undertaken. While the results are promising enough to support further investigation into this area as progress in the reconstruction of both phyla continues, further analysis is ultimately beyond the scope of the present paper.

The next crucial step in Austronesian and Kra-Dai comparison will involve the completion of the reconstruction of all high-level Kra-Dai branches and the ultimate reconstruction of the Proto Kra-Dai lexicon itself. The integration of this research program with related non-linguistic disciplines such as genetics, archaeology, and paleoclimatology is highly encouraged, as a more refined model of prehistory in Southeast Asia comes increasingly within reach.

References

Adelaar, K. Alexander & Nikolaus P. Himmelmann. *The Austronesian Languages of Asia and Madagascar*. London/New York: Routledge.

Baxter, William H., and Laurent Sagart. 2011. Baxter-Sagart Old Chinese reconstruction, version of 20 February 2011. URL http://crlao.ehess.fr/document.php?id=1217.

Benedict, Paul. 1975. Austro-Thai Language and Culture, with a Glossary of Roots. HRAF Press.

Blust, Robert. 1995. Austronesian Comparative Dictionary. http://www.trussel2.com/ACD/acd

_____. 2002. Kiput Historical Phonology. *Oceanic Linguistics* 41 (2): 384-438.

_____. 2006. The Origin of the Kelabit Voiced Aspirates: A Historical Hypothesis Revisited. *Oceanic Linguistics* 45 (2): 311-38.

_____. 2009. *The Austronesian Languages*. Canberra : Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.

Lase, Apolonius. 2011. *Kamus Li Niha: Nias-Indonesia*. PT Kompas Media Nusantara. Jakarta.

Norquest, Peter. 2007. *A Phonological Reconstruction of Proto-Hlai*. PhD dissertation: University or Arizona.

Norquest, Peter & Sean Downey. 2013. Expanding the PAn Consonant Inventory. *Journal of the Southeast Asian Language Society* 6: 99-145.

_____. Forthcoming. Austronesian implosives.

Ostapirat, Weera. 1999. *Proto-Kra*. PhD dissertation: University of California, Berkeley.

Paz, Consuelo J. 1981. *A Reconstruction of Proto-Philippine Phonemes and Morphemes*. Dilliman, Quezon City: Cecilio Lopez Archives of the Philippine Languages and the Philippine Linguistic Circle, University of the Philippines.

Pittayaporn, Pittayawat. 2009. *The Phonology of Proto-Tai*. PhD dissertation: Cornell University.

Schmidt, Wilhelm. 1906. Die Mon–Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens [The Mon–Khmer Peoples, a Link between the Peoples of Central Asia and Austronesia]. Archiv für Anthropologie, Braunschweig, new series, 5:59-109.

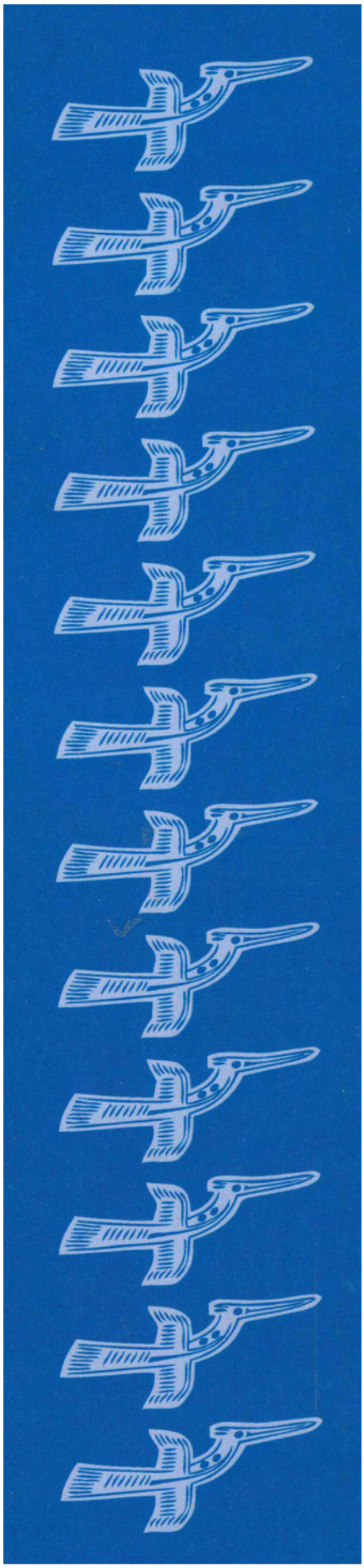
Shorto, H. L. 1976. In Defense of Austric. Computational Analyses of Asian and African Languages 6:95-104.

Wolff, John U. 2010. *Proto-Austronesian Phonology with Glossary* (2 volumes). New York: Southeast Asia Program Publications, Cornell University.

NORQUEST, Peter. 2013. A revised inventory of Proto Austronesian consonants: Kra-Dai and Austroasiatic evidence. *Mon-Khmer Studies* 42:102-126

Mon-Khmer Studies

Volume 42



Authors: Charles Thomas TEBOW II and Sigrid LEW
Title: *A phonological description of Western Bru, Sakon Nakhorn variety, Thailand*
Pages: 127-139
Date Received: 20/5/2013
Revised Text Accepted: 22/11/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A Phonological Description of Western Bru, Sakon Nakhorn Variety, Thailand

Charles Thomas TEBOW II
Payap University, Linguistics Institute

Sigrid LEW
Payap University, Linguistics Institute

Abstract
This paper provides a phonological analysis for a Western Bru variety spoken in Northeast Thailand labelled Bru Sakon Nakhorn (Bru SN). Previous descriptions of Western Bru varieties differ in the amount of distinctive vowel qualities, the presence of onglides linked to phonation and vowel height, the contrastive status of the feature nasalisation, and vowel contrast in reduced syllables. The present analysis identifies contrastive onglides, lack of contrastive nasalization, and predictable vowel qualities in reduced syllables. It further argues that the consonants often described as palatal plosives or alveolo-palatal affricates in Mon-Khmer languages are alveolo-palatal plosives. The vowel system indicates that diphthongs are phonologically short vowels. Ongliding related to vowel height and ponation type is not present. Furthermore, this variety appears to differ from closely related So in distinguishing onglides and offglides. Spectrograms and minimal pairs reveal that they are contrasting phonemes, not allophones, indicating that Bru SN clearly retains this vowel contrast, in line with other Bru varieties.
Keywords: Katuic, Western Bru, phonology
ISO 639-3 language codes: brv

1. Introduction

Bru belongs to the Katuic branch of the Austroasiatic languages and is spoken in Vietnam, Laos, Cambodia, and Thailand. The Katuic family derives its name from the Katu language and has a special status within the Mon-Khmer group because it is rich in both ancient lexical and phonological retentions as well as innovations regarding the vowel system (Sidwell 2005). There are different views on which Katuic subgroup Bru belongs to. The categorization as West, North, or North-East Katuic is mainly based on whether lexical or phonological parameters were chosen for comparison (Sidwell 2009). Peiros (1996) considers Bru, Kui and Pacoh as separate subgroups next to Katu. Miller & Miller (1996) label Bru, So, Tri, Makong, Siliq and Katang as North Katuic. Luang-Thongkum (2001) categorizes Bru, So and Pacoh as North-East Katuic. Sidwell (2005) agrees with Diffloth & Zide (1992) in identifying Bru, So, Kui and Souei as West Katuic, placing them next to the Ta'oih, Katu, and Pacoh dialect chains. Under West Katuic, Lewis et al. (2013) distinguishes Eastern Bru in Laos and Western Bru in Thailand next to So in Laos and Khua in Vietnam under the Brou-Sou branch. Independent from the actual classification, So commonly appears in the same subcategory as Bru (Diffloth & Zide 1992, Miller & Miller 1996, Thongkum 2001, Sidwell 2005).

Apart from comparative studies on Katuic languages including Bru, there are grammatical observations on the Eastern Bru variety Bru Tri of Vietnam and Laos by J Miller (1964) and C Miller (1964), vowel descriptions for the same variety (J Miller 1976; Phillips, Miller & Miller 1976; Vương 1999) and phonological descriptions for a Western Bru variety in North-East Thailand by Luang-Thongkum (1979) and Green (1996). Gainey (1985) compared the phonological inventories of Kui, Bru and So varieties in northeastern Thailand. A grammatical description of So was provided by Migliazza (1998). Miller and Miller compared the phonological inventories of six Bru varieties in three provinces of northeast Thailand (1995), provided a Bru dictionary (1995), and conducted a lexical comparison of So and Bru dialects in Thailand (Miller & Miller 1996). Grammatical observations (Engelkemier 2010) and a discourse study (Tebow 2010) are available for one of these Bru varieties, referred to as Kok Sa-ak in Miller and Miller's 1996 study and found in Sakon Nakhorn.

TEBOW II, Thomas Charles and Sigrid Lew. 2013. A phonological description of Western Bru, Sakon Nakhon variety, Thailand. *Mon-Khmer Studies* 42:127-139

As Luang-Thongkum (1979) points out, the language name Bru (found with various spellings in the literature) refers to phonologically different varieties, depending on whether they are found in Thailand or in Vietnam. In the Vietnam variety, register is contrastive only for long monophthongs and diphthongs and basically manifests itself in diphthongisation; here onglides mark the tense register, and monophthongs are found in the lax register (cf. J Miller 1967; Philipps, Miller, and Miller 1976; Vương 1999). For the Thai Khong Chiam variety, Luang-Thongkum (1979) identifies contrastive modal to slightly tense and breathy phonation for all short and long vowels. In the context of breathy phonation, open vowels are realised with slight onglides. For the same variety, Green (1996) identifies onglides also for tense close and mid vowels.

Next to different transcriptions of monophthongs depending on the register, Luang-Thongkum's (1979) and Green's (1996) studies also differ in the amount of contrastive vowel qualities. Luang-Thongkum (1979) identifies a symmetrical system with 11 vowel qualities comprised of 4 unrounded front and 4 rounded back vowel heights, and 3 unrounded back vowel heights next to three offglides /iə, ɯə, uə/ and two onglides /ia, ua/. Green (1996) describes closed, close-mid and open vowels with all expected length and register distinctions but postulates a rather irregular set of mid and open-mid vowels with gaps for opening degree, length, and phonation contrast, next to vowel contrast for /u/ and /a/ in reduced syllables. Miller and Miller (1996) identify 3 front vowel heights, 4 central vowel heights, and even 5 back vowel heights for six Bru varieties in Thailand.

Lastly, Gainey (1985) calls nasalization contrastive for Bru and So in Thailand whereas Luang-Thongkum (1979) identifies some contrastive but mostly context-bound nasalization, and Green (1996) similarly argues that nasalisation contrast is predictable or doubtful. This study will examine the phonological inventory of the Bru variety in Sakon Nakhorn province, hereafter Bru SN, in order to determine where in the varying phonological descriptions of Western Bru it fits in regard to the amount of distinctive vowel qualities, the presence of onglides linked to phonation and vowel height, the feature nasalisation, and vowel contrast in reduced syllables since there appears to be no consensus in the previous studies.

After providing a short overview of Bru SN in its linguistic context, a description of Bru SN syllable and word structure will follow. Bru SN phonology with its 24 consonants and 15 contrastive vowels further distinguished by length and two phonation types will then be presented and discussed in the third section. It will be argued that obstruents typically described as palatal stops or alveolo-palatal affricates best are interpreted as alveolo-palatal stops. Whereas diphthongs correlate to long vowels in the related Kui varities (Van der Haak & Woykos 1990), they pattern like short vowels. Special consideration will be given to the two types of diphthongs found in this language, onglides and offglides, as postulated for Proto Katuic by Sidwell (2005). The distinction of onglides and offglides may not be found in this very closely related language, depending on the analysis (Gainey 1985 vs. Migliazza 2003). Explanations for this deviation from So are discussed in the final part.

2. Bru SN Linguistic Background

The particular variety underlying this phonological description is spoken in the Phang Khone and Phanna Nikhom sub-districts of Sakon Nakhorn province of Thailand. It has approximately 5000 speakers and was initially named after the village where the development took place, Bru Khok Sa-at, because it has been found to represent the dialects spoken in all eight Bru villages along the shore of the Nam Un reservoir.[1] Bru SN speakers in Thailand are generally bilingual or trilingual. The older generation (40+ years) have had little formal education and speak Bru in the home and in the village. In situations where there are Isan speakers present, they will speak Isaan. Bru SN who are younger than 40 years speak both Bru and Isaan in the home and in the village, Isaan in the wider community, and standard Thai in formal situations and outside of the Isaan district as they have had more formal education (Choo et. al. 2012).

Gainey (1985) notes that Kui, Bru and So are lexically and phonologically similar. All three varieties have eight out of 24 compared phonological phenomena in common, with Kui and So sharing ten of those. Bru and So, however, share 17 out of 24 phonological features, including

[1] The eight villages are Khok Sa-at, Kham Wae, Nong Hai Yaay, Nong Hai Noi, Naa Lao, Naa Than, Huay Bun and Hin Taek.

contrastive nasalization, identical vowel inventories, more variety in the consonant and vowel inventory of reduced syllables, final voiceless and glottalized approximants, and a /tʰr-/ cluster. Bru SN and the So language of the Kusuman district of Northeast Thailand in fact are very similar, as seen in the 91 percent of cognates they share (Miller & Miller 1996) and in the ease of acquired intelligibility between the two languages (Choo 2012). Migliazza (2003) reports that the So people came to Thailand from central Laos within the last 150 years due to political and economic conditions there. According to the story teller of Khok Sa-at (personal communication with Tebow), the neighboring So of Kusuman district and the Bru of Dong Luang were in the same migration from Laos a few generations ago, settling in different regions. Miller & Miller (1996) state that the Bru of Khok Sa-at originate from the Meung Wang area in Laos near the Vietnamese Border. Based on the author Tebow's personal observations and in agreement with C. Miller (personal correspondence), Bru SN is not mutually intelligible with the Eastern Bru variety Bru Tri of Laos and Viet Nam (250 kilometers distant), nor is it mutually intelligible with the Bru of Don Luang in Mukdahan province (100 kilometers away).

The data for this description comes from a lexicon of 3000 words collected in the village between February 2007 and April 2008. Further text collection has continued from October 2008 to May 2012. A corpus of 3100 words has been recorded and processed using FieldWorks and Speech analyser (SIL 2007).

3. Bru SN Word Structure

Bru SN is an isolating language with very little derivation or inflection. A limited set of verbs can be marked as plural iterative with the minor syllable prefix /r-/. A larger set of verbs can also be preceded by the causative marker /a-/. Compounding is a common way of word formation, like /tʰrɛː/ 'shirt' and /a.laːj/ 'pants' joining to form /tʰrɛː.a.laːj/ 'clothes', or /sɛːm/ 'younger sibling' and /aːj/ 'older sibling' forming /sɛːm.aːj/ 'siblings.' Compound words can have two to four syllables. The longest words in the data corpus are loanwords with up to four syllables, as in /seː.naː.bɔ.diː/ 'advisor' is coming from Thai. The shortest word would consist of a vowel as in /ɔː/ 'good'.

Bru SN follows the word structure typical for mainland southeast Asia and has monosyllabic as well as quasi-disyllabic (cf. Michaud 2012) or sesquisyllabic words (cf. Matisoff 1973), the latter consisting of a full syllable and a preceding phonologically reduced syllable. Sesquisyllables can be either full words with a bound grammatical prefix, originally cluster-initial loanwords with epenthetic vowels to avoid any non-native consonant sequences, or old compounds where the first classifying element fossilised (Matisoff 1989); over time it entirely or partially lost its meaning and, consequently, some of its phonological features that were not needed to distinguish the former meaning anymore. This appears to be the reason why only a limited inventory of consonants is found in reduced syllables. Vowel contrast and contrast of suprasegmental features like length, tone, or voice quality is neutralised in this context so that reduced syllables can be considered mora-less (Herr 2011). Therefore they may be entirely deleted, especially in Austroasiatic languages (Michaud 2012). Nasal initials often turn into syllabic nasals in this context, causing the so-called prenasalization which is another areal feature found in languages of all types in mainland southeast Asia (cf. Henderson 1965).

Full syllables show the full range of initial and final consonants and have contrastive vowel quality, length and register. All consonants except glottalized approximants /wˀ, jˀ/ and the voiceless palatal approximant /j̊/ can serve as onsets. Full syllables can have initial clusters formed by plosives liquids /l, r/. Codas consist of voiceless plosives and fricatives, the nasals, and all approximants including the liquids. Syllable kernels are made up by a long or short monophthong or a diphthong with either modal or breathy voice. The shape of the full syllable is C1(C2)V(:)(C3).

In reduced syllables, phonation, vowel quality and length are neutralized. The consonant inventory is restricted to initial /p, pʰ, t, ʈ, k, kʰ, m, s, r, l/, a short open central vowel with the obligatory glottal onset [ʔĕ], marked with the glottal stop in this paper, and optional final nasals /m, n/, with the place of articulation assimilating to the following full syllable onset. Reduced syllables with a final nasal as in [tĕmpoːr] 'yank' are rare, though. Even though the glottal stop is a redundant feature, occurring as an abrupt onset for initial vowels and following final short vowels with modal voice, it is chosen to mark the reduced syllable onset in line with the contrastive

consonant onsets, since the vowel is not contrastive in this position, either. Unlike Green's (1996) findings for Bru Khong Chiam, vowel quality is predictable in reduced syllables. The default is is a short near-open central vowel [ĕ] as in [ʔĕ.mʌʔ] /ʔ.mə/ 'who', whereas a near-close back [ŭ] occurs with velar /k/ as in [kŭ.tɛʔ] /k.tɛ/ 'earth', and near-close front [ĭ] follows alveolar and alveolo-palatal /s, ɬ/ as in [sĭ.baːw] 'shout' or [ɬĭ.pi̤ᵃk] 'mongoose'. Since there is no contrast between the three vowel qualities in reduced syllables, there are some dialectal differences for the vowels in reduced syllables. For example, in the dialect of Naa Lao village located 10 kilometers from Khok Sa-at village, the vowel is not assimilated to back articulation in velar-initial reduced syllables, as in [kĕ.tɛʔ] 'earth'. Because vowels in reduced syllables differ from area to area and are not contrastive but fully predictable, they do not occur in the phonemic transcriptions, following current practice (cf. Svantesson & House 2006, Gafos 1999, Shaw 1993). The syllable breaks indicate that the consonant is the syllable onset, not the first consonant of a cluster onset.

There are exceptions for minor syllables with a sibilant onset in loan words. If a borrowed word starts with [sĕ] the vowel quality is kept, as in [sĕ.lɔp] 'to faint', and not changed to the predictable vowel form in reduced syllables. Since the phonological process of vowel reduction and ultimately deletion in initial syllables of sesquisyllabic words does not get applied to /s/-initial loanwords, it could be argued that these borrowed words are not perceived as sesquisyllabic but disyllabic. The vowel in the first syllable appears to be perceived as a phoneme because it does not follow the pattern of Bru SN sesquisyllabic words where the initial sibilant is always followed by a high close vowel.

A mere reflex of a former nasal-initial reduced syllable is found in the prenasalization of initial plosives of full syllables. These homorganic nasals such as /n.to/ 'tree-bark' or /ŋ.kọ̄ːŋ/ 'to crawl' are perceived as a modification of the following plosive and are considered to be one syllable by native speakers. Obstruent- and liquid initial sesquisyllabic words like [ʔĕ.ɬɔː] 'dog', on the other hand, are counted as two syllables. Accordingly, the compound word /m.pɛ.s.muːt/ (lit. 'mother ant' = 'witch') with a prenasalized full syllable followed by a reduced syllable is considered to have three syllables: [ᵐpɛʔ.sĭ.muːt]. Examples for mono- and sesquisyllabic Bru SN words are given below.

V	/ɔ/	'grandfather'
V:	/ɔː/	'good'
CV	/po/	mouth
CV:	/kiː/	'loom'
CVC	/sɔk/	'hair'
CV:C	/ɬu̯ːt/	'wipe'
CV	/pra/	'money'
CCV:	/plaː/	'flame'
CCVC	/plɔŋ/	'to blow'
CCV:C	/kʰlaːp/	'wing'
c.CV	/ʔ.ɬɔː/	'dog'
c.CVC	/ʔ.laj/	'3rd PL'
c.CCVC	/ʔ.blɯh/	'ask'
c.CVC	/k.baŋ/	'bowl'
c.CCV	/t.kloː/	'lay head down'
c.CCVC	/k.trih/	'shake'
cn.CV	/kn.ɬɛː/	'key'
cn.CVC	/tm.poːr/	'to yank'

TEBOW II, Thomas Charles and Sigrid Lew. 2013. A phonological description of Western Bru, Sakon Nakhon variety, Thailand. *Mon-Khmer Studies* 42: 127-139

4. Bru SN Phonology

Bru SN has 23 consonants and 46 vowels, including contrastive length and phonation. All Bru SN phonemes with examples will be described in the three sections below.

4.1 Consonants

20 of the 23 distinctive Bru SN consonants can occur in syllable-initial position. Aspiration is contrastive for all plosives. Voicing is contrastive for bilabial and alveolar but not for alveolo-palatal and velar stops. The alveolo-palatal stop has been transcribed with the symbol for palatal stops /c/ by many researchers (e.g. Thomas 1967, Sidwell 2005). Since the plosive has a relatively long release noise, an affricate symbol /tɕ/ like used in Thai (Tingsabadh & Abramson 1993) could be used instead. This is misleading for two reasons.

- The short [i]-like transition preceding it when it occurs in syllable-final position clearly indicates that the closure is not alveolar; a final /t/ would not be preceded by this transitional vowel.
- Even though frication is allowed for finals, as in final /h/, final plosives always are unreleased. This favors the interpretation of the alveolo-palatal obstruent as a plosive, not an affricate, because in syllable-final position this sound is always unreleased and thus patterns like a plosive.

For the present description, the non-IPA symbols for alveolo-palatal plosives have been chosen because they represent the actual sound more closely than the symbol for the palatal stop.

The rhotic is realised as an alveolar flap in the onset and as a trill in the coda of a syllable. In clusters, it can be a trill or an alveolar approximant. Table 1 below shows all Bru SN consonant phonemes.

Table 1: Bru SN Consonants

Manner/Place of Articulation	bilabial	alveolar	alveolo-palatal	post-palatal
plosive	b p pʰ	d t tʰ	ʨ ʨʰ	k kʰ
nasal	m	n	ɲ̟	ŋ
approximant	w wˀ	l r	j jˀ ʲ̊	
fricative		s		h

Apart from /wˀ jˀ ʲ̊/ every consonant occurs in syllable-initial position. Examples for syllable-initial clusters are: /pra/ ‘money’, /pʰriː/ ‘to unroll’, /brih/ ‘speck’, /tran/ ‘animal’, /tʰrɔː/ ‘rice’, /kruᵃp/ ‘to cover’, /plɔŋ/ ‘to blow’, /blɔŋ/ ‘alcohol’, /klɔːk/ ‘white’, /kʰlaːp/ ‘wing’. The bilabial approximant /w/ is realized as a fricative [ʋ] when it precedes the closed front vowel /i/, as in [ʋil] ‘village’. The bilabial fricative [f] in words borrowed from Thai such as /faràŋ/ ‘foreigner’ and /fŭn/ ‘dust’ is a possible allophone of the aspirated voiceless bilabial plosive /pʰ/ that these loan words usually are produced with, as in the Bru rendering /pʰ.raŋ/ and /pʰon/. Examples for initial consonants are given below.

/bak/ ‘cut’	/dop/ ‘dive’	/ʨu̟ːt/ ‘wipe’	/kaːŋ/ ‘middle’
/pak/ ‘pierce’	/top/ ‘basket’	/ʨʰu̟ːt/ ‘sting‘	/kʰaːŋ/ ‘flat blade’
/pʰaːk/ ‘scrape’	/tʰok/ ‘flick.rope’	/rɑ̟ŋ/ ‘rich’	/sɑ̟ŋ/ ‘to weigh’
/mᵉɑ̟/ ‘rain’	/naj/ ‘this’	/ɲ̟ᵉam/ ‘cry’	/ŋᵉɑ̟m/ ‘sweet’
/wᵉa/ ‘borrow’	/lᵉɑ̟j/ ‘stripe’	/jɛːŋ/ ‘gold’	/hɤ̟ŋ/ ‘critical’

Bru unreleased voiceless oral stops, all nasal stops, and all approximants including the liquids occur in syllable-final position. The glottalized approximants /wˀ jˀ/ as well as the voiceless voiceless approximant /ʲ̊/ occur only as codas; they are not allophones of the non-

glottalized or voiced approximants since they form clear contrast and are found with both phonation types. Examples for final consonants are seen in the following words.

/klạp/ 'touch'	/mạt/ 'eye'	/maṭ/ 'beautiful'	/mạk/ 'love'
/l^{e}ạm/ 'spread'	/l^{e}ạn/ 'thresh'	/m^{e}ạɳ/ 'borrow'	/r^{e}ạŋ/ 'abandon'
/saːw/ 'grab'	/ku.jaːl/ 'wind'	/klaːj/ 'pass by'	/ʔ.paj̊/ 'cotton tree'
/k.jaːwʔ/ 'step'	/t.mir/ 'lips'	/r.taːjʔ/ 'scatter'	/sah/ 'bail'

4.2 Vowels

Bru SN has 14 contrastive vowel qualities. Nine of them are monophthongs and five are diphthongs. For the monophthongs, length is phonemic, totalling 24 vowels. Only short monophthongs and modal voice diphthongs can be followed by a glottal stop, indicating that diphthongs are phonologically short vowels. For the onglides /^{e}a, ^{o}a/ the phonetic final glottal stop is an artefact indicating modal voice. Phonation is phonemic for both monophthongs and diphthongs which leads to a total of 46 contrastive vowel varieties. The occurrence of breathy vowels is slightly limited: There is no short /ẹ/ in smooth syllables and no long /ạː/ in checked syllables. Syllable-initially, vowels are preceded by a fully predictable glottal stop, meaning there are no vowels without a glottal onset. In syllable-final position its occurrence is limited to short modal vowels.

Ladefoged and Maddieson (1996) have organised phonation types on a continuum from maximal glottal opening to maximal glottal constriction. The five laryngeal settings in this continuum which they find to be sufficient to describe languages are: 1. Breathy, 2. Slack, 3. Modal, 4. Stiff and 5. Creaky. They argue that for consonants, languages contrast phonation types which are more than one degree in distance. For vowels, Jalapa Mazatec has been reported to have a three way contrast between breathy, modal and creaky voice. Tonal Mpi contrasts modal and stiff voice. For many languages with contrastive phonation the voice quality is slack or slightly breathy, also called lax by Keating et al. (2010), versus stiff or slightly creaky, also called tense (Keating et al. 2010). Bru SN contrasts modal and breathy voice like the related language So. The Eastern Bru language Bru Tri, on the other hand, contrasts creaky and modal voice for the same cognates. Thus all three languages have a lax and a tense register, placed at different points on the continuum of glottal constriction.

All Bru SN vowels, including length and phonation distinctions, are shown in Table 2. Unlike observed for Bru Khong Chiam (Luang-Thongkum 1979, Gainey 1985), nasalisation is not contrastive in Bru SN. Also, there are not four but only three front and back vowel heights, resembling Green's analysis (1996). The vowel inventory will be discussed in the sections below.

Table 2: Bru SN Vowels

	Front		Back centralized		Back	
Close	i	i̤	ɨ	ɨ̤	u	ṳ
	iː	i̤ː	ɨː	ɨ̤ː	uː	ṳː
Mid	e	e̤	ə	ə̤	o	o̤
	eː	e̤ː	əː	ə̤ː	oː	o̤ː
Open	ɛ	ɛ̤	a	a̤	ɔ	ɔ̤
	ɛː	ɛ̤ː	aː	a̤ː	ɔː	ɔ̤ː
Diphthongs	i^{a}	i̤^{a}	ɨ^{a}	ɨ̤^{a}	u^{a}	ṳ^{a}
	^{e}a	^{e}a̤			^{o}a	^{o}a̤

The diphthongs in Bru SN are comprised of three offglides or rising diphthongs, and two onglides or falling diphthongs, like Bru Khong Chiam described by Luang-Thongkum (1979) and Green (1996). Neither onglides nor offglides distinguish length; like the short monophthongs, they

can be followed by glottal stops and thus pattern like short vowels. The existence of five diphthongs is not surprising since they follow general Bru phonology. On the other hand, for So which shares many cognates with Bru SN there is no consensus on its dipththong inventory. It has been described as having three three offglides and two onglides (Gainey 1985) but also only three diphthongs which are all offglides (Migliazza 2003). Even though the long vowels /a:/ and /ɔ:/ in Thai or Lao words often are replaced by breathy diphthongs, as seen in words like /tʰa:/ → /t^{e}a̤/ 'to spread', /na:j/ → /n^{e}a̤j/ 'mister', /pʰɔ:/ → /p^{o}a̤/ 'enough' or /kᵏɔ:n/ → /k^{o}a̤n/ 'hammer' (tones are not marked as they differ in the Tai-Kadai sources and have no impact on the Bru SN vowel quality under study), only approximately one third of the data with onglides could be the product of a loanword mechanism. The breathy offglide /i̤^{a}/ is found in 16 words of which 4 are Tai-Kadai loan words. Breathy /^{e}a̤/ is found in 91 words of which 32 are borrowed from Thai or Lao where it replaces the long /a:/. Only in these cases, about a third of all occurrences, the onglides could be the product of a loanword mechanism which may be linked to voiced onsets in Proto Tai or Proto Southwest Tai, e.g. Bru SN /k^{o}a̤n/ 'hammer' < PT *ɣo:l (Pittayaporn 2009) or *ɣɔn (Li 1960), or PSWT *ɣɔɔn (Li 1960) or *ɣɔn (Jonsson 2009). However, the majority of Katuic languages adopted loanwords after Proto Katuic split up into its daughter languages, and it is not possible to determine when particular words have been borrowed (Peiros 1991). Furthermore, the majority of words with onglides are Bru cognates.

For Khmer it is hypothesised that the prolonged formant transition in breathy vowels allows the reinterpretation as diphthongs (Wayland & Jongman 2002). That there is a link between long open monophthongs and onglides is supported by the fact that Sidwell (2005) often offers two reconstructions with both of these vowel types, as in Proto Katuic *baam, *biam 'chew' and *atɔɔŋ, *atuaŋ 'beans'. A possible explanation for why Bru SN and closely related So possibly differ in their diphthong inventory might be found by taking a look at Proto Katuic as reconstructed by Sidwell (2005). He postulates six diphthongs, *ie and *ia with front, *i̤ə and *i̤a with central, and *uo and *ua with back commencing vowels. The Bru SN offglides /i^{a}/ and /u^{a}/ could be the equivalents of PK *ie and *uo, and the onglides /^{e}a/ and /^{o}a/ would be the counterparts of PK *ia and *ua with the open and thus more prominent central vowels. Bru SN might have lost two of these six diphthongs already, a hypothesis underlined by the fact that onglide /^{e}a̤/ is frequent but has no breathy counterpart, and there are only 10 lexemes which contain the central diphthong /i̤^{a}/ in the data corpus. The example /ŋ^{e}a̤m/ 'sweet' and its PK root *ŋaam do not support this hypothesis. A comparative study to investigate whether Bru SN and other related varieties like Khong Chiam possibly preserve older vowel distinctions than So appears promising but is not possible within the framework of this paper. Furthermore, an analysis of So varieties in Thailand to verify the conflicting results in Gainey's (1985) older and Migliazza's (2003) more current analysis seems necessary.

Examples for all Bru SN vowels including phonation and length contrast are given below.

/k.trih/	/kɻt/	/trɨŋ/	/kɻŋ/	/buh/	/tu̱p/
'shake'	'pick off'	'jungle'	'stare'	'burn'	'bury'
/ki:/	/kɻ:/	/kɨ:jˀ/	/kɻ:p/	/su:n/	/mu̱:t/
'loom'	'that'	'small'	'cave'	'angry'	'enter'
/tet/	/pe̱t/	/ə:t/	/jə̱:ŋ/	/ʔ.pok/	/ŋo̱k.ŋ^{o}a̱ŋ/
'stuck to'	'prance'	'located'	'far'	'to cover'	'praise'
/ʔ.ke:l/	/te̱:t/	/nəŋ/	/ŧə̱ŋ/	/mo:k/	/mo̱:k/
'knee'	'chant'	'also'	'straight'	'hat'	'glutton'
/ʔ.lɛh/	/tɛ̱h/	/an/	/ʔ.la̱ŋ/	/pɔt/	/kɔ̱k/
'argue'	'leech'	'3rd sg'	'clear'	'unhook'	'curve'
/ʔ.dɛ:l/	/nɛ̱:w/	/mat/	/ma̱t/	/ʔ.pɔ:t/	/ʔ.pɔ̱:h/
'line up'	'thing'	'future'	'eye'	'play violin w/ bow	'bubble'
/li^{a}n/	/ti̱^{a}/	/pi^{a}t/	/pi̱^{a}ŋ/	/tu^{a}ŋ/	/ŋu̱^{a}ŋ/
'to study'	'before'	'bored'	'straw'	'lake'	'before'

TEBOW II, Thomas Charles and Sigrid Lew. 2013. A phonological description of Western Bru, Sakon Nakhon variety, Thailand. *Mon-Khmer Studies* 42: 127-139

/m.nᵉa/	/mᵉạ/	/nᵒaŋ/	/kᵒạŋ/
‘stink’	‘rain’	‘assistant’	‘custom’

4.2.1 Monophthongs

Length and phonation are contrastive for monophthongs. The symbol /a/ stands for an open central to back centralized vowel [a̰]. The phonetically most unstable vowel is the mid central vowel /ə/. For the modal alternant, its length is neutralized in open syllables where it is realized long in isolation but short in connected speech. As to vowel quality, this phoneme has two allophones, a long close-mid back centralized [ɤ̰ː] and a short open-mid back centralized [ʌ̰] which sometimes is realized as close-mid [ɤ̰] if the syllable ends on a velar plosive. Apart from this conditioned variation, there is both inter and intra-speaker variation found for the height of this vowel. Another phonological process is observed for the open-mid rounded back vowel /ɔ/. It is realized as an open rounded back vowel [ɒ] if it is short and therefore followed by a glottal stop. In every other environment the vowel quality is close-mid [ɔ] as in [ʔɔː] ‘good’ and [pɔt] ‘unhook’ vs. [ʔɒʔ] ‘grandfather’. Green’s (1996) findings of an irregular system of contrastive open and mid back and central vowels for Bru Khong Chiam are not applicable to Bru SN; the interpretation of this variety’s data resembles the one provided by Luang-Thongkum (1979).

4.2.2 Diphthongs

Whereas there is no length distinction, phonation is contrastive for onglides and offglides, apart from /ᵉạ/ which always occurs with breathy voice. Examples for contrastive phonation are /liᵃn/ ‘to study’ vs. /ʔ.liᵃ̤ŋ/ ‘backwards, on back’, or /kuᵃt/ ‘small frog’ vs. /kṵᵃt/ ‘bottle’.

The three offglides commence with prominent close vowels [i, ɨ, u] and end on a central or back centralized near-open target vowel [ɐ] indicated through a superscript /ᵃ/. The front and back offglides /iᵃ/ and /uᵃ/ have ongliding counterparts /ᵉa/ and /ᵒa/. These commence with a short close-mid height [ĕ, ɤ̆], and the more prominent target vowel in these rising diphthongs is a clear open central vowel [a]. Waveforms for the minimal pair /krṵᵃng/ ‘city’ and /krᵒạng/ ‘fence’ show that both onglide and offglide are about 140 msec long, but the near-open central target vowel in the offglide (between the cursors) shown in Figure 1 is only about 50 msec long. In contrast, the open central vowel for the onglide in Figure 2 (between the cursors) is more prominent in both amplitude and about 120 msec long. There is an audible difference in quality and length for the first part of the onglides and offglides, seen in the spectrograms for these minimal pairs as well. Note that the rhotics in these tokens are realised as alveolar approximants, not trills.

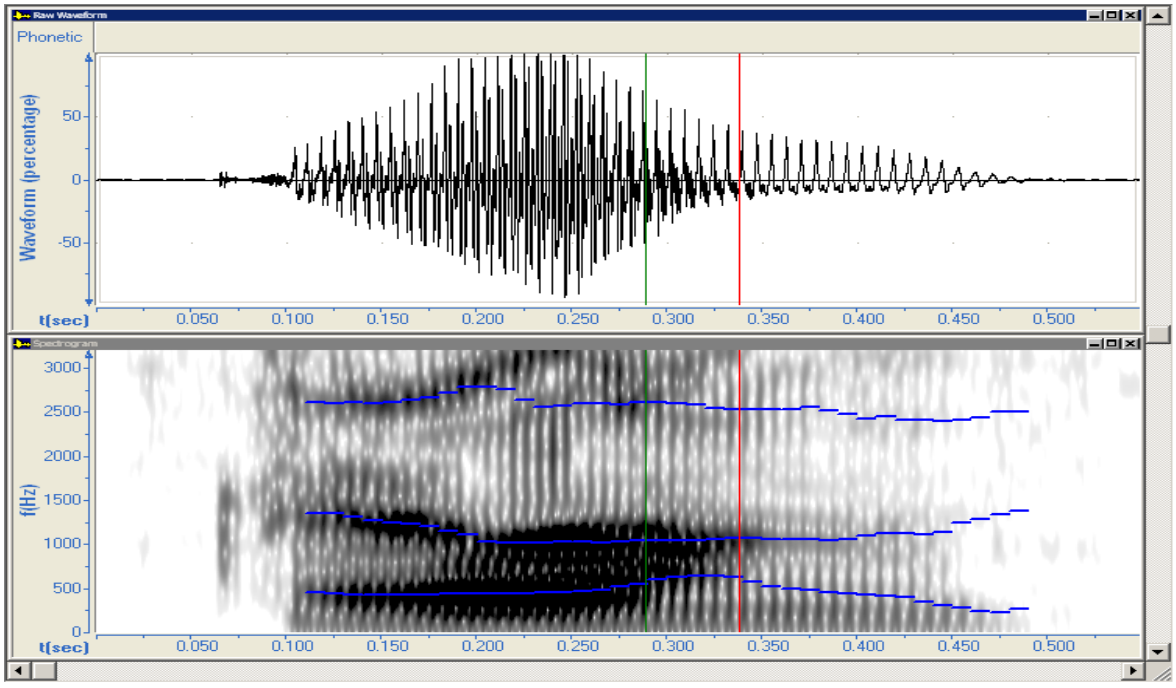


Figure 1: Waveform and spectrogram for back offglide /uᵃ/ in ***krṵᵃng*** ‘city’

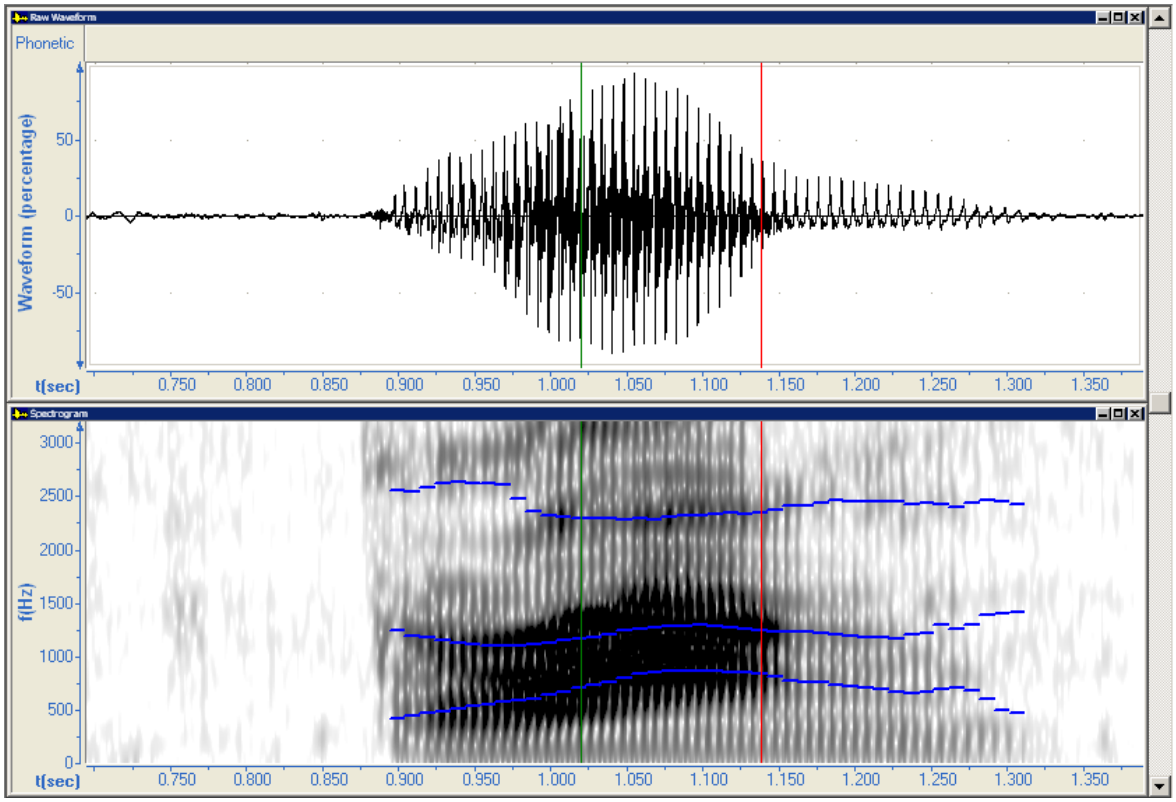


Figure 2: Waveform and spectrogram for back onglide /ᵒa/ in *kroo̤ng* ‘fence’

The front diphthongs are realized in a similar way, seen in Figures 3 and 4 below. Figure 3 with an offglide in /ʔ.li̤ᵊŋ/ ‘backwards, on back’ has a vowel length of about 200 msec (between the cursors), with a commencing close element of about 130 msec, making up more than half of the diphthong. The onglide in /ɬᵉa̤ng/ ‘craftsman’ in Figure 4 is about 260 msec long, with the commencing close element of about 100 msec being less than half as long as the whole vowel. Again, a difference in vowel quality for both elements in the back onglides and offglides is not only heard but also indicated through amplitude and waveform seen in the spectrograms below.

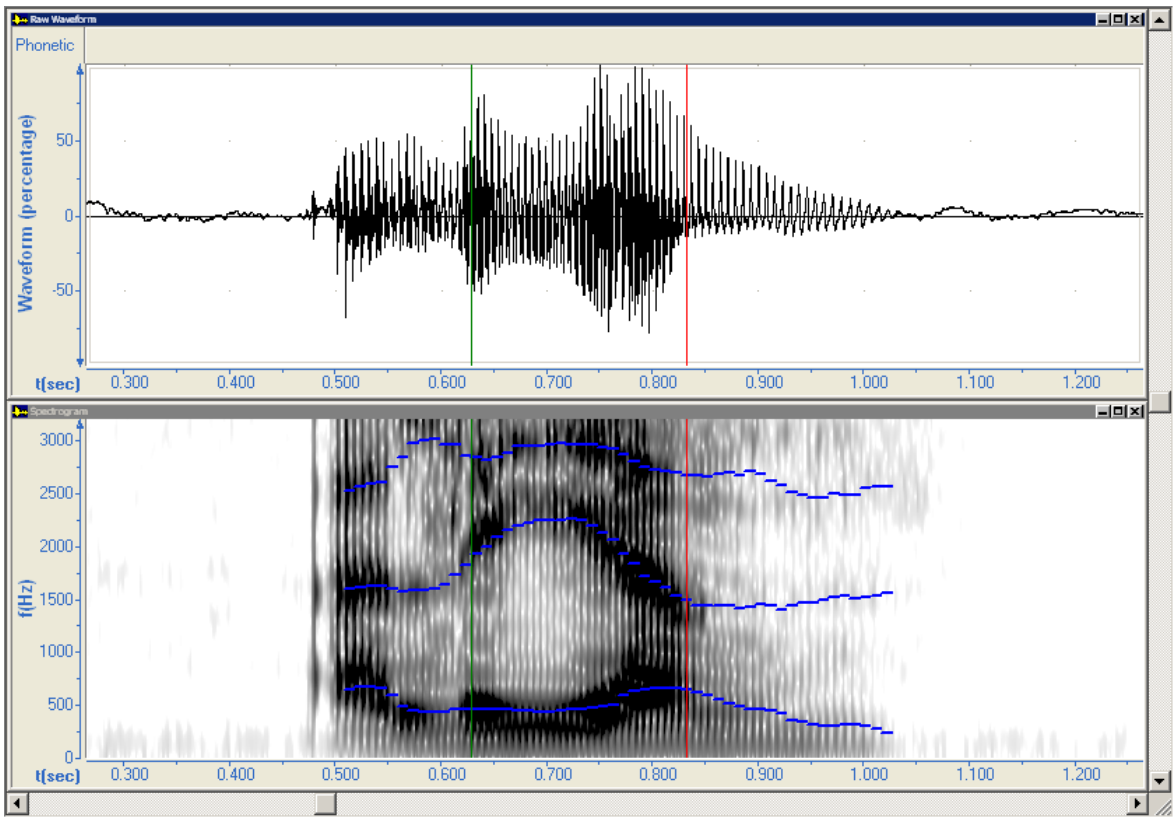


Figure 3: Waveform and spectrogram for front offglide /iᵃ̆/ in ʔ.liᵃ̆ŋ ‘backwards, on back’

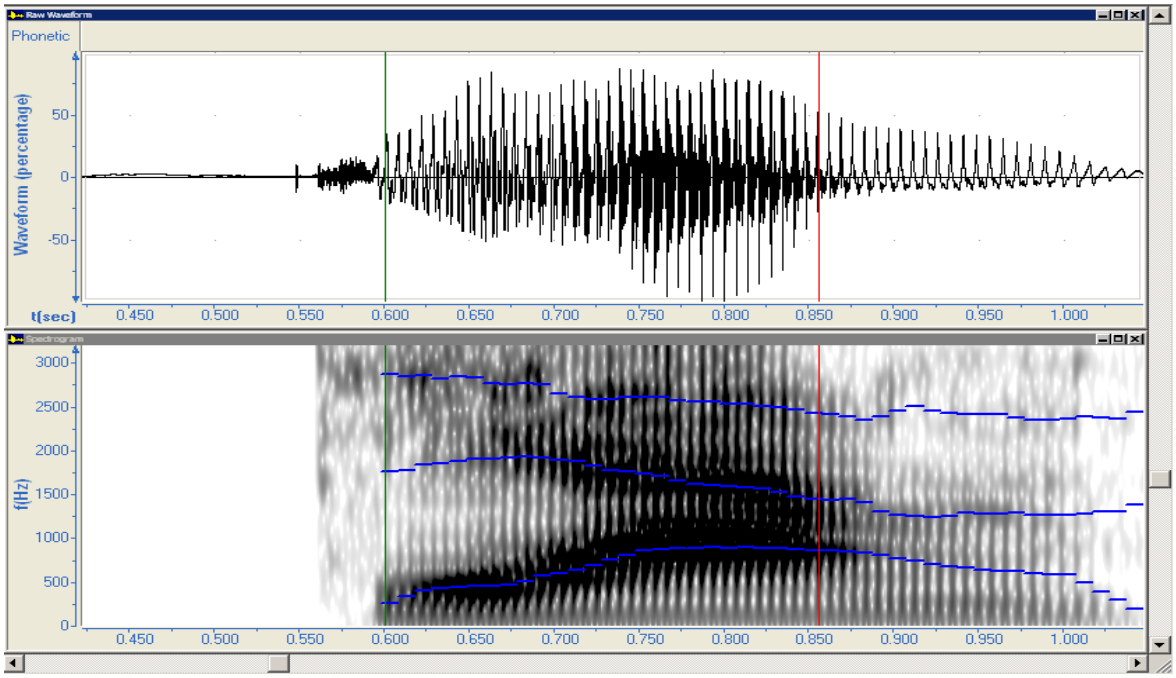


Figure 4: Waveform and spectrogram for front onglide /ᶜa/ in ʈᶜaŋ ‘craftsman’

The close commencing vowels of the offglides are clearly longer than the commencing close-mid vowel of the onglide. It is not possible to interpret the short close element in these diphthongs as a semivowel because it would lead to clusters formed by three consonants which is prohibited by Bru SN syllable structure.

5. Summary and conclusions

The present analysis of Bru SN phonology shares all features described by Luang-Thongkum (1979) for the Western Bru variety Khong Chiam in northeast Thailand, except that Bru SN has only three, not four front and back vowel heights. The other deviation is that nasalization, also reported by Gainey (1985), is not contrastive in Bru SN which supports Green's (1996) observations for Bru Khong Chiam. Vowel quality in reduced syllables is fully predictable, apart from Thai or Lao loanwords with initial sibilants where the original open or near-open central vowel is kept. There are two contrastive onglides as generally reported for Bru languages but the allophonic feature of ongliding for open vowels as a side effect of breathy voice (Luang-Thongkum 1979) or for close and mid tense vowels (Green 1996) is not present in Bru SN.

Onglides appear to be a register phenomenon, liked to phonation and vowel length. The development of onglides cannot be related to merely voicing of initials since onglides are found in both registers for Bru Tri of Vietnam. According to Huffman's (1976) classification of states for Mon-Khmer register development, this would mean that Bru Tri is developing a 'restructured' vowel system with a complete consonant merger. The allophonic development of onglides linked to voice quality in Bru Khong Chiam suggests that this variety, too, is on its way towards a restructured vowel system. Bru SN has no register-related ongliding and is still is a true register language.

The Bru SN vowel system might differentiate more diphthongs than neighboring So; Gainey's (1985) description shows the same inventory of three offglides and two onglides but according to Migliazza (2003), the onglides [^{e}a] and [^{o}a] are not found in this closely related language. A synchronic analysis of So varieties in Thailand and historical-comparative studies are necessary to explain the possible vowel inventory deviation in these otherwise very similar languages.

6. References

Choo, Marcus. 2012. *An investigation of intelligibility between So varieties in Northeast Thailand: The Bru in Khok Saat.* Chiang Mai: Payap University. Online version: http://li.payap.ac.th/images/stories/survey/So_RTT_Report.pdf

Choo, Marcus, Herington, Jennifer, Ryan, Amy, & Simmons, Jennifer. 2012. *Sociolinguistic Survey of Bru in Phan Khon and Phanna Nikhom Districts, Sakon Nakhon Province, Thailand.* Chiang Mai: Payap University. Online version: http://li.payap.ac.th/images/stories/survey/sociolinguistic_survey_of_bru_in_thailand.pdf

Engelkemier, Jennifer M. 2010. *Aspects of Bru Khok Sa-at grammar based on narrative texts.* Unpublished Master's thesis. Chiang Mai: Payap University.

Diffloth, Gerard and Norman Zide. 1992. Austro-Asiatic languages. In: Bright, William et al. (eds.) *International Encyclopedia of Linguistics.* 1, New York: 137-142 Chiang Mai: Payap University thesis.

Gafos, Adamantios I. 1999. *The articulatory basis of locality in phonology*. New York/London: Taylor & Francis.

Gainey, Jerry. 1985. *A comparative study of Kui, Bruu and So phonology from a genetic point of view.* Unpublished Master's thesis. Bangkok: Chulalongkorn University.

Green, Julie. 1996. *A preliminary description of Bru (Khong Chiam) phonology.* Bangkok: Thammasat University and the Summer Institute of Linguistics.

Henderson, Eugénie J. A. 1965. The topography of certain phonetic and morphological characteristics of South East Asian languages. *Lingua 15*:400-434.

Herr, Kristen Elizabeth. 2011. The phonological interpretation of minor syllables, applied to Lemi Chin. Unpublished Master's thesis. Chiang Mai: Payap University.

Huffman, Huffman, Franklin E. 1976. The register problem in fifteen Mon-Khmer languages. *Oceanic Linguistics Special Publications 13*:575-589.

Jongman, Allard and Ratree Wayland. 2002. Registrogenesis in Khmer: A phonetic account. *Mon-Khmer Studies 32*:101-115.

Jonsson, Nanna L. (1991). *Proto-Southwestern Tai.* Unpublished Doctoral dissertation. Albany, NY: State University of New York at Albany.

Keating, Patricia, Esposito, Christina M., Garellek, Marc, Khan, Sameer ud Dowla, Kuangm, Jiangjing. 2010. Phonation contrasts across languages. Los Angeles: *UCLA working paper in phonetics* 108, pp. 188-202.

Ladefoged, Peter and Ian Maddieson. 1996. *The sounds of the world's languages*. United Kingdom: Blackwell Publishing.

Li, Fang-Kuei. 1960. A tentative classification of Tai dialects. In S. Diamond (ed.) *Culture in history: Essays in honor of Paul Radin* (pp. 951-959). New York: Columbia University Press.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.). 2013. *Ethnologue: languages of the world, seventeenth edition*. Dallas: SIL International. Online version: http://www.ethnologue.com/.

Lễ, Vương Hữu. 1999. A new interpretation of the Bru Vân and Kiêu vowel system. *Mon-Khmer Studies 29*:27-106.

Luang-Thongkum, Theraphan. 2001. *Phasaa khong naan chon phau nay khwaeng sekawng laaw tai khwaam ruu phuen thaan kaan wicay lae phannaa* [Languages of the Tribes in Xekong Province Southern Laos: a foundation for research and development]. Bangkok: The Thailand Research Fund.

Luang-Thongkum, Theraphan. 1979. The distribution of sounds in Bruu. *Mon-Khmer Studies 8*:221-293.

Matisoff, J. A. (1989). The bulging monosyllables, or the mora the merrier: echo-vowel adverbialization in Lahu. *South-East Asian Linguistics: Essays in honour of Eugénie JA Henderson*. London: School of Oriental and African Studies.

Matisoff, James A. 1973. Tonogenesis in Southeast Asia. In Larry M. Hyman (ed.) *Consonant Types and Tone* (pp. 71-95). Southern California Occasional Papers in Linguistics, No. 1.

Michaud, Alexis. 2012. Monosyllabicization: patterns of evolution in Asian languages. In Thomas Stolz, Nicole Nau, & Cornelia Stroh (eds.) *Monosyllables: From phonology to typology* (pp. 115-130). Berlin: Akademie-Verlag.

Migliazza, Brian. 2003. So stories: a preliminary analysis of texts in a Mon-Khmer language. *Mon-Khmer Studies 33*:67-158.

Migliazza, Brian. 1998. *A grammar of So: A Mon-Khmer language of Southeast Thailand.* Unpublished PhD Thesis. Bangkok: Mahidol University.

Miller, Carolyn P. 1964. The substantive phrase in Brôu. *Mon-Khmer Studies 1*:63-80.

Miller, John D. 1964. Word classes in Brôu. *Mon-Khmer Studies 1*:41-62.

Miller, John D. 1976. An acoustical study of Brou vowels. *Phonetica 17*:149-77.

Miller, John & Carolyn Miller. 1995. Notes on phonology and orthography in several Katuic Mon-Khmer groups in northeast Thailand. *Mon-Khmer Studies* 24:27-51.

Miller, John and Carolyn Miller. 1995. Trilingual Bru-Thai-English Dictionary. Unpublished Manuscript.

Miller, John & Carolyn Miller. 1996. Lexical comparison of Katuic Mon-Khmer languages with special focus on So-Bru groups in Northeast Thailand. *Mon-Khmer Studies 26*:255-290.

Phillips, Richard, John Miller & Carolyn Miller. 1976. The Brũ vowel system: Alternate analyses. *Mon-Khmer Studies 5*:203-217.

Peiros, Ilia. 1991. Katuic comparative dictionary. *Pacific Linguistic 132*. Canberra: Australian National University.

Pittayaporn, Pittayawat. 2009. *The phonology of proto-Tai*. Unpublished Doctoral dissertation. Cornell: Cornell University.

Shaw, Patricia A. 1993. The prosodic constituency of minor syllables. In Eve V. Clark, Erin Duncan, Donka F. Farkas, and Philip Spaelti (Eds.) *Proceedings of the Twelfth West Coast*

Conference on Formal Linguistics, Vol. 12 *(pp. 117-132).* Stanford: Center for the Study of Language and Information Publications.

Sidwell, Paul. 2005. Proto-Katuic Phonology and the subgrouping of Mon-Khmer languages. In Sidwell, P. (ed) *SEALSXV: papers from the 15th meeting of the Southeast Asian Linguistics Society* (pp.193-204.). Canberra: Pacific Linguistics.

Sidwell, Paul. 2005. *The Katuic Languages: classification, reconstruction and comparative lexicon. Munich:* Lincom Europa.

Sidwell, Paul. 2009. *Classifying the Austroasiatic languages. History and state of the art.* Munich: Lincom.

Svantesson, Jan Olof, & House, David. 2006. Tone production, tone perception and Kammu tonogenesis. *Phonology 23*(2), 309-333.

Tingsabadh, M. R., and Arthur S. Abramson. "Thai." *Journal of the International Phonetic Association* 23.01 (1993): 24-28.

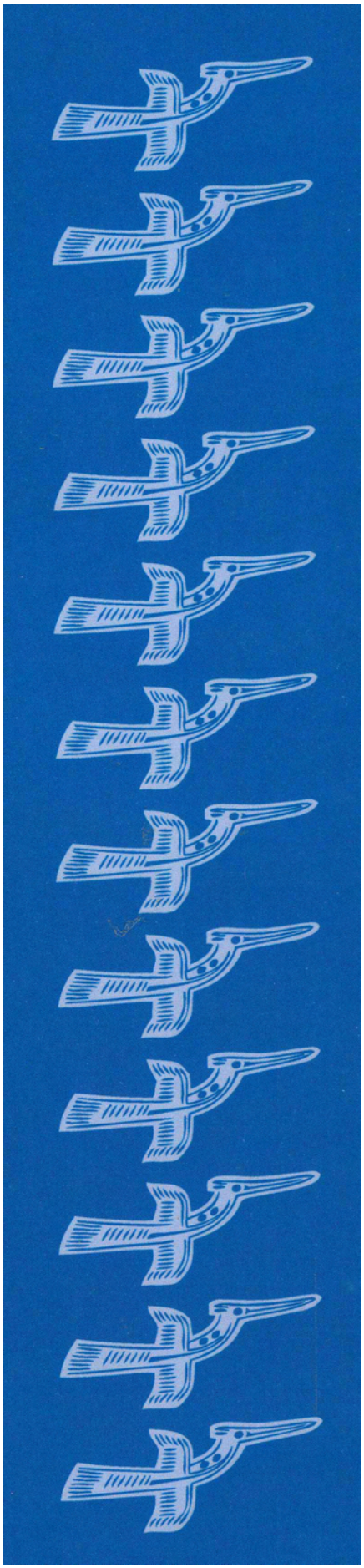
Thomas, David D. 1966. Mon Khmer subgroupings in Vietnam. In N. Zide. (ed.), *Studies in comparative Austroasiatic linguistics* (pp. 1994-202). The Hague: Mouton.

Tebow, Charles T. 2010. *An analysis of participant reference in Bru narrative texts as spoken in the village of Khok Saat.* Unpublished Master's thesis. Chiang Mai: Payap University.

Vương, Hữu Lễ. 1999. A new interpretation of the Bru Van Kiéu vowel system. *Mon-Khmer Studies 29*:27-106.

Mon-Khmer Studies

Volume 42



Author: Jonathan SCHMUTZ
Title: *The Ta'oi Language and People*
Pages: i-xiii (Notes, Reviews, Data-Papers)
Date Received: 4/2/2013
Revised Text Accepted: 23/4/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

The Ta'oi Language and People

Jonathan SCHMUTZ
Payap University

Abstract
This paper provides a review on the linguistic and cultural background of the Ta'oi people in Laos and Vietnam from the available literature. Starting with an overview of the geographic location, historical and cultural context and linguistic nature, the paper pays special attention to the confusing amount of ethnonyms and glossonyms[1] referring to these people and their language.
Keywords: Ta'oi culture, glossonyms
ISO 639-3 language codes: tth, tto

1. Introduction

Katuic languages, belonging to the Mon-Khmer branch of the Austroasiatic language family (cf. Sidwell 2009), are found in Cambodia, Laos, Thailand, and Vietnam. Today, the Katuic population is estimated at over one million people, with the Katuic groups typically being divided into 15-20 distinct language varieties (Sidwell 2005a). Research on Katuic languages includes comparative analyses (Gregerson 1976, Diffloth 1989), historical reconstructions (Thomas 1976, Peiros 1996, Theraphan 2002, Sidwell 2005a,b), and phonologies, e.g. on Pacoh (R. Watson 1964), Katu (Wallace 1969) and Bru Tri (Phillips et al. 1976). Grammar sketches are available as well, e.g. J. Miller (1964) on Bru Tri, Costello (1969) on Katu, S. Watson (1976) and Alves (2006) on Pacoh, and Solntseva (1996) on Ta'oi. Discourse analyses have been performed by Burusphat (1993) on Kui, R. Watson (2000) on Pacoh, and Migliazza (2003) on So. Despite this, there is still much to be learned about the Katuic. In fact, it is still unclear what ethnic and language groupings exist, how much the languages continue to be spoken, and the social and linguistic interactions that take place between the different varieties.

Several sources speak of the minority peoples of Laos or Vietnam in broad terms, grouping the Ta'oi[2] together with other similar minority groups into "Lao Theung" or "Kha" groups. The most thorough source that deals specifically with the Ta'oi is Robert L. Mole's book *The Montagnards of South Vietnam: A Study of Nine Tribes* (1970). Schliesinger (2003), Chazée (2002), Đặng et al. (2000), and Laos's Department of Ethnic Affairs' *The Ethnic Groups in Lao P.D.R.* (2008) each give brief anthropological sketches of all of the people groups of Laos or Vietnam, including the Ta'oi[3]. Linguistic publications on Ta'oi have been limited in scope (Watson 1969, van der Haak 1993, Solntseva 1996). Some word lists have been taken (Nguyễn Văn Lợi et al. 1986, Theraphan 2001, Ferlus ND, Miller 1988), and the Mon-Khmer Languages Project website (sealang.net/monkhmer/) provides wordlists from six different sources comprising a total of 1194 entries. There is also a Vietnamese-Ta'oi-Pacoh dictionary (Nguyễn et al. 1986). As with many Katuic varieties, it is unclear what Ta'oi varieties exist, the extent that they are spoken, and the sociolinguistic relationship between each variety and with closely related languages. Due to infrastructural and administrative limitations, direct access to Ta'oi language communities is difficult. Here a literature review of the geographic, linguistic and anthropological research on the Ta'oi people in Laos and Vietnam will be given. Special attention will then be paid to the vast number of ethno- and glottonyms, followed by recommendations for further research.

2. Geography

The country of Laos is composed of sixteen provinces and one municipality. These provinces are further broken up into 139 districts. The lowest of the administrative divisions in

[1] This paper uses the term 'ethnonym' to refer to the name of an ethnic group, 'autonym' to refer to what speakers call their ethnic group, and 'glossonym' or 'glottonym' to refer the name of a language.

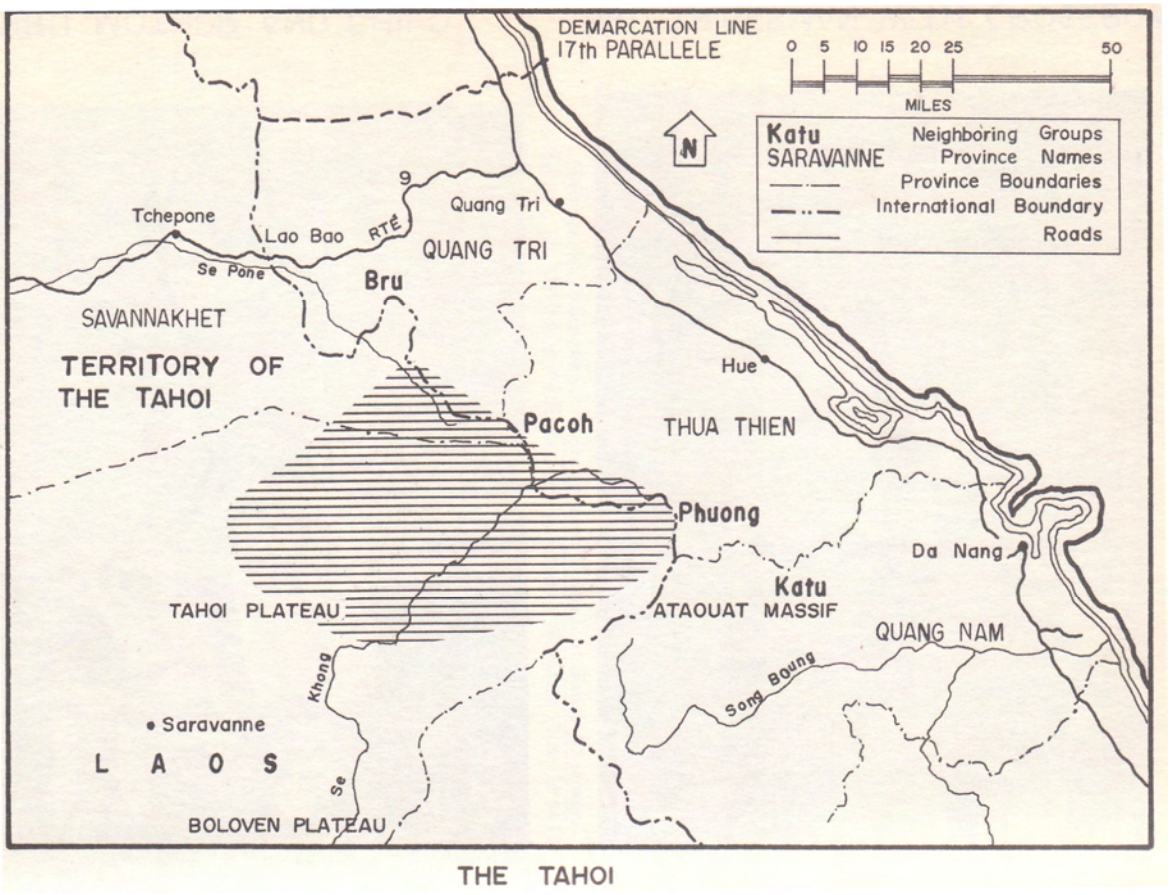
[2] There are many versions of spelling this language, which are discussed in section 6.1 of this paper. This paper adopts the spelling "Ta'oi" except where quoting directly from another source.

[3] Some of these sources mention other texts that may have primary research but were out of print, in a language other than English, or both.

SCHMUTZ, Jonathan. 2013. The Ta'oi Language and People.
Mon-Khmer Studies (Notes, Reviews, Data-Papers). 42:i-xiii

Laos is the village. The total number of villages in Laos is estimated at being between 10,000 and 11,000 (Messerli et al. 2008). There is one informal administrative division called a *khet* ("zone"), which is found between the village and district levels. A *khet* is typically comprised of several villages. There are reported Ta'oi groups located in the Lao districts of Sepone and Nong in Savannakhet Province, Ta Oi, Toumlaan, Salavan, and Lao Ngam in Salavan Province, Bachiang Cehaleunsook, Paksong, and Phathoomphone in Champasak Province, Thateng and Kaleum in Sekong Province, and Sanamxay in Attapeu Province (Steering Committee for Census of Population and Housing 2006).

Vietnam has 58 provinces and 5 municipalities. Provinces are further divided into 548 rural districts and 47 urban districts. These are divided into a further 1448 precincts and 9050 communes (General Statistics Office of Vietnam 2011). In Vietnam, there are reported to be Ta'oi groups living in A Lưới district in Thừa Thiên-Huế province and Hương Hóa district in Quảng Trị province (Đặng et al.2000). The map below (from Mole, 1970) shows the traditional homelands of the Ta'oi people.



Map 1: Taken from *The Montagnards of South Vietnam* by Robert L. Mole, 1970.

Most Katuic groups in Laos are found from Khammuan province south, and Ta'oi is no exception. The southern provinces of Laos for the most part are mountainous along the eastern borders which they share with Vietnam. The land slopes gradually west to the Mekong River, which forms much of the border with Thailand. Plains used for paddy rice cultivation are located in Savannakhet and Champasak. The plains in Savannakhet are watered by the Xe Banghiang River, a major tributary of the Mekong. The Mekong River winds south along the border between Laos and Thailand before cutting eastward through Laos just before Pakse and making its way across Southwest Laos into Cambodia.

Research on provincial accessibility (Messerli et al. 2008) shows that as many as 50% of the locations in Southern Laos are more than 5 hours travel from provincial capitals. Many of these areas have Ta'oi-speaking populations. As Chazée (2002:85) states, "The majority of the Taoy remains isolated from the market and development opportunities."

Traditionally, the Ta'oi in Vietnam are found within 20 miles of the border with Laos in the rugged mountainous area of the Annamite Mountain Chain. The Annamite Mountains in this area are drained on the Vietnamese side by "numerous small streams that become part of the Song Da Giang" (Mole 1970:76) as it travels toward the Pacific. On the Lao side, the mountains give way to the Ta'oi Plateau, and are drained by the Se Pone and the Se Khong which both make their way down to the Mekong River.

3 Sociolinguistic Background and Endangerment

The countries of Laos and Vietnam are both rich in linguistic and ethnic diversity. Laos is known to have languages from the Mon-Khmer, Tai-Kadai, Sino-Tibetan, and Hmong-Mien families, while Vietnam has these as well as Austronesian languages (Lewis 2009). Language contact resulted in broad cross-directional linguistic exchanges (Choo 2009) which is prevalent not only between languages of the same family, but can be seen as strongly linking the minority and national languages. For example, in Laos and Thailand, Huffman (1976) reports vocabulary borrowing up to 20% from Thai and Lao, and he also has Vietnamese at about a 20% cognate level with several Katuic languages. The linguistic borrowings along with the continued rise of the national languages in Laos and Vietnam cause concern about the future of minority speech varieties, including those from the Katuic branch. Despite the Mon-Khmer languages being the most numerous in Laos, Enfield (2006) states that, "One cause for urgency in linguistic research in Laos is language endangerments... almost all are endangered..." (473).

This endangerment can be traced to a myriad of factors, including a desire to better oneself economically by learning the national language and the resettlement of people into villages with multiple ethnicities present. The effects of language learning still need to be studied in depth within the context of Laos and Vietnam. The possibility exists for both bilingualism and the loss of minority language in the subsequent generations with language contact scenarios such as the ones springing up across Southeast Asia. Choo (2009:10) asserts that, "It will not be clear whether frequent contact with the lowland Lao correlates directly with decreased mother tongue vitality until a proper study is done."

From personal observations and communication with Ta'oi speakers, it appears that the Ta'oi language is maintaining vitality in villages where almost all of the people are Ta'oi, but is losing vitality in mixed villages where there are several minority languages as well as native Lao speakers present. In this context, the younger generation is often growing up speaking Lao as their mother tongue, as only some are able to understand or speak the language of their parents.

4 Society and Culture

The Ta'oi tend to organize maximally at the village level, around the family as a cohesive unit. In the past, the extended family would all live in the same long-house, but now there is a trend toward the nuclear family having their own house in the same village or nearby. Schliesinger (2003b:90) states that, "The difference between rich and poor people in Ta Oi society is not great. The Ta Oi have a well-developed spirit of mutual assistance within the community."

The Ta'oi are a patrilineal and patrilocal society, and a new bride will take the lineage of her husband upon marriage. Young people are free to choose whom they want to marry. However, there is a bride price that must be paid for the marriage to take place. If the groom's family is poor, "the dowry may be reduced by consent of the bride's family but it must always include at least one buffalo and some food" (Mole 1970:83). At least in the past, there was polygamy among the Ta'oi with men able to take multiple wives if they could afford the bride price. It is not uncommon for Ta'oi to marry outside of their people group, especially in more recent times. According to Schliesinger (2003b), the Ta'oi are marrying among neighboring groups with greater frequency in order to obtain better farming techniques. Each Ta'oi family has a certain totemic plant or animal that is associated with their lineage. Chazée (2002:56) says that, "Meat or vegetables from the name of one's lineage line are not eaten and the same house is not shared between two persons of different lineages." This is played out in a marriage relationship by the new bride taking on the lineage of her husband's family and thus the totemic symbol.

In the past, men were considered superior to women, but this has changed (Đặng et al. 2000). However, the roles of men and women are quite distinct, with women taking care of most of the activities concerning food including planting and harvesting and meal preparation. The men will hunt and fish, clear jungle, and build buildings as necessary (Mole 1970). These days, the women often give birth in their own houses, but before, they would have to go out into the forest alone to deliver the baby and give initial care unassisted before returning to the village (Schliesinger 2003b).

4.1 Ceremonies, myths and beliefs

The Ta'oi are traditionally an animistic culture who make sacrifices to appease the spirits, or *ʔyaaŋ*, of the village. The two main *ʔyaaŋ* that the Ta'oi are concerned with are the spirit of the rice paddy and the spirit of the sky. These are thought by the people to be the two most powerful spirits that control the fate of the tribe. Spirit houses are set up in the center of the village for both of these spirits, with a sacrificial post connected or nearby where sacrifices are made. Sacrifices range from alcohol, rice, or chickens up to a buffalo. Sacrifices are made to appease spirits, gain their favor, or to secure their cooperation. Mole (1970:87) states that, "The Tau-oi believe that the spirits work through, and control, the various natural forces so that harm or prosperity may be given to an individual or the village as the spirits please." There are also many taboos that are in place so that the spirits will not be offended. If the taboo is broken, a sacrifice must be made to appease that spirit that is thought to have been offended. These sacrifices are made by the village shaman who will also perform any ceremony necessary for the sacrifice.

While sacrifices to appease the spirits can take place at any time, there are two ceremonies that occur on an annual basis. The first takes place in February and is to honor the spirit of the village. The second occurs in October and informs the spirits that they are going to clear new land for rice paddies.

The Ta'oi are thought to practice black magic or sorcery through the use of incantations and spells. This makes them feared by the other people groups in Laos. There are many among the lowland Lao that are afraid to take up posts among the Ta'oi because of this fear, though smelling nice is thought to protect against the spells. Thus, shampoo, toothpaste, deodorant, and talcum powder are all thought to be important when living in Ta'oi areas.

The majority of Katuic ethnographic studies (Chazée 2002, Schliesinger 2003b, Mole 1970) state that the Katuic peoples are animist who believe in the powers of the spirit world. However, Buddhism is gradually influencing the beliefs of the Katuic people, while some communities are turning to Christianity.

4.2 Houses and villages

Ta'oi villages are for the most part found between 300 and 1000 meters above sea level. The Ta'oi will often share their villages with other minority peoples. In their traditional homeland, it is often with the Katu that they share. More recently, many Ta'oi villages have moved down to lower elevations, and they are now found in mixed villages with Katu, Kui, Katang, Alak, Loven, and Lao (Schliesinger 2003b).

Traditionally, Ta'oi villages were either in a circular shape used for defense or had long-houses, "radiating like the spokes of a wheel" (Mole 1970:80). In either case, the center of the village was a communal house that was used for meetings and for guests and a spirit house for the village spirit. Attached to the spirit house would be a pole where sacrifices would be made. In modern times, defense is no longer a consideration and the houses are more often built along a road with smaller houses used more for individual families.

According to Mole (1970), the traditional long-houses could be as big as 600 feet in length. These long-houses would have extended families all living together in the same long-house with a corridor running the length of the house with rooms coming off of one or both sides. The size of the long-house would be determined by the size of the extended family, village space permitting.

4.3 Costume, Crafts, and the Arts

The Ta'oi would traditionally weave cloth of red or blue to use for their various needs. Their traditional costume had, "embroidered patterns similar to those of Kriang and Kahtoo" (Lao for National Construction 2008:118), but is becoming less common. In recent times, the women typically wear the Lao *sinh* with a blouse, while the men wear trousers. Schliesinger (2003b) found that Ta'oi still wear traditional pearl necklaces. The Ta'oi traditionally would file their teeth, tattoo their bodies to ward off evil spirits (Mansfield 2000), and stretch their earlobes, but these practices are now less common.

The Ta'oi also practice the crafts of woodworking, carving, and basket making (Schliesinger 2003b), and are noted for their wooden masks and statues (Chazée 2002). Mole notes that the Ta'oi favor a lizard motif and that it can be found in the design of almost all of their houses. He states, "Sometimes the lizard motif is intricately carved on the ends of the main roof beam as a work of art. In this regard the Tau-oi seem to have a talent of woodworking and skillful carving that surpass that of most other tribes" (Mole 1970:80).

The Ta'oi also have music, dance, and poetry that are unique to their culture. They play instruments such as the bronze gong, khaen[4], and drum at different occasions. According to one source, they have a particular type of song that, "They sing to express their joys and sorrows and, to declare their love" (Đặng et al. 2000:87). These songs are accompanied with musical instruments such as the khaen. They also have poems, folktales, and proverbs that tell about their past, their culture, and their livelihood (Lao Front for National Construction 2008).

4.4 Agriculture and economy

The traditional method of farming uses the shifting or slash-and-burn techniques. In recent years, the government of Laos has been promoting resettlement projects throughout Laos in order to help stop slash-and-burn agriculture as well as the cultivation of opium (Evrard and Goudineau 2004). Some of the resettled communities who have resettled in rice friendly areas have incorporated paddy cultivation into their agricultural practices. Glutinous rice is the most popular crop among the Katuic groups, planted mostly for their own dietary needs and supplemented through hunting and gathering. Other crops include cassava, sweet potatoes, corn, and other vegetables. The Ta'oi have also started growing cash crops such as coffee, tea, soy beans, castor beans, tobacco, sesame, red chillies, fruit, and opium (Mansfield 2000). In addition to crops, the Ta'oi supplement their diet by foraging, hunting, fishing, and raising domesticated animals such as chickens, pigs, and buffalo. In the past, they were even known to hunt and domesticate elephants.

The Ta'oi formerly relied solely on natural conditions for their crops. The crops were watered by rainfall alone, and they would use no other fertilizer besides the ashes from the burning of the remnants of the last harvest. According to Mole (1970), this was because they believed that the spirits of the paddy and the rice would not like it if other fertilizers were used. In more recent times the Ta'oi have moved into some mixed villages, where they are rapidly taking on new farming techniques from other groups. Chazée (2002:85) states that, "The majority of the Taoy remains isolated from the market and development opportunities, but start to mix with other minorities with more productive farming systems. The integration seems rapid, and there is acculturation."

5. History and Migration

The Mon-Khmer people are thought to be the original inhabitants of Southeast Asia. Originally, the Mon-Khmer people, such as the Ta'oi, inhabited more of the lowland regions of Southeast Asia, but were pressed further up into the hills with the expansion of the Lao/Tai groups from what are now the southern provinces of China starting in the 14th and 15th centuries AD. As the Lao peoples moved further and further south into the region, the Mon-Khmer people were forced higher and higher away from the more arable land. This culminated with the 1431 AD capture of the Khmer capital by the Siamese, causing the Khmer peoples to retreat into the more remote hills (Mole 1970).

[4] The khaen is a traditional reed pipe instrument that is used in many parts of Southeast Asia.

The Ta'oi were a warlike people who would often raid the Lao villages, which would cause the Lao to respond in kind. In the late 19th century, the Ta'oi and some of the other tribes took to slaving: "While killing anyone who resisted, the Tau-oi kidnapped women and children of neighboring tribes and supplied Montagnard slaves for the markets at Bassac, Attopeu, Phnom Penh, Bangkok and other trade centers." (Mole 1970:78). The Vietnamese paid tribute to the Ta'oi in order that they might pass through their territory safely. In 1897 the French entered into negotiations with the Ta'oi, ending the slave trade and the violence associated with it.

During the time of colonial rule, the French enacted a policy of what was called a corvée labor system. The men 19-60 years old of the Lao Theung groups were required to pay 1 piastre a year as well as serve 10 days a year laboring for the French (Evans 2002). Parts of the road from Salavan to Ta'oi still have some of the paving stones laid down by these workers to this day (Osborne 2012).

At the turn of the century an indigenous rebel movement directed against the French was growing on the Boloven Plateau. This 'holy man movement' is expounded upon in works such as "The Holy Man in the History of Thailand and Laos" (Wilson 1997). Under the leadership of first Ong Keo and later Ong Kommadam, attempts were made to force out the French. This movement attracted Ta'oi support.

During the Vietnam War, the country of Laos was used as a staging ground and not so secret battleground by both the Pathet Lao and their North Vietnamese allies and the United States. For the most part, the United States limited its involvement to supplying those fighting the communist forces and bombing throughout the country, concentrating on the Ho Chi Minh Trail, which ran south along the border mountains where many of the Katuic people lived. These bombings along with the fighting that took place throughout the country caused many people, including Ta'oi, to leave their homes. Stuart-Fox (1997:144) finds that, "At one time or another as many as three-quarters of a million people, a quarter of the entire population, had been driven from their homes to become refugees in their own country."

During the late 1950's and early 1960's as the anti-royalist forces moved into the mountainous regions of the Annamite cordillera, Ta'oi villages from what is now Ta'oi District were moved away from the advancing forces and re-settled along the roads closer to Pakse. As early as 1967, the administrative center of Salavan, which was the closest to the Ta'oi, was under the control of the Pathet Lao. The de-population of the area was a military strategy designed to deny local food supply or support to an advancing army (Osborne 2012).

Since the end of the war, there have been two factors that have affected the movements of minority peoples. The first is the migrations that have taken place as people look to build better lives for themselves in a different area. Schliesinger (2003b:88) states that, "Since the end of the Vietnam War there is a trend for most Katuic-speaking people who lived near the mountainous, malaria infested, remote and inaccessible Laos-Vietnam border region, to migrate westwards onto the plains as far as the outskirts of Pakse close to the Mekong River."

Many of the villages that have sprung up from this migration are mixed villages, or villages that contain multiple ethnicities. The Ta'oi have tended to form villages with the Katu, Kriang (Ngeq), Katang, and others. This has had the effect of a greater reliance on the Lao language as these groups will use Lao among those outside of their own language community.

The second factor which has affected the movements of minority groups are the resettlement projects that the government of Laos has taken on. Besides creation of infrastructure such as dams, the reasons that are given for these resettlement projects are opium eradication, security concerns, access and service delivery, cultural integration and nation building, and swidden agriculture reduction" (Baird and Shoemaker 2007:870). Because opium has to be grown at higher elevations, the resettlement of villages to the lowlands allows for the eradication of opium production. Resettlement because of security concerns was more valid in the aftermath of the Vietnam War, when rebel groups were more active. In moving these villages out of the highlands, the government hopes to be able to better provide services such as education and health services to villages that are easier to access via the road system. The government is also trying to eliminate traditional slash-

and-burn farming techniques by moving villages to areas where rice paddy farming is a viable option.

The final reason for resettlement is the one that deals most directly with topics covered in this paper, that is, cultural integration and nation building. In resettling minority groups, such as the Ta'oi, into areas that are traditionally ethnic Lao areas, the minority groups are being encouraged to integrate into the wider Lao society and language. As mentioned before, this is true not only in villages where there is an ethnic Lao population, but also where mixed villages are created from several different ethnicities. This creates the need for a lingua franca, which in this case is Lao.

6. Linguistic Background of Ta'oi

The Katuic languages were first listed as a Mon-Khmer branch of its own by Thomas and Headley (1970). This list contains 17 suspected Katuic varieties, and was lexically based as much of the subsequent research has been. As research has increased, different researchers have postulated divisions that have built from the work of Ferlus (1974). For example, the Ethnologue (Lewis 2009) lists 19 Katuic languages according to its classifications, and two of Thomas and Headley's Katuic languages have since been recognized as Bahnaric (Sidwell 2005a). Ferlus (1974) and Therapan (2000) whose classifications are based on lexical considerations distinguish only West and East Katuic, but do not conform in the assignment of individual languages. Miller & Miller (1996) whose divisions are based on a lexicalstatistical analysis of 50 wordlists differentiate North, West, and Central Katuic. Sidwell's (2005a) historical phonological comparison postulates the four Katuic branches West-Katuic, Ta'oi, Katu, and Paco. Table 1 below shows different Katuic groupings and related varieties in each, according to the aforementioned sources.

Table 1: Katuic groupings according to various sources (taken from Choo 2009).

Source	Proposed Katuic grouping
Ferlus 1974	West Katuic: Kui, Souei, Bru, So East Katuic: Katu, Kantu, Phüöng, Ta-Oi, Kriang etc.
Miller & Miller 1996	North Katuic: So, Bru, Tri, Makong, Siliq, Katang West Katuic: Sui/Suoi/Suai, Nheu, Kui, Kuay Pacoh: Pacoh Central Katuic: Ong, Ir, Ta-oih (implied from body of paper) Ngeq: Ngeq Katu (Laos): Katu (Laos) Katu (Vietnam): Katu (Vietnam)
Theraphan 2002	West Katuic: Kui, Souei East Katuic (North): Bru, So, Pacoh East Katuic (Central): Ta'Oi, Chatong, Kriang East Katuic (South): Dakkang, Triw, Kantu, Katu
Sidwell 2005(a)	West Katuic: Kui, Souei, Bru, Sô, etc. Ta'Oi: Ta'Oi, Katang, Talan/Onh/Ir/Inh, Kriang/Ngeq, Chatong Katu: Kantu, Katu, Phuong, Triw, Dakkang Pacoh: Pacoh

Based on these groupings, Ta'oi falls either into the East Katuic (Ferlus), Central Katuic (Miller), East Katuic [Central] (Theraphan), or into a Ta'Oi-Kriang grouping (Sidwell). An in-depth discussion of Ta'oi linguistic classification would exceed the framework of this paper. What remains clear is that further research on the Ta'oi language is needed. One of the biggest areas for further studies is the relationship between Ta'oi dialects, as well as establishing the relationship between Ta'oi, Ong, Ir, Chatong, Katang-Ta'oi, Pacoh, and Cantua.

6.1 Ethnonyms and Glossonyms

The most recent census in Laos took place in 2005 (Messerli et al. 2008), and it lists the population of Laos at 5.6 million people, although since this number is likely to have grown closer to between 6 and 7 million (World Factbook 2012). The 2005 census (National Statistics Centre 2007) states that ethnic Lao compose 55% of the population. The remaining 45% are made up of the ethnic minorities, among which the government of Laos officially recognizes 49 ethnic groups with 160 sub-groupings. These have been ethnolinguistically classified into four families: Lao-Tai (Tai-Kadai), Mon-Khmer (Austroasiatic), Hmong-Mien, and Sino-Tibetan. Officially recognized Katuic languages in Laos are the following varieties: Katang, Makong (including Bru), Tri, Ta'oi, Katu, Kriang, Souay (also known as Kuy), and Pacoh.

One older method of classifying ethnic minorities in Laos that is still sometimes used is based on the geographic altitude at which they typically live, started by P.S. Nginn in the early 1960s, but not widely used until after 1975 (Schliesinger 2003a). It divides the ethnic groups of Laos into three groups: the *Lao Loum* 'Lao below' who traditionally live in the lowlands up to approximately 400m above sea level, the *Lao Theung* 'Lao above' who traditionally lived at the middle altitudes of approximately 800-1400m, and the *Lao Soung* 'Lao high' traditionally lived in the higher mountainous regions, those above 1400m from sea level. For the most part, the groups are broken up ethnolinguistically, with the Lao-Tai groups in the *Lao Loum,* the Austroasiatic (including Katuic) groups in the *Lao Theung,* and the *Lao Soung* being composed of the Hmong-Mien and Tibeto-Burman populations (Chazée 2002). This method is losing relevance as more and more of the people move out of their traditional homelands and into those traditionally occupied by other groups.

According to Chazée (2002), ethnonyms for ethnic groups in Laos are a challenge. Some of the ethnic groups do not have an autonym. This has been found to be especially true with regards to Austroasiatic groups (e.g. Katuic). Even if they do have an autonym, they are often called by a different name by others. One example of this is the Makong. The Ethnologue (Lewis 2009) lists Makong as an alternate name for So. However, the government of Laos puts Makong as a primary ethnonym with sub-groups: Trui, Phoua, Maroih, and Trong; but not So. Studies done by the Nam Theun Project researchers (Ovenden 2007) frequently list the Brou (Bru) as representative of Makong. However, Bru is listed as a distinct variety from So in the Ethnologue; and Bru is not even found in the official list of ethnic groups recognized by the government of Laos[5].

The Ta'oi people are similar to the Makong in having a large number of ethnonyms. In addition to the ethnonyms, there are also the glossonyms (names of the language). For the most part, the ethnonyms and glossonyms are synonymous. However, there are a few exceptions where names of dialects are different from any known ethnonyms. Both ethnonyms and glossonyms from various sources are included in Table 1 below. The Ethnologue (Lewis 2009) has two listings for the Ta'oi, Upper and Lower Ta'oih. The reason for the Upper and Lower Ta'oi distinction is unknown. For Upper Ta'oih, it lists Kantua, Ta Hoi, Ta-Oi, Ta-Oy, and Tau Oi as alternate names, and it lists Pasoom, Kamuan', Palee'n, Leem, and Ha'aang (Sa'ang) as dialects. For Lower Ta'oih it lists the alternate name of Tong and the dialects as being Tong and Hantong'. Mole (1970) lists Tau-oi as having synonyms of Ta Hoi, Tahoi, Ka-Ta-Oi, Ta-oih, and Toi-Oi. The most extensive lists come from Schliesinger who has gathered a number of sources and lists the language as Ta Oi with Taoy, Ta Oy, Ta Oih, Ta Hoi, Ta Uat, Taoey, Tau Oi, Tau-oi, Tauat, Atuat (probably after the Atouat Mountain in Laos), and Ta Liat as alternate names and Ong, Ir (or Yir), Tong, and Hantong as subgroups (Schliesinger 2003a). Also in this book he quotes Nguyen Duy Thieu who also lists Bru, Paco, Oong, In, Canay, Cado, Zir, Toong, Kha Paco, T'rau, and Lao Thong as other names of Ta'oi. In his notes, Thieu states that, "Oong means mountain, Canay means mouse, Cado means wild banana and Toong is a village name" (Schliesinger 2003a:90). In Vietnam, the name Ta'oi is also used for the Paco (Pacoh, Pako), Can Tua, and Ba Hi people (Đặng et al. 2000). Sidwell (2005a) has the name as Ta'Oi with alternate spellings of Taoih, Ta-Oy, and Ta Hoi, and lists Ong/Ir/Talan

5 See also Enfield (2006: 486), who shares a similar problem of ambiguity in identifying ethnonyms Perhaps coincidentally, the example he gives is also Brou-Makong-So.

as different names for one dialect and Chatong as another dialect. Names of Ta'oi found in the literature are listed in Table 2 below which shows the different names used by the various sources.

Table 2: Ta'oi ethnonyms and glossonyms from published literature

Source	Primary Name Used	Additional Ethnonyms Listed	Additional Glossonyms listed
Ethnologue (2009)	Ta'oih	Upper: Kantua, Ta Hoi, Ta-oi, Ta-Oy, Tau Oi Lower: Tong	Upper Dialects: Pasoom, Kamuan', Palee'n, Leem, Ha'aang (Sa'ang) Lower Dialects: Tong, Hantong'
Mole (1970)	Tau-Oi	Ta Hoi, Tahoi, Ka-Ta-Oi, Ta-oih, Toi-Oi	
Schliesinger [2](2003a and 2003b)	Ta Oi	Ta-Oi, Ta-oi, Ta Oy, Ta oy, Taoy, Taoey, Ta Oih, Ta-oih, Ta Hoi, Thoi, Ta Uat, Tauat, Tai-Oi, Tai Oih, Tau Oi, Tau-Oi, Tau-oi, Kha Ta Hoi, Kha Tahoi, Atuat, Bru, Cado, Canay, In, Kantua, Kha Paco, Lao Thong, Oong, Paco, Ta Liat, Zir, Toong, T'rau, Ting Subgroups: Hantong, Ir (or Yir), Ong, Tong	
Chazée 2002	Taoy	Ta-oih, Ta-oy, Ta Hoi	
Đặng et al. (2000)	Ta-ôi	Tôi-ôi, Ta-ôih, Ta-hoi, Tà-uất (Atuất) Subgroups: Pa-cô, Can-tua, Ba-hi	
Solntseva (1996)	Taoih	Ta-ôih, Ta-uôih, Ta-uôt, Pa-koh, Ba-hi, Pa-hi	Ta-ôih, Ta-uôih, Ta-uôt
van der Haak (1993)	Ta'uaih		Ta'oih, Katang-Ta'oih
Sidwell (2005a)	Ta'Oi	Taoih, Ta-Oy, Ta Hoi	Dialects: Ong/Ir/Talan, Chatong

While there are some differences between the ethnonyms and glossonyms, the names used for both the people and their language can be grouped into basic categories: Ta'oi (plus variations), names that are thought to be related languages, and names that are suspect that come from a single original source. The first category is "Ta'oi" plus variations. This would include what are thought to be the two main dialects of Ta'oiq and Ta'uas or Ta'uaih (van der Haak 1993). There is a third group that van der Haak calls Katang-Ta'oih that may be a dialect of Ta'oi, Katang, or it may be another language entirely. Up until this point there has not been sufficient research to determine the relationships between Ta'oi, Katang, and Katang-Ta'oi.

In the second grouping of names that are thought to be from related languages, we have Bru, Ong variations (Tong, Hantong), Ir variations (Yir, In, Zir), and different spellings of Pacoh. Pacoh at least is a different language (Alves 2006) and Bru is sometimes used as a term for a larger section of the Katuic population. In Vietnam, the Pacoh are included under the umbrella of Ta'oi, as are Can-tua and Ba-hi (Đặng et al. 2000). According to Richard Watson, Pahi (Ba-hi), Kado (Cado), and Pacoh were dialects of the same language, although Kado, at least, has grown apart to the extent that it is no longer mutually intelligible with Pacoh, except for those people who have a lot of contact. Cantua is a Pacoh name for Ta'oi, though it is unknown whether it is actually a dialect within Ta'oi (personal communications). There seems to be at least an ethnographic difference between the Ta'oi and the Ir and Ong, and they seem to consider themselves different

[6] Schliesinger draws from many sources. He has tables that include ethnonyms from each source as well as a table that conglomerates most of the others in the back of the first volume of his Ethnic Groups of Laos. All quoted sources have been added to this table.

groups. Whether their languages are separate languages or just dialects of a single language remains to be determined. One final glossonym of note is Chatong, which Sidwell (2005a) reports as being a member of the Ta'oi subgroup, which also contains Kriang (Ngeq) and Ta'oi. Other names that are listed were found only from a single source and are suspected to be names of villages or other geographical areas rather than actual ethnonyms or glossonyms. From the above sources, the following divisions in Table 3 are therefore suggested.

Table 3: Dialects of Ta'oi

Primary Dialects:	**Dialects in Need of Data:**	**Related but Separate Languages or Dialects**
Ta'oiq	Katang-Ta'oi	Bru
Ta'uas	Ong	Pacoh
	Ir	Pahi
	Cantua	Cado
	Chatong	

7. Outlook

The Mon-Khmer language Ta'oi in the Katuic sub-group is spoken in Laos and Vietnam. It has many names, and further research is needed to determine the relationships especially between Ta'oi, Katang, and Katang-Ta'oi. As an endangered language, Ta'oi would benefit greatly from phonological and grammatical descriptions, an orthography, and literature development, which may help prevent extinction. In addition, an updated anthropological study with historical data is recommended as much has changed in recent years for many minority groups in Laos and Vietnam, including the Ta'oi.

References

Alves, Mark J. 2006. A Grammar *of Pacoh: A Mon-Khmer Language of the Central Highlands of Vietnam*. Canberra, Pacific Linguistics.

Baird, I. and B. Shoemaker. 2007. Unsettling experiences: Internal resettlement and international aid agencies in Laos. *Development and Change* 38: 865–888. http://www.gapeinternational.org/wp-content/uploads/2011/08/Unsettling-Experiences.pdf (Accessed 17 December 2012).

Burusphat, Somsonge. 1993. Kui narrative repetition. *Mon-Khmer Studies (MKS)* 22: 149-162. http://sealang.net/sala/archives/pdf8/somsonge1993kui.pdf (Accessed 20 December 2012).

Chazée, Laurent. 2002. *The Peoples of Laos: Rural and Ethnic Diversities*. Bangkok, White Lotus Press.

Choo, Marcus. 2009. Payap University unpublished document.

Costello, Nancy. 1969. The Katu noun phrase. *MKS* 3:21-35. http://sealang.net/sala/archives/pdf8/costello1969katu.pdf (Accessed 20 December 2012).

Đặng Nghiêm Vạn, Chu Thái Sơn, and Lưu Hùng. 2000. *Ethnic minorities in Vietnam*. Hanoi, Thế Giới Publishers.

Diffloth, Gérard. 1982. Registres, dévoisement, timbres vocaliques: leur histoire en Katouique. *MKS* 11:47-82.

Diffloth, Gérard. 1989. Proto-Austroasiatic creaky voice. *MKS* 15:139-154. http://sealang.net/sala/archives/pdf8/diffloth1986proto.pdf (Accessed 20 December 2012).

Enfield, Nick J. 2006. Languages as historical documents: the endangered archive in Laos. *South East Asia Research* 14.3: 471-488. http://coreservice.mpdl.mpg.de:8080/ir/item/escidoc:60114/components/component/escidoc:60115/content (Accessed 20 December 2012).

Evans, Grant. 2002. *A Short History of Laos: The Land in Between*. Chiang Mai, Silkworm Books.

Evrard, Olivier and Yves Goudineau. 2004. Planned resettlement, unexpected migrations and cultural trauma in Laos. *Development and Change* 35.5:937-962.

http://laos.efeo.fr/IMG/pdf/EFEO_Resettlement_evrard_goudineau.pdf (Accessed 7 December 2012)

Ferlus, Michel. ND. Lexique comparatif:Ta-Oy, Ong, Katang, Krieng. SEAlang Mon-Khmer Database Project. http://sealang.net/monkhmer/database/ (Accessed 7 December 2012).

Ferlus, Michel. 1974. Lexique Souei-Franç. *Asie du Sud-Est et Monde Insulindien* 5.1:141-159.

General Statistics Office of Vietnam. 2012. Number of administrative units as of 31 December 2011 by province. http://www.gso.gov.vn/default_en.aspx?tabid=466&idmid=3&ItemID=12833 (accessed 12 December 2012).

Gregerson, Kenneth J. 1976. Tongue-root and register in Mon-Khmer. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta eds., *Austroasiatic Studies Volume I*, pp. 323-370. Honolulu: The University Press of Hawaii. [Paper originally presented at 1st International Conference on Austroasiatic Linguistics (ICAL), Honolulu, Hawaii, 2-6 January 1973].

Huffman, Franklin E. 1976. The relevance of lexicostatistics to Mon-Khmer languages. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta eds., *Austroasiatic Studies Volume I*, pp. 539-574. Honolulu: The University Press of Hawaii. [Paper originally presented at 1st International Conference on Austroasiatic Linguistics (ICAL), Honolulu, Hawaii, 2-6 January 1973]. http://sealang.net/sala/archives/pdf8/huffman1976relevance.pdf (Accessed 21 December 2012).

Lewis, M. Paul (ed.), 2009. *Ethnologue: Languages of the World, Sixteenth Edition*. Dallas: SIL International. http://www.ethnologue.com/ (Accessed 21 December 2012).

Lao Front for National Construction. 2008. *The Ethnic Groups in Lao PDR*. Vientiane: Department of Ethnic Affairs.

Mansfield, Stephen. 2000. *Lao Hill Tribes: Traditions and Patterns of Existence*. New York: Oxford University Press.

Messerli, Peter. 2008. Socio-economic atlas of the Lao PDR – an analysis based on the 2005 population and housing census. In Peter Messerli, Andreas Heinimann, Michael Epprecht, Phonesaly Souksavath, Thiraka Chantalanouvong, and Nicholas Minot eds. Swiss National Center of Competence in Research (NCCR) North-South, University of Bern, Bern and Vientiane: Geographica Bernensia.

Migliazza, Brian. 2003. So stories: a preliminary analysis of texts in a Mon-Khmer language. *MKS* 33: 67-158. http://sealang.net/sala/archives/pdf8/migliazza2003so.pdf (Accessed 21 December 2012).

Miller, John and Carolyn Miller. 1996. Lexical comparison of Katuic Mon-Khmer languages with special focus on So-Bru groups in Northeast Thailand. *MKS* 26: 255-290. http://sealang.net/sala/archives/pdf8/miller1996lexical.pdf (Accessed 21 December 2012).

Miller, John. 1964. Word classes in Brôu. *MKS* 1: 41-62. http://sealang.net/sala/archives/pdf8/miller1964word.pdf (Accessed 21 December 2012).

Miller, John . 1988. Ta-oih - Saravan Wordlist. Unpublished ms.

Mole, Robert L. 1970. *The Montagnards of South Vietnam: A Study of Nine Tribes*. Tokyo: Charles E. Tuttle.

Osborne, Mark. 2012. Personal communications.

Ovenden, M. 2007. Social development plan, volumes 2 & 3. Nam Theun 2 Power Company. Vientiane: Nam Theun 2 Power Company. http://www.namtheun2.com/index.php?option=com_content&id=111 (Accessed 3 January 2013.

National Statistics Centre. 2007. Chapter 1: population size and composition. Vientiane: Lao Department of Statistics. http://www.nsc.gov.la/Products/Populationcensus2005/PopulationCensus2005_chapter1.htm (Accessed 12 October 2009).

Nguyễn Văn Lợi, Đoàn Van Phúc, Phan Xuân Thàn. 1986. *Sách hoc Tiêng Pakôh-Taôih*. Hanoi: Tinh Bình tri Thiên.

Peiros, Ilia. 1996. *Katuic Comparative Dictionary*. Pacific linguistics series C-132. Canberra: Australian National University.

Phillips, Richard, John Phillips and Carolyn Miller. 1976. The Bru vowel system: alternate analyses. *MKS* 5: 203-217. http://sealang.net/sala/archives/pdf8/phillips1976bru.pdf (Accessed 21 December 2012).

Schliesinger, Joachim. 2003a. *Ethnic Groups of Laos, Volume 1: Introduction and Overview*. Bangkok: White Lotus Press.

Schliesinger, Joachim. 2003b. *Ethnic Groups of Laos, Volume 2: Profile of Austro-Asiatic Speaking Peoples*. Bangkok: White Lotus Press.

SEAlang: Mon-Khmer Languages Project. Paul Sidwell and Doug Cooper, Project Directors. http://sealang.net/monkhmer/ (Accessed 3 January 2013).

Sidwell, Paul. 2005a. *The Katuic Languages: Classification, Reconstruction & Comparative Lexicon*. Pre-Publication Draft. 307 pp. http://www.academia.edu/1540083/The_Katuic_Languages_classification_reconstruction_and_comcomparat_lexicon (Accessed 21 December 2012).

Sidwell, Paul. 2005b. Proto-Katuic phonology and the sub-grouping of Mon-Khmer languages. *Papers from the 15th annual meeting of the Southeast Asian Linguistics Society (SEALS) 2003*, ed. by Paul Sidwell, 193-204. Canberra: Pacific Linguistics, the Australian National University.

Sidwell, Paul. 2009. *Classifying the Austroasiatic Languages: History and State of the Art*. Munich: Lincom Europa.

Solntseva, V. Nina. 1996. Case-marked pronouns in the Taoih language. *MKS* 26: 33-36. http://sealang.net/archives/mks/pdf/26:33-36.pdf (Accessed 21 December 2012).

Steering Committee for Census of Population and Housing. 2006. Results from the population and housing census. 2005. Vientiane.

Stuart-Fox, Martin. 1997. *A History of Laos*. Cambridge: Cambridge University Press.

Thomas, David D. and Robert K. Headley Jr. 1970. More on Mon-Khmer subgroupings. *Lingua* 25: 398-418.

Thomas, Dorothy M. 1976. *A phonological reconstruction of Proto-East-Katuic*. MS thesis. University of North Dakota.

Theraphan L-Thongkum. 2001. *Languages of the Tribes in Xekong Province, Southern Laos*. The Thailand Research Fund. Bangkok, Chulalongkorn University.

Theraphan L-Thongkum. 2002. The role of endangered Mon-Khmer languages of Xekong Province, southern Laos, in the reconstruction of Proto-Katuic. In *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society*, edited by Marlys Macken. 407-430. Arizona State University, Program for Southeast Asian Studies.

Van der Haak, Feikje. 1993. Ta'uaih phonology and orthography: a preliminary statement. *MKS* 22: 257-264. http://sealang.net/archives/mks/pdf/22:257-264.pdf (Accessed 21 December 2012).

Wallace, Judith M. 1969. Katu phonemes. *MKS* 3: 64-73. http://sealang.net/archives/mks/pdf/3:61-75.pdf (Accessed 3 January 2013).

Watson, Richard. 1964. Pacoh phonemes. *MKS* 1: 135-148. http://sealang.net/sala/archives/pdf8/watson1964pacoh.pdf (Accessed 3 January 2013).

Watson, Richard. 1969. A note on Ta-oi, Nge', and Nyaheun. *MKS* 3: 130. http://sealang.net/sala/archives/pdf8/watson1969note.pdf (Accessed 3 January 2013).

Watson, Richard. 2000. Pakoh expository discourse. *Proceedings of the 5th International Symposium on Languages and Linguistics (ISLL*), 354-369. Ho Chi Minh City: Vietnam

National University. http://sealang.net/sala/archives/pdf4/watson2000pakoh.pdf (Accessed 3 January 2013).

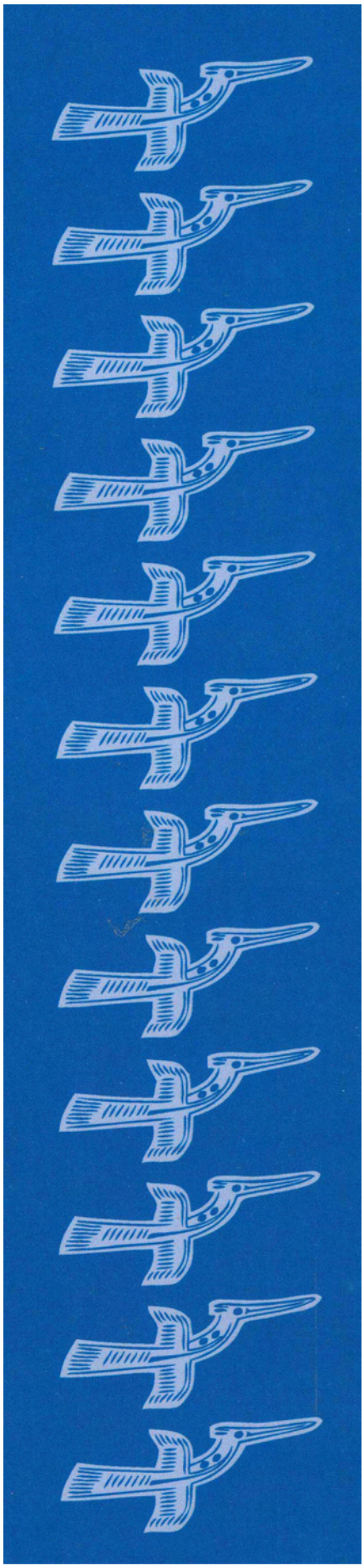
Watson, Saundra K. 1976. The Pacoh noun phrase. *MKS* 5: 219-231. http://sealang.net/sala/archives/pdf8/watson1976pacoh.pdf (Accessed 3 January 2013).

Wilson, Constance M. 1997. The holy man in the history of Thailand and Laos. Journal of Southeast Asian Studies. 28.2:345-64.

World Factbook. 2012. Laos. Washington, D.C.: Central Intelligence Agency. https://www.cia.gov/library/publications/the-world-factbook/geos/la.html (Accessed 3 January 2013).

Mon-Khmer Studies

Volume 42



Author: Darren C. GORDON
Title: *A selective Palaungic linguistic bibliography.*
Pages: xiv-xxxiii
Date received: 4/6/2013
Revised text accepted: 18/10/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

A selective Palaungic linguistic bibliography

Darren C. GORDON
Simpson University

Abstract
This paper is an effort to present a selective bibliographic compilation of Palaungic linguistic resources, as well as relevant cultural resources, totaling 341 bibliographic entries. It is expected that the resulting catalogue will list a significant portion of materials pertaining to Palaungic languages. However, this collection of resources should be considered a living document. It is assumed that there are other available resources still to be included, as well as new resources to be added. Exploring the scope of available Palaungic materials provides a forum through which those who work with or have an interest in Palaungic languages, and the people who speak them, can consult, utilize, and contribute together.
Keywords: Palaungic, bibliography
ISO 693-3: vwa, bgk, blr, bvp, cno, dnu, huo, xko, kkn, lbn, lwl, lcp, zng, mml, mqt, pce, rbb, pll, pnx, ril, stu, tlq, uuu, prk, wbm, yin

1. Introduction

This project began as research and compilation of published materials concerning Palaungic languages while I was teaching at Payap University, Chiang Mai in 2006. This research was originally published as a working paper (Research Project #206) at Payap University, July 2006. Much of the research was conducted at Payap University, Chiang Mai University, David Thomas Library in Bangkok, online, and through colleagues working with or with knowledge of Palaungic languages and potential resources. Since that time it has been revised minimally, both in 2007 and in 2009. However, both the original working paper and the subsequent revisions were not widely available to researchers working with Palaungic languages. In order to make this bibliography of Palaungic languages more accessible for use as well as amendable to the contributions from a great many others, the bibliography has been revised and presented here.

In addition to making this research more accessible, I want to make explicit two other primary reasons for this research. First, this project is aimed at benefiting those who work among Palaungic peoples; a benefit in both knowing what materials are available as well as identifying what areas of research may still be lacking. It is hoped that the bibliographic information presented will become a helpful resource for those who are working with or have an interested in Palaungic languages. Secondly, this project is intended to highlight, as others have, the need for further exploration as to the scope of the Palaungic language family. There has been increasing work in this regard (Sidwell 2009, 2011), but significant questions remain.

The bibliographic compilation presented here is an attempt to capture published sources of language-related work relevant to Palaungic languages. Many times this means that they are the primary focus of the research, but other times the focus on Palaungic languages within the research is secondary, though considered to be of interest and relevance to the Palaungic researcher. Besides linguistic research, some cultural studies are also included, especially when it seems that there is relevant language embedded within the cultural research presented (e.g. Sprenger regarding Lamet) Additionally, most of the sources collected have come from English, with only a few from French, German, Thai, Lao, and Chinese. This is seen, not as a lack of resources in these languages, only a deficiency of the researcher. It is hoped that even with this published bibliography that more non-English sources could be offered as suitable for future inclusion and made available.

2. Palaungic languages

Palaungic languages comprise one branch of the Austroasiatic language family found interspersed throughout Mainland Southeast Asia as depicted in figure 1. The identification of the related languages known as Palaungic find their beginning with Schmidt (1904, 1906). His identification listed four language clusters: Palaung, Wa, Riang, and Danaw. Later, Sebeok (1942) identified these languages as Salowen Basin, totaling five language clusters, adding Khamûk (or

GORDON, Darren C. 2014. A selective Palaungic linguistic bibliography.
Mon-Khmer Studies (Notes, Reviews, Data-Papers). 42:xiv-xxxiii

Khmu) and Le-met, yet leaving out Danaw. The inclusion of Lamet is later affirmed by Diffloth (1977a) and Mitani (1978) and is undoubtedly the result of Izikowitz's anthropological work among the Lamet, though Sebeok neglects to identify his work in his bibliography. A decade later, Shafer (1952) also presents Palaungic as comprising 5 language clusters. His list is identical to Schmidt, with the inclusion of Angkou (after Palaung) and moving Riang before Palaung (ostensibly to show a greater relationship between them). Pinnow (1959) increases the Palaugnic language clusters to six with the only difference being the inclusion of Lawa.

Figure 1: Palaungic languages (in yellow) are located in northern Thailand, Myanmar, southern China and Laos. Source: Adapted from Encyclopædia Britannica (1997), fair use for research purposes

A growth of interest in Palaungic languages in the 1960s and 1970s resulted in more classifications being offered for these languages (Thomas 1964, Thomas 1969, Thomas and Headly 1970, Thomas 1973, Diffloth 1974, Ferlus 1974, Diffloth 1977, Mitani 1978). Much of this work was conducted using a lexicostatisical methodology, as historical reconstruction was rather underdeveloped. Perhaps the most important paper of this period was Diffloth (1977), which was then followed by an equally significant monograph length study three years later (1980). The former presents a classification of Palaungic languages based on select phonological developments, while the latter reconstructs the lexicon and phonology of the Waic sub-branch of Palaungic languages.

Diffloth's (1977) classification (see Figure 2) is a strongly nested tree with the highest branch separating Danaw from the rest of the branch, in this respect repeated by Sidwell (2011) also based upon historical phonology. Later, for reasons that are not explained, Diffloth (1982) revised back his classification, demoting Danaw to a sister of Palaung-Riang, and recapitulating Mitani's (1978) division of Palaungic into Eastern and Western sub-branches (see Figure 3).[1]

[1] The author wishes to acknowledge the contribution of Dr. Paul Sidwell in the understanding and writing of this section, analysing the research of this era as it pertains to the development of Palaungic classification.

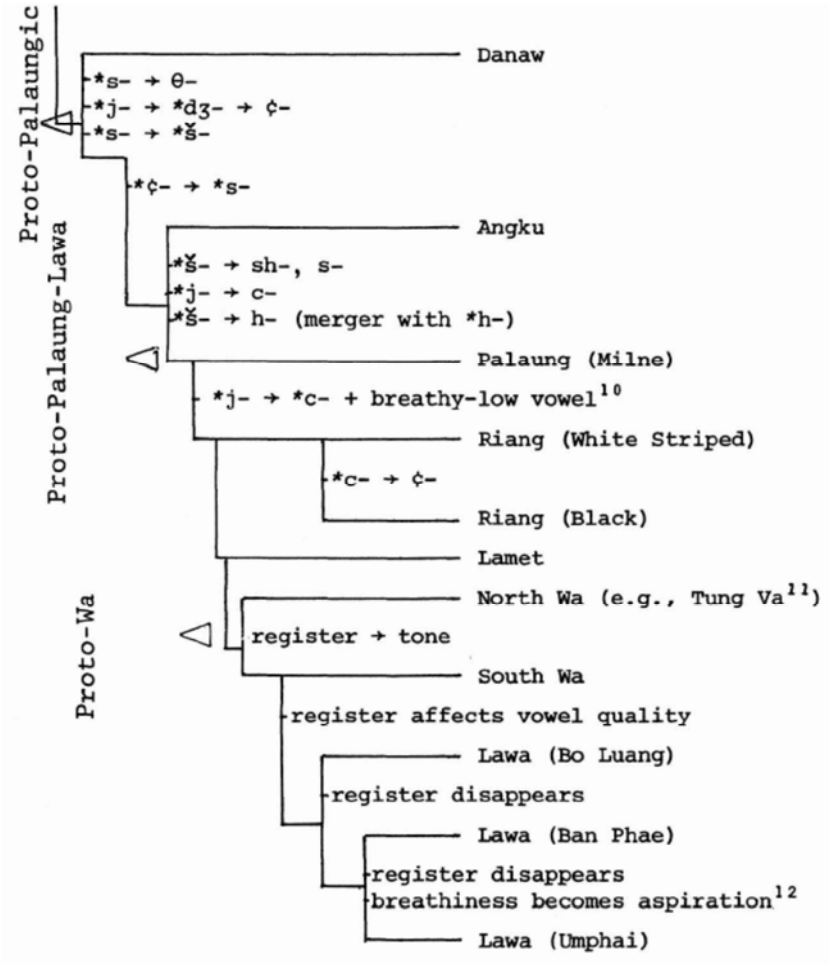


Figure 2: *Palaungic classification of Diffloth (1977) with historical phonological justifications.*

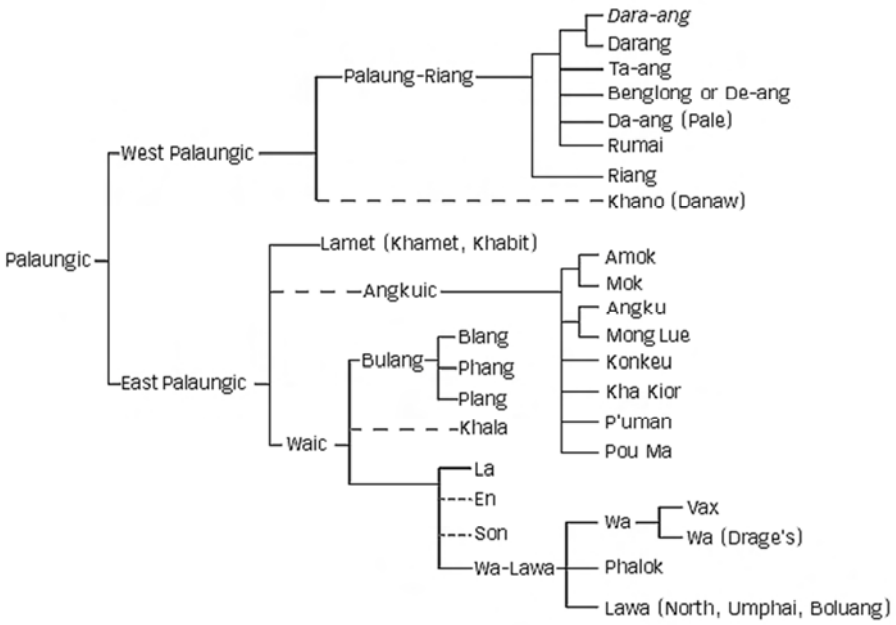


Figure 3: *Palaungic classification of Diffloth (1982) with minor modification by Kasisopa (2003), reproduced from Deepadung (2009).*

In order to emphasize the difficulty that yet remains among researchers attempting to delineate and classify Palaungic languages, there are two other recent classifications given for comparison in figures 4 and 5.

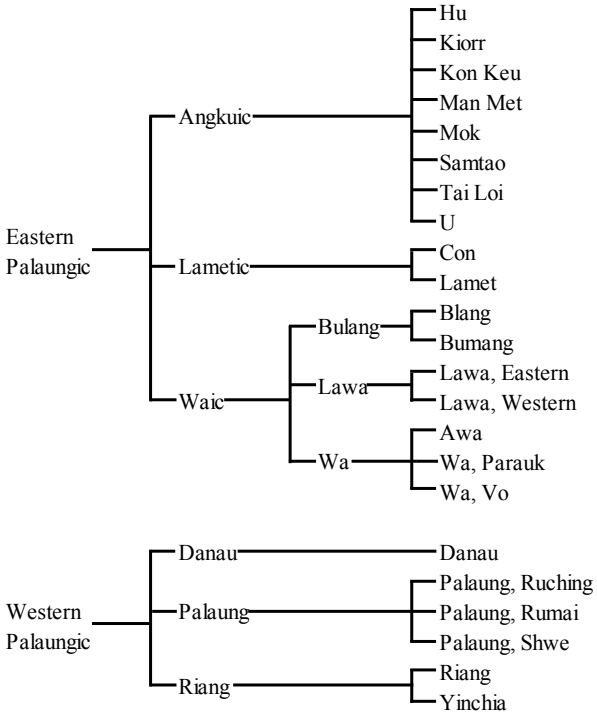


Figure 4: *Palaungic Classification based on Palaungic languages listed in Lewis, Simons, and Fennig (2013).*

As can be seen from the classifications presented by Sidwell (2011) and Lewis, Simons, and Fennig (2013), divergent issues such as whether Angkuic should be considered Eastern or Western Palaungic, and the position of Danaw, reveal a significant need for continued comparative research on the Palaungic languages. This is further revealed in the differences of Palaungic subgroupings as found in the Khasi-Palaungic relationship suggested by Sidwell (2011). It has also been suggested (e.g. Kingsada 2003) that Mang, along with the possibility of Khang, comprise a North Palaungic group and, additionally, Sidwell (2009:132-133) discusses proposals to link Pakankic (Bolyu, Bugan, Mang (?)) with Palaungic.

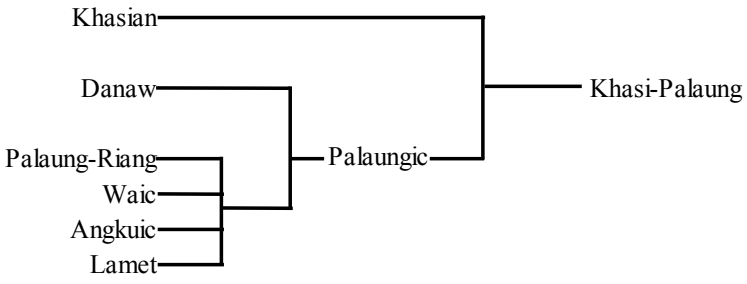


Figure 5: *Khasi-Palaung relationship as proposed by Sidwell (2011).*

The incongruity of both the internal and structural classification of Palaungic languages makes the work of bibliographical compilation more complicated. Since, even up to the present, there still remains a variance as to which languages are identified and included in the Palaungic branch, one must still decide which language will be considered for inclusion within the bibliography. As a guide, the most recent Palaungic classifications, although divergent, present a fairly reasonable framework for these decisions. Other languages that are only singularly

mentioned by one researcher as Palaungic have not been considered (as an example, Schliesinger (2003) identifies Keu (Akeu) as Palaungic whereas most other researchers identify this language as Tibeto-Burman (i.e. Chazée 1999)). In other cases, there are languages, such as Bid (Phsin, Bit, or Kha Bit), that are listed by some researchers as Palaungic (Schliesinger 2003, Proschan 1996, Ferlus 1996) and omitted by others (Parkin 1991). When there are several researchers who consider the possible inclusion of a language (often lesser known) as being Palaungic (e.g. Mang as proposed by Diffloth (1996), Kingsada (2003), and Sidwell (2013)), these have been included in the bibliography.

3. Concluding Remarks

Although it is hoped that this project is comprehensive, it is assumed that there will inevitably be some valuable additions and corrections needed. In light of this, this Palaungic bibliographic compilation should be viewed as a living document that will continually be amended and updated. Therefore, all comments, corrections and additional contributions of bibliographic information are welcome and can be sent to: dgordon@simpsonu.edu. Also, to inquire about an updated version of this bibliography, please send an email to the preceding address.

In addition, it is acknowledged that another desirable improvement upon the effectiveness of this project would be to include annotations for many of the entries, as well as categorizing the resources according to the primary topic covered in each bibliographic item. Therefore, an annotated version of these sources is forthcoming and is currently planned for publication in 2015. Along with annotations, these sources will also be organized according to the principal linguistic domain found within each bibliographic entry.

With all the efforts in putting this bibliography together, it is a reminder that there is much left to be learned about Palaungic languages. Hopefully, above all, this project will support and encourage the furtherance of interest and community-centered research among Palaungic peoples.

4. Selective Palaungic Bibliography

Adams, Karen Lee. 1989. *Systems of Numeral Classification in the Mon-Khmer, Nicobarese and Aslian Subfamilies of Austroasiatic*. Canberra: Australian National University.

Adams, Karen Lee. 1991. The influence of Non-Austroasiatic languages on numeral classification Austroasiatic. *Journal of-the American Oriental Society* 111(1):62-81.

Adams, Karen Lee. 1992. A comparison of the numeral classification of humans in Mon-Khmer. *Mon-Khmer Studies* 21:107-129.

Anderson, G.D.S. and N.H. Zide. 2002. Issues in Proto-Munda and Proto-Austroasiatic nominal derivation: The bimoraic constraint. In M. Macken (ed.), *Papers from the Tenth Annual Meeting of the Southeast Asian Linguistics Society*, p. 55-74. Arizona State University, Program for Southeast Asian Studies.

Anderson, John. 1871. *A report on the expedition to Western Yunnan via Bhamô*. Calcutta: Superintendent of Government Printing.

Anderson, John. 1876. *Mandalay to Momien: a narrative of the two expeditions to Western China of 1868 and 1875 under Colonel Edward B. Sladden and Colonel Horace Browne*. London: Macmillan.

Antisdel, C.B. 1911. Elementary studies in Lahoo, Akha (Kaw) and Wa languages. *Journal of the American Oriental Society 1*(1):62-81.

Alves, Mark J. 2001. Distributional properties of causative verbs in some Mon-Khmer languages. *Mon-Khmer Studies* 31:107-120.

Atsushi, Yamada. 2001. Wago ni okeru 'saiki daiyōshi/joshi' tiʔ no kinō to yōhō" (The function and use of the 'reflexive pronoun/particle' tiʔ in Wa). *Proceedings of the 123rd Meeting of the Nihon gengogaku gakkai*.

Atsushi, Yamada. 2003. Wayu Baraoke fangyan zhong de fanshen daici tiʔ yu dongci cizu lianjie jiegou (The reflexive pronoun tiʔ and serial verb constructions in Parauk Wa). In Hiroaki Kitano (ed.) *Endangered Languages of the Pacific Rim, Descriptive and Theoretical Studies in*

Minority Languages of East and Southeast Asia, p. 155-167. ELPR Series A3-016, March, 2003.

Bastian, Adolf. 1868. *Reise durch Kambodja nach Cochinchina (Die Voelker des Oestlichen Asien: Studien und Reisen von Dr. Adolf Bastian, Vierter Band)*. Jena: Hermann Costenoble.

Benedict, Paul K. 1975. *Austro-Thai: Language and Culture*. New Haven: HRAF Press.

Benedict, Paul K. 1976. Austro-Thai and Austroasiatic. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic Studies*, Vol. 1, p. 1-36. Honolulu: University of Hawaii.

Bernatzik, H. A. and E. Bernatzik. 1958. *The Spirits of the Yellow Leaves*. London: Robert Hale.

Bigandet, Paul Ambrose. 1858. A comparative vocabulary of Shan, Kakying and Pa-laoung. *Journal of the Indian Archipelago*, 1:221-232.

Blagden, Charles Otto. 1894. Early Indo-Chinese influence in the Malay Peninsula, as illustrated by some of the dialects of the aboriginal tribes. JSBRAS. 27:21-56.

Block, Karen L. 1994. *Discourse grammar of first person narrative in Plang*. MA thesis. University of Texas at Arlington.

Block, Karen L. 1996a. Plang: some possessive noun phrases. Technical Paper #20, Payap University.

Block, Karen L. 1996b. What makes a story in Plang? *Mon-Khmer Studies* 26:357-385.

Buakaw, Supakit. 2009. The phonetic realization of denasalized final nasal consonants in No Lae Palaung. *Journal of Language and Culture*, 28(1):31-46.

Buakaw, Supakit. 2012. A phonological study of Palaung dialects spoken in Thailand and Myanmar, with focuses on vowels and final nasals. Ph.D dissertation. Mahidol University

Cabaton, Antoine. 1905. Dix dialectes indochinois recueillis par Prosper Odend'hal. Etude linguistique par Antoine Cabaton. *Journal Asiatique*. 10e sér 5:265-344.

Cameron, A. A. 1911. A note of the Palaungs of the Kodaung Hill Tracts of the Momeik State. *Census of India*, vol. IX, Part, Appendix A: v-xlii.

Chareonwong, Kitsana. 1987. Karn samruat chumchon Lua 5 amphor in Chiang Mai (The survey of the Lua in 5 districts in Chiang Mai province). Chiang Mai: Payap University.

Chareonwong, Kitsana. 1988. The relationship between Lua and Princes of Chiang Mai during Rattanakosin period. Chiang Mai: Payap University.

Charoenma, Narumol. 1980. *The sound systems of Lampang Lamet and Wiang Papao Lua*. MA thesis, Mahidol University, Bangkok.

Charoenma, Narumol. 1982. The phonologies of a Lampang Lamet and Wiang Papao Lua. *Mon-Khmer Studies* 11:35-45.

Chaturabhawd, Preecha. 1980. *People of the hills*. Bangkok: Edition Duang Kamol [chapter on Lawa, pp. 128-153].

Chazée, Laurent. 1999. *The peoples of Laos: rural and ethnic diversities*. Bangkok: White Lotus Press.

Chen, Guoqing. 1999. A preliminary study of evolution of minor syllables of Cambodian and Wa language. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 4: 32-37.

Chen, Guoqing. 2000. A comparison between Cambodian and Wa in morphological processes. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 6: 31-37.

Chinese Academy of Sciences. 1981. *Va-English dictionary*. Nationality Research Institute, Chinese Academy of Social Sciences, Beijing.

Coggin-Brown, J. 1911-12. Some customs of the Tawnpeng Palaungs. *Proceedings of the University of Durham Philosophical Society*, 4.1:1-8.

Condominas, Georges. 1974. Notes sur l'histoire lawa. A propos d'un lieu dit *Lua (Lawa)* en pays Karen (Ampoe Chom Thong, Changwat Chieng Mai). In *Commemorative volume for the museum*, p. 146-164. Bangkok.

Condominas, Georges. 1990. *From Lawa to Mon, from Saa' to Thai: historical and anthropological aspects of Southeast Asian social spaces*. Canberra: Australian National University.

Conver, Lynn C. 1999. A sketch of the phonology of a Lamet dialect. *Mon-Khmer Studies* 29:35-56.

Daladier, Anne. 2011a. Les Khasi. In Emilio Bonvini, Joëlle Busuttil and Alain Peyraube (eds.) *Dictionnaire des langues*, p. 1121-1128. Paris: PUF.

Daladier, Anne. 2011b. The group Pnaric-War-Lyngngam and Khasi as a branch of Pnaric. *Journal of Southeast Asian Linguistics Society*, 4.2:169-206.

Đặng Nghiêm Vạn, Chu Thái Sơn, and Lưu Hùng. 1993. *Ethnic Minorities in Vietnam*. Hanoi: The GIOI Publishers.

Davies, Henry Rudolph. 1894. *Diary of the expedition with the Chinese deputies to find the Huchi, Tienma, and Hanlung gates*. Rangoon: Superintendent, Government Printing.

Davies, Henry Rudolph. 1909. *Yün-nan: the link between India and the Yangtze*. Cambridge University Press.

Deepadung, Sujaritlak. 2009. Ethnicity and the Dara-ang (Palaung) in Thailand. *Journal of Language and Culture* 28(1):7-29.

Deepadung, Sujaritlak. 2011. Names and sub-groups of the Palaung: A preliminary field survey in Yunnan and Kengtung. *MANUSYA* 14(2):88-104.

Deepadung, Sujaritlak. 2013. Personal pronouns in Palaung dialects. *SIL Language and Culture Documentation and Description 22*, SIL Digital Resource.

Deepadung, Sujaritlak and Patpong, P. 2010. Dara-ang: Language, culture and ethnic identity maintenance at the Thai-Myanmar border. Research Institute for Languages and Cultures of Asia, Mahidol University. [in Thai]

Diffloth Gérard. 1974. Austro-Asiatic languages. *Encyclopedia Britannica*, 15th edition, vol. 2, 480-484.

Diffloth, Gérard. 1976a. An appraisal of Benedict's views on Austroasiatic and Austro-Thai relations. *Southeast Asian Studies*, Discussion Paper #82.

Diffloth, Gérard. 1976b. Proto-Mon-Khmer final spirants. In Hashimoto Mantaro (ed.) *Genetic relationship, diffusion and typological similarities of East and Southeast Asian languages*, p.210-229. Tokyo: The Japan Society for the Promotion of Science.

Diffloth, Gérard. 1977a. Mon-Khmer initial palatals and "substratumized" Austro-Thai. *Mon-Khmer Studies* 6:39-57.

Diffloth, Gérard. 1977b. *Proto-Waic and the effects of register on vowel gliding*. Presented at the 10th International Conference on Sino-Tibetan Languages and Linguistics, Washington D.C.

Diffloth, Gérard. 1980. The Wa languages, *Linguistics of the Tibeto-Burman Area 5*(2):1-182.

Diffloth, Gérard. 1982a. *On the Palaungic Branch of Mon-Khmer*. Presented at the 15th Annual International Conference on Sino-Tibetan Languages and Linguistics, Beijing.

Diffloth, Gérard. 1982b. *Subclassification of Palaungic and notes on "P'uman"*. Presented at the 15th Annual International Conference on Sino-Tibetan Languages and Linguistics, Beijing.

Diffloth, Gérard. 1985. Non-universals of registers. Paper presented at the Conference on Language Typology and Universals, University of Wisconsin, Milwaukee, March 29-30, 1985.

Diffloth, Gérard. 1991a. Linguistic prehistory from a Mon-Khmer perspective. Paper presented at the conference, *The High Bronze Age of Southeast Asia and South China*, Hua Hin, Thailand, 14-19 January.

Diffloth, Gérard. 1991b. Palaungic vowels in Mon-Khmer perspective. In Jeremy H. C. S. Davidson (ed.) *Austroasiatic Languages: Essays in Honor of H. L. Shorto*, p. 13-27. SOAS: University of London.

Diffloth, Gérard. 1992. On the Bulang (Blang, Phang) languages. *Mon-Khmer Studies* 18-19:35-43.

Diffloth, Gérard. 1996. Mang, a small north Mon-Khmer language of great historical significance, In Suriya Ratanakul and Suwilai Premsrirat (eds.) *Pan-Asiatic Linguistics: Abstracts of the 4th*

International Symposium on Languages and Linguistics, p. 41. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University.

Diffloth, Gérard. 2005. The contribution of linguistic palaeontology to the homeland of Austro-asiatic. In Sagart, Laurent, Blench Roger and Alicia Sanchez-Mazas (eds.) *The Peopling of East Asia: Putting Together Archaeology, Linguistics and Genetics*, p. 79-82. New York, NY: RoutledgeCurzon.

Diffloth, Gérard. 2011. Les langues Austro-Asiatiques. In Emilio Bonvini, Joëlle Busuttil and Alain Peyraube (eds.) *Dictionnaire des langues*, p. 1099-1102. Paris: PUF.

Diffloth, Gérard and Norman H. Zide. 1976. *Austroasiatic Number Systems*. Special issue of *Linguistics*, vol. 174. The Hague: Netherlands.

Diffloth, Gérard and Norman H. Zide. 2003. Austroasiatic languages. In William Frawley and William Bright (eds.) *International Encyclopedia of Linguistics*, 2nd ed., Vol. 1, p. 177-181.

Dissawarotham, Phijitra. 1986. The phonology of Plang as spoken in Baan Huay Nam Khun. MA thesis, Mahidol University, Bangkok.

Drage, Godfrey. 1907. *A few notes on Wa*. Rangoon: Superintendent, Government Printing.

Dryer, Matthew S. 1996. Mon-Khmer word order from a cross-linguistic perspective. In Karen Adams and Thomas J. Hudak (eds.) *Papers from the sixth annual meeting of the Southeast Asian Linguistic Society (SEALS VI 1996)*, p. 83-99. Tempe: Arizona State University.

Egerod, Søren. 1965. An English-Atayal vocabulary. *Acta Orientalia*, vol. 29.

Ferlus, Michel. 1974a. Délimitation des groups linguistiques austroasiatiques dans le Centre indochinos, *Asie du Sud-Est et Monde Insulindien* 5.1: 15-23.

Ferlus, Michel. 1974b. Les langues du groupe austroasiatique-Nord. *Asie du Sud-Est et Monde Insulindien* 5.1: 39-68.

Ferlus, Michel. 1978a. Comparaison de deux dialects lamet et restitution de la Proto-Langue. tapuscrit.

Ferlus, Michel. 1978b. Reconstruction de /TS/ et /TŠ/ en Mon-Khmer. *Mon-Khmer Studies* 7:1-38.

Ferlus, Michel. 1980. Formation des registres et mutations consonantiques dans les langues Mon-Khmer. *Mon-Khmer Studies* 8:1-76.

Ferlus, Michel. 1988. Les langues austrasiatiques. In Nicole Revel (ed.) *Le riz en Asie du Sud-Est*, 1:81-96; 2:327-332; 3:charts 26-34. Paris: Ecole des Hautes Estudes en Sciences Sociales.

Ferlus, Michel. 1996a. Du taro au riz Asie du Sud-Est, petite histoire d'un glissement sémantique. *Mon-Khmer Studies* 25:39-49.

Ferlus, Michel. 1996b. Langues et peuples viet-muong. *Mon-Khmer Studies* 26:7-28.

Ferlus, Michel. 2009. Le mot "sang" en austroasiaque. *Mon-Khmer Studies* 38:25-39.

Ferlus, Michel. 2010. The Austroasiatic vocabulary for rice: its origin and expansion. *Journal of the Southeast Asian Linguistics Society*, 3.2:61-76.

Ferrell, Raleigh. 1971. Le P'u-man, langue austroasiatique. *Bulletin de la Societe de Linguistique de Paris*. 6.1:405-412.

Flatz, Gebhard. 1970. The Khalo' or Mae Rim Lawa: A remnant of the Lawa population of Northern Thailand. *Journal of the Siam Society* 58.2:87-104.

Funke, Friedrich Wilhelm. 1960. Die Stellung der Lawā in der Kulturgeschichte Hinterindiens. *Tribus* 9:138-146. [The position of the Lawā in the cultural history of Further India].

Garnier, Francis. 1873. *Voyage d'exploration en Indo-Chine effectué pendant les années 1866, 1867, et 1968 par une commission française, présidée par Doudart de Lagrée, et pubié. . . sous la direction de. . . Francis Garnier*, 2 vols. Paris: Hachette.

Giaphong, Suchada. 2004. Plang grammar as spoken in Huay Namkhun village, Chiang Rai province. MA thesis. Mahidol University.

Greenberg, Joseph H. 1972. Numeral classifiers and substantival number: problem in the genesis of a linguistic type. *Working Papers on Language Universals*, Stanford University, Department of Linguistics.

Grierson, Sir George Abraham. 1903-28. *Linguistic Survey of India*. 13 vols. Calcutta: Office of the Superintendent of Government Printing.

Grimes, Barbara. 2003. Northern Mon-Khmer languages. In *International Encylopeda of Linguistics* (Vol. 3). Oxford: Oxford University Press.

Gussenhoven, Carlos. 2004. *The Phonology of Tone and Intonation*. Cambridge University Press.

Hall, Elizabeth. 2010. A Phonology of Muak Sa-aak. MA thesis, Payap University Thailand.

Hamp, Eric P. 1976. On Mon-Khmer, it's kin, and principles. In *Austroasiatic Studies*, ed. Philip N. Jenner, Laurence C. Thompson and Stanley Starosta, vol. 2:423-429. Honolulu: University of Hawaii Press.

Haudricourt, A. G. 1966. The limits and connections of Austroasiatic in the Northeast. In Norman H.Zide (ed) *Studies in comparative Austroasiatic linguistics*. The Hague: Mouton.

Harding, H. I. 1927. K'ala language. Addendum to Davies' vocabularies of Mon-Khmer languages of Yunnan and Western Tibet. *Journal of the Burma Research Society* 17:165-168.

Harvey, G. E. 1933. *Wa Précis*. Rangoon.

Harvey, G.E. 1957. The Wa people of the Burma-Chinese border. *St. Anthony's Papers* [Oxford] 2:126-135.

Hayes, La Vaughn H. 1996. Another look at final spirants in Mon-Khmer. *Mon-Khmer Studies* 26:51-64.

Hopple, Paulette. 1988. In their own language: Finding communication centers among selected dialects of Wa. In *Ekasan sammanaa thaang wichaakaan ruang Lawa nay laannaa, 7-8 Minaakhom 2531 (Seminar papers of different academic disciplines about Lawa in Laana, 7-8 March, 1988)*, p. 33-41, Chiengmai, Thailand: Chiang Mai Teachers College (Faculty of Humanities and Social Sciences, Chieng Mai Teacher's College, Lanna United Colleges).

Howard, Michael C. 2005 *Textiles of the highland peoples of Burma, vol. II*. Bangkok: White Lotus.

Howard, Michael C., and Wattana Wattanapun. 2001. The Palaung in Northern Thailand. Chiang Mai: Silkworm Books.

Hpuk lai sigang si siyeh pa hkrau (Wa New Testament). 1950. Rangoon, Burma: British and Foreign Bible Society.

Hutchinson, E. H. 1935. The Lawa in Northern Siam. *Journal of the Siam Society* 27(2):153-182.

Huffman, Franklin E. 1971. Phonology and vocabulary of Mon-Khmer languages, Unpublished ms (field notes). Ithaca, NY, Cornell University.

Huffman, Franklin E. 1976a. The register problem in fifteen Mon-Khmer languages. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic Studies*, vol. 2, p. 575-590. Honolulu: University of Hawaii Press.

Huffman, Franklin E. 1976b. The relevance of lexicostatistics to Mon-Khmer languages. . In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic Studies*, vol. 2, p. 539-574. Honolulu: University of Hawaii Press.

Huffman, Franklin E. 1977. An examination of lexical correspondences between Vietnamese and some other Austroasiatic languages. *Lingua* 43.2:171-98.

Huffman, Franklin E. 1978. The centrality of Katuic-Bahnaric to Austroasiatic. Presented at the 2nd International Conference on Austroasiatic Linguistics, Mysore, India.

Izikowitz, Karl Gustav. 1941. Fastening the soul: some religious traits among the Lamet. *Goteborg Högskolar Årskrift* 47.14:3-32.

Izikowitz, Karl Gustav. 1943. The community house of the Lamet. *Ethnos* 8:19-60.

Izikowitz, Karl Gustav. 1951. Lamet, jungle peasants of Indochina. *Eastern Anthropologist* 4:124-132.

Izikowitz, Karl Gustav. 1969. Neighbors in Laos. In F. Barth. Bergen (ed) *Ethnic Groups and Boundaries*. Scandinavian University Books for Univeritetsforlaget.

Izikowitz, Karl Gustav. 2001. *Lamet: Hill Peasants in French Indochina*. Bangkok: White Lotus (previously published by Goteborg: Etnologiska Studier 17 in 1951).

Janzen, Hermann. 1976a. Structure and function of clauses and phrases in Pale. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic Studies*, vol. 2, p. 669-691. Honolulu: University of Hawaii.

Janzen, Hermann. 1976b. The system of verb-aspect words in Pale. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic Studies*, vol. 2, p. 659-667. Honolulu: University of Hawaii Press.

Janzen, Hermann. 1978. A phonological description of Pale in comparison with Gold- and Rumai-Palaung. Presented at the 2nd International Conference on Austroasiatic Linguistics, Mysore, India.

Janzen, Hermann, and M. Janzen. 1972. Grammar analysis of Pale clauses and phrases. *Journal of the Burma Research Society* 55:47-99.

Jindasri, Nusit. 1993. Karn pathibat ngarn kab Lua. *Journal of Hill Tribe Research Center*, 17 (1-2 Jan-June 1993).

Kasisopa, B. 2003. Phonological study and genetic classification of Dara-ang (Palaung) spoken at Nawlae village, Fang district, Chiang Mai province. M.A. thesis, Chulalongkorn University. [In Thai]

Kauffman, Hans E. 1971. Stone memorials of the Lawā (Northwest Thailand). *Journal of the Siam Society* 59.1:129-151.

Kauffman, Hans E. 1972a. Die Samang der Lawā und ihre Menhire (Nordwest-Thailand). *Ethnologische* 1:223-231 [The Samang of the Lawā and their menhirs].

Kauffman, Hans E. 1972b. Some social and religious institutions of the Lawā (Northwest Thailand). *Journal of the Siam Society* 60.1:237-306.

Kauffman, Hans E. 1977. Some social and religious institutions of the Lawā (Northwest Thailand): part II. *Journal of the Siam Society* 65.1:181-226.

Kauffman, Hans E. 1980. Some social and religious institutions of the Lawā (Northwest Thailand): part III. *Journal of the Siam Society* 68.1:87-125.

Kingsada, Thongpheth. 2003. Languages and ethnic classification in the Lao P.D.R. In Hayashi Yukio and Thongsa Sayavongkhamdy (eds.) *Cultural Diversity and Conservation in the Making of Mainland Southeast Asia and Southwestern China: Regional Dynamics in the Past and Present*, p. 21-38. Bangkok: Amarin Printing and Publishing Company Ltd.

Kingsada, Thongpheth and Ryuichi Kosaka. 1999. Khabit [word list]. In Thongpheth Kingsada and Tadahiko Shintani (eds.) *Basic vocabularies of the languages spoken in Phongxaly, Lao P. D. R.*., p. 297-317. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Komonkitiskun, Jiranan. 1985. *Some general characteristics of Lawa grammar (La-up dialect)*. MA thesis, Mahidol University, Bangkok.

Komonkitiskun, Jiranan. 1992. Lawa Pronouns. *Mon-Khmer Studies* 21:199-205.

Kunstadter, Peter. 1965a. *The future of upland tribal people in the nations of Southeast Asia: Lua' and Karen hill peoples of Northwest Thailand*. Paper presented at the Annual Meeting of the American Anthropological Association, Denver, CO (November).

Kunstadter, Peter. 1965b. *The Luaq (Lawa) of Northern Thailand: aspects of social structure, agriculture and religion*. Princeton: Center of International Studies, Research Monographs 21.

Kunstadter, Peter. 1966a. Living with Thailand's gentle Lua. *National Geographic Magazine* 130.1:122-152 (July).

Kunstadter, Peter. 1966b. Residential and social organization of the Lawa of Northern Thailand. *Southwestern Journal of Anthropology* 22:61-84.

Kunstadter, Peter. 1966c. The Luaq and Skaw of Maehongson Province, Northwestern Thailand. In Peter Kunstadter (ed.) *Southeast Asian tribes, minorities and nations,* vol. 2, Chapter 17. Princeton: Princeton University Press.

Kunstadter, Peter. 1967. The Lua' and Sgaw Karen of Maehongson province, Northwestern Thailand. In Peter Kunstadter (ed.) *Southeast Asian tribes, minorities and nations*, vol. 2, p. 639-674. New Jersey: Princeton University Press.

Kunstadter, Peter. 1969a. Hill and valley populations in Northwestern Thailand. In Peter Hinton (ed.) *Tribesmen and Peasants in Northern Thailand*. Tribal Research Centre, Chiang Mai, Thailand.

Kunstadter, Peter. 1969b. *Socio-cultural change among upland peoples of Thailand: Lua' and Karen – two modes of adaption*. Proceedings of the 8th International Congress of Anthropological and Ethnological Sciences, 1968, Tokyo and Kyoto, vol. II, Ethnology, Tokyo 232-235.

Kunstadter, Peter. 1974. Usage et tenure des terres chez les Lua' (Thailande). *Études Ruales* 53-6:449-466.

Kunstadter, Peter. 1978. Subsistence agriculture economics of Lua' and Karen hill farmers, Mae Sariang District, Northwestern Thailand. In Peter Kunstadter (ed.) *Farmers in the Forest*. Honolulu: University Press of Hawaii.

Kunstadter, Peter. 1983a. Animism, Buddhism, and Christianity: religion in the life of the Lua' people of Pa Pae, North-West Thailand. In John McKinnon and Wanat Bhruksasri (eds.) *Highlanders of Thailand*, p. 135-154. Kuala Lumpur: Oxford University Press.

Kunstadter, Peter. 1983b. Highland populations in Northern Thailand. In John McKinnon and Wanat Bhruksasri (eds.) *Highlanders of Thailand*, p. 15-45. Kuala Lumpur: Oxford University Press.

Kunstadter, Peter. 1984. Cultural ideals, socioeconomic change, and household composition: Karen, Lua', and Hmong, and the Thai in the Northwestern Thailand. In R. McNetting, R. R. Wilke and E. J. Arnold (eds.) *Households: Comparative historical studies of the domestic group*. Berkeley.

Lai Maku (Gospel of Mark-Wa). 1951. Rangoon, Burma: British and Foreign Bible Society.

Lai Yohan (Gospel of John-Wa). 1951. Rangoon, Burma: British and Foreign Bible Society.

LeBar, Frank M., G. C. Hickey, and J. K. Musgrave. eds. 1964. *Ethnic groups of Mainland Southeast Asia*. New Haven: Human Relations Area Files Press.

Lefèvre-Pontalis, Pierre. 1892. Nores sur quelques populations du nord de l'Indochine. *Journal Asiatique* 8.19:237-269.

Lefèvre-Pontalis, Pierre. 1896. Nores sur quelques populations du nord de l'Indochine (2e série). *Journal Asiatique* 9.8:129-303.

Le May, Reginald. 1926. *An Asian arcady: the land and people of Northern Siam*. Cambridge: William Heffer.

Lewis, M. Paul, Gary F. Simons, and Charles D Fennig. (eds.). 2013. Ethnologue: Languages of the World, Seventeenth edition. Dallas, Texas: SIL International. Online version: http://www.ethnologue.com.

Li Dao Yong. 1993. Mang cun kao cha [Research on Mang villages]. Zhongyang minzu xueyuan xuebao 1993/1:61-66.

Li Dao Yong, Nye Hsi Jen, Chyou Eh Feng. 1986. *Sketch of the Bulang Language* [in Chinese]. Beijing: Minorities Publishing House.

Li Yangsong. 1989. Oracle chicken bones of Wa people in Ximeng, Yunnan. *Tai-Yunnan Project Newsletter*, No. 4, March 1989. Canberra: Australian National University.

Lindell, Kristina. 1974. A vocabulary of the Yuan dialect of the Kammu language. *Acta Orientalia*. 36:191-207.

Lindell, Kristina, Jan-Olof Svantesson, and Damrong Tayanin. 1978. Two dialects of the Rəmeet (Lamet) language. *Cahiers de Linguistique Asie Orientale* 4:5-22.

Logan, James Richardson. 1858. Notes on the Pa Laong. *Journal of the Indian Archipelago* 2:233-236.

Lowis, C. C. 1906. A note on the Palaungs of Hsipaw and Tawngpeng. *Ethnographic Survey of India, Burma*. no.1.

Luce, Gordon H. 1950 *Rə-oe rak tsən Markɔs*. British and foreign Bible Society, Rangoon.

Luce, Gordon H. 1965. Danaw, a dying Austroasiatic language. *Lingua* 14:98-129.

Luce, Gordon H. 1972. Riang-Lang. Parable of the Prodigal Son. In Jacques Barrau, et al. (eds.) *Langues et techniques, nature et société, Vol II: Approche ethnologique et naturaliste (Volumes presented to André G. Haudricourt on his 60th birthday)*, p.169-172. Paris: Klincksieck.

Luce, Gordon H. 1973. *Old Burma – Early Pagan, Vol. IV: Language supplement (including Glossary of Old Mon (of Burma), Glossary of pre-standard Old Burmese, Glossary of Sanskrit/Pali loanwords as found in Burma's inscriptions of the 11th and 12th centuries A.D.; also 1 chart comparing Austro-Asiatic languages, and 28 charts comparing Old Burmese with Old Chinese, Classical Tibetan, and 8 languages of the Burma/Lolo groups)*, assisted by Dr. and Mrs. Chang (of Cambridge) U Tin Htway (of Heidelberg) , and Sandra Harding, Unpublished ms [copy deposited in Wason Collection, Olin Graduate Library, Cornell University, Ithaca, NY].

Ma Yin, ed. 1994. *China's Minority Nationalities*, 2nd edition. Beijing: Foreign Languages Press.

Maddieson, Ian and Peter Ladefoged. 1985. 'Tense' and 'lax' in four minority languages of China. *Journal Phonetics* 13:433-454.

Maeda, Seiji. 1981. Position de thèse: Les Lawa, la plus ancienne population de la Thaïlande – étude monographique et socio-juridique. *Asie du Sud-Est et Monde Insulindien* 11.1-2:167.

Maeda, Seiji. 1988. Anthropologie juridique des Lawa (Thaïlande): étude intensive et étude extensive. Ph.D thesis. Universite de Paris.

Mak, Pandora. 2012. *Golden Palaung: A Grammatical Description*. Asia-Pacific Linguistics Open Access monographs, A-PL 003; SEAsian Mainland Languages E-Series (SEAMLES), 002. Canberra: Asia-Pacific Linguistics (SEAMLES), available at http://hdl.handle.net/1885/9558

Matras, J. and Michel Ferlus. 1971. Les langues austroasiatiques. *Asie du Sud-Est et Monde Insulindien* 2(4):55-93.

McKinnon, John and Wanat Bhruksasri eds. 1983. *Highlanders of Thailand*. Singapore: Oxford University Press.

Milne, Mary Lewis (Leslie). 1921. *An elementary Palaung grammar*. Oxford: Clarendon Press.

Milne, Mary Lewis (Leslie). 1924. *The home of an Eastern clan*. Oxford: Clarendon Press.

Milne, Mary Lewis (Leslie). 1931. *A Dictionary of English-Palaung and Palaung-English*. Rangoon, Superintendent, Government Printing and Stationary.

Mitani, Yasuyuki. 1965. A descriptive and comparative study of the Khamet phonology. *Southeast Asian Studies* 3.3:22-51.

Mitani, Yasuyuki. 1966. Descriptive study of the Lawa language (Bo Luang district). *Southeast Asian Studies* 4.2:230-252.

Mitani, Yasuyuki. 1972a. A short history of Lawa. *Southeast Asian Studies* 10.1:131-168.

Mitani, Yasuyuki. 1972b. Studies in the Lawa phonology. *South East Asian Studies* 10.2: 174-196.

Mitani, Yasuyuki. 1976. Descriptive and comparative studies of the Khamet phonology. In *Southeast Asia: nature, society, and development (Kyoto U. Monograph of the Center for Southeast Asian studies, English ser. 9)*, p. 76-95. Honolulu: University of Hawaii Press.

Mitani, Yasuyuki. 1977. Palaung dialects: a preliminary comparison. *Southeast Asian Studies* 15.2:192-212.

Mitani, Yasuyuki. 1978a. On the reconstruction of Proto-Lawa vowels. Proceedings of the International Congress if Anthropological and Ethnological Sciences 10.2:84-85 (Abstract).

Mitani, Yasuyuki. 1978b. Phonological studies of-Lawa: description and comparison. Ph. D. dissertation, Cornell University.

Mitani, Yasuyuki. 1978c. Problems in the classification of Palaungic. Presented at the 2nd International Conference on Austroasiatic Linguistics, Mysore, India.

Mitani, Yasuyuki. 1979. Vowel correspondences between Riang and Palaung. In T.L. Thongkum, V. Panupong, P. Kullavanijava and M.R. Kalaya Tingsabadh (eds.) *Studies in Thai and Mon-Khmer Phonetics and Phonology in Honour of Eugénie J.A. Henderson*, p. 142-150. Bangkok: Chulalongkorn University Press.

Mitra, A.K. 1956. The Riang of Tripura. *Bulletin of the Department of Anthropology* 5.2:21-120.

Mukherjee, B. 1959. Trends and changes in the Riang culture since independence. *Bulletin of the Department of Anthropology* 8.2: 103-108.

Mukherjee, B. 1960. Cross-cousin marriage among the Riang. *Anthropos* 55:734-738.

Needham, Rodney. 1960. Alliance and classification among the Lamet. *Sociologus* 10:97-108.

Ngô, Đức Thịnh. 1978. Dân tộc Mảng [The Mảng ethnicity]. In *Các dân tộc ít người ở Việt Nam (các tỉnh Phía Bắc) [The ethnic minorities of Vietnam (Northern Region)*, 139-144. Hànội: Nhà Xuất bản Khoa học Xã hội.

Ngô, Đức Thịnh and Trương Văn Sihn. 1973. Vài nét vềự phân bố địa lý các nhóm ngữ ở Lào hiện nay [About the distribution of linguistic groups in contemporary Laos]. *Ngôn-ngữ* 1973/1:16-23.

Nguyễn, Loan. 1976. Vài nét về tiếng Mảng Tây Bắc Việt Nam [About the Mảng language of Northwestern Vietnam]. *Ngôn-ngữ* 1976/1:53-62.

Nguyễn, Phú Phông. 2011. Les Vietnamien. In Emilio Bonvini, Joëlle Busuttil and Alain Peyraube (eds.) *Dictionnaire des langues*, p. 1108-1114. Paris: PUF.

Nguyễn, Trúc Bình. 1972. Nhóm Kháng [The Kháng]. In *Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc Việt Nam [The ethnic minorities of Austroasiatic origin in Northwest Vietnam]*, 251-207. Hànội: Nhà Xuất bản Khoa học Xã hội.

Nguyễn, Văn Huy. 1975. Về nhóm Kháng ở bản Quảng Lâm [On the Kháng in the village of Quảng Lâm]. In *Về vấn đề xác định thành phần các dân tộc thiểu số ở miền bắc Việt Nam [On the identification of ethnic minorities in North Vietnam]*, 429-443. Hànội: Nhà Xuất bản Khoa học Xã hội.

Nimmanaheminda, Kraisi. 1963. The Mrabri language. *Journal of the Siam Society* 51.2:179-184. [with comparative vocabularies including Khamed (Lamet)].

Nimmanaheminda, Kraisi. 1967. The Lawa guardian spirits of Chiangmai. *Journal of the Siam Society* 55.2:185-225.

No author. 1917. *Linguistic survey of Burma*. Rangoon: Superintendent of Government Printing.

No author. 1984. *Wa language textbook vol. 1*. Kunming: Yunnan Ethnic Publishing House.

Obayashi, Taryo. 1963. Field work among the Lawa and Sgau-Karen of Northwestern Thailand. *Minzokugaku Kenkyu* 28:223-229.

Obayashi, Taryo. 1964a. Aus dem religiösen Lebender Lawa in Nordwest-Thailand. *Memoirs of the 7th International Congress of Anthropological and Ethnological Sciences*, Moscow, 9:416-419. [Notes on the religious life of the Lawa in northwestern Thailand].

Obayashi, Taryo. 1964b. The Lawa and Sgaw Karen of Northwestern Thailand. *Journal of the Siam Society* 52.2:199-216.

Obayashi, Taryo. 1966. Death customs of the Lawa in northwestern Thailand. *Memoirs of the Institute for Oriental Culture*, 41:247-268.

d'Oréans, Henri. 1898. *From Tonkin to India by the sources of the Irawadi*. Translated by Hamley Bent. London, Methuen.

Ostapirat, Weera. 2009. Some phonological criteria for Palaung sub-grouping. *Journal of Language and Culture* 28(1):63-76.

Parkin, Robert. 1991. *A guide to Austroasiatic speakers and their languages*. Oceanic Linguistics Special Publication, vol. 23. Honolulu: University of Hawaii Press.

Paulsen, Debbie Lynn. 1989. A phonological reconstruction of Proto-Plang. MA thesis, University of Texas at Arlington.

Paulsen, Debbie Lynn. 1991. *Tone and intonation in Plang*. Presented at the 24th International Conference on Sino-Tibetan Languages and Linguistics, Bangkok, Thailand.

Paulsen, Debbie Lynn. 1992a. A phonological reconstruction of Proto-Plang. *Mon-Khmer Studies* 18-19:160-222.

Paulsen, Debbie Lynn. 1992b. *A preliminary peek at Plang Particles*. Working Papers in Linguistics, Payap University.

Paulsen, Debbie Lynn. 1996a. *Phonology in Plang*. PYU Working Papers in Linguistics 1:126-142.

Paulsen, Debbie Lynn. 1996b. *Tone and intonation in Plang*. PYU Working Papers in Linguistics 1:53-65.

Paulsen, Debbie Lynn and Karen Block. 1997. A preliminary report on independent clause structures in Plang. Technical Paper #27, Payap University.

Peiros, Ilia. 1996. The Vietnamese etymological disctionary and 'new' language families. Bangkok: Institute for Language and Culture for Rural Development, Mahidol University.

Peiros, Ilia. 1997. A Lawa Dictionary. Salaya: Institute of Language & Culture for Rural Development. Mahidol University, Bangkok.

Peiros, Ilia. 1998. *Comparative Linguistics in Southeast Asia*. Pacific Linguistics Series C-142. Canberra, Australian National University.

Phan, Hu'u Dat. 1978. Les premiers acquis de l'ethnologie vietnamienne. *Asie du Sud-Est et Monde Insulindien* 9.1-2:155-177.

Phratham lailuang phrayesu khrit (Lawa New Testament). 1971. Bangkok: Bible Society of Thailand and Laos.

Phuk lai pa rom dee bleeh gum lox loe ceu pwi (Vo/Wa language—Dictionary). 1997. Chiang Mai: Payap University, Applied Linguistics Training Program.

Phuk lai sigang si plak lox Lai Ruthax (Book of Ruth-Wa). 1998. Chiang Mai: Payap University, Applied Linguistics Training Program.

Pinnow, Heinz Jürgen. 1959. *Versuch einer Lautlehre der Karia-Sprache*. Wiesbaden, Otto Harrassowitz, xviii.

Pinnow, Heinz Jürgen. 1965. Personal pronouns in Austroasiatic, *Lingua* 14.

Pitchford, V. C. 1937. The wild Wa states and Lake Nawngkio. *The Geographical Journal* 90:223-232.

Premsrirat, Suwilai. 2001. Austroasiatic languages and peoples. In *Annotated Bibliography for Ethnic Minorities in Thailand* N°2. Nakhon Pathom, ILCRD, Mahidol University [in Thai].

Proschan, Frank. 1996. A survey of Khmuic and Palaungic languages in Laos and Vietnam. In *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics, Jan 8-10, 1996*, vol. 3:895-919. Bangkok: Mahidol University.

Purnell, Herbert C. 1972. Toward a contrastive analysis between Thai and hill tribe languages: some phonetic data. In Jimmy G. Harris and Richard B. Noss (eds.) *Tai phonetics and phonology*, p. 113-130. Bangkok: Central Institute for English Language.

Qiu Efeng, Li Dao Yong, and Nie Xizhen. 1980. A brief description of the Va language. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 1:58-69.

Qixiang, Yan. 1981. The Va script. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 4:77-78.

Qixiang, Yan. 1983. A brief description of the Benglong language. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 5:67-80.

Qixiang, Yan, and Zhou Zhizhi. 1995. China's Mon-Khmer languages and the Austroasiatic language family [中国孟高棉语旌语言与南亚语系]. Beijing: Central Nationalities University Press. [In Chinese]

Qixiang, Yan, Zhou Zhizhi, LiDao Yong, Wang Jingliu, Qu Ming, Tao Wenjun, and Tian Kaizheng. 1981. *Pug lai cix ding Yiie si ndong lai vax mai lai hox / Wa-Han jianming cidian* [Concise Wa-Chinese dictionary]. Kunming: Yunnan minzu chubansh [Yunnan Ethnic Publishing House].

Rangsit, M. C. Sanidh. 1942/45. Beitrag zur Kenntnis der Lawasprachen von Nord-Siam. *Anthropos* 37/40:689-710. [Contribution to the knowledge of the Lawa language in northern Siam].

Rangsit, M. C. Sanidh. 1945. Die Lawa, ein hinterindisches Bergvolk. *Atlantis* 17:493-496. [The Lawa, hill people of Further India].

Ratanakul, Suriya. 1984a. A study of compound nouns in Lawa, White Meo, and Sgaw Karen. Presented at the 17th Annual International Conference on Sino-Tibetan Languages and Linguistics, University of Oregon, Eugene.

Ratanakul, Suriya. 1984b. Lawa lae lua? [Lawa and Lua?] (in Thai). *Journal of Language and Culture*, Mahidol University 4.1:56-78.

Ratanakul, Suriya. 1985a. Sociocultural aspects of the Lawa language. *Journal of the Siam Society* 73:183-204.

Ratanakul, Suriya. 1985b. The Lawa ləsɔm ʔlɛ poetry. *Journal of the Siam Society* 73.1-2:183-204.

Ratanakul, Suriya. 1988. *Languages in Southeast Asia, Part 1* (in Thai). Bangkok: Research Center of Southeast Asian Cultures, Mahidol University.

Ratanakul, Suriya. 1996. Lawa ləsɔm ʔlɛ poetry revisited. *Mon-Khmer Studies* 26:387-410.

Ratanakul, Suriya and Lakhana Daoratanahongse. 1985. The phonology of Lawa. In Suriya Ratanakul, David Thomas and Suwilai Premsrirat (eds.) *Southeast Asian linguistics studies presented to Andre G. Haudricourt*, p. 264-306. Bangkok: Mahidol University.

Ratanakul, Suriya and Lakhana Daoratanahongse. 1986. *Lawa-Thai Dictionary*. Institute of Language and Culture for Rural Development, Mahidol University, Parpphim Press.

Ratanakul, Suriya and Lakhana Daoratanahongse. 1987. *Thai-Lawa Dictionary*. Institute of Language and Culture for Rural Development, Mahidol University, Parpphim Press.

Ratanakul, Suriya, Sujaritlak Deepadung, and Phattiya Yimrewat. 1996. *Phasaa le? watthanatham lawa?: suksa cak bot sonthanaa 30 bot*. Bangkok: Sahadhammika Co, Ltd.

Rattanapitak, Ampika. 2009. Palaung wordlist. *Journal of Language and Culture*. 28(1)77-122.

Rattanavong, Houmphanh. 2003. History of population in Laos and Essay on Ethnic Classification. In Hayashi Yukio and Thongsa Sayavongkhamdy (eds.) *Cultural Diversity and Conservation in the Making of Mainland Southeast Asia and Southwestern China: Regional Dynamics in the Past and Present*, p. 1-20. Bangkok: Amarin Printing and Publishing Company Ltd.

Sacchiero, G. B. 1890. Alcuni cenni sulla Tribu dei Palaung. *Bollettino della Società Geografica Italiana* 1890:920-925.

Sanae, Sopha.1994. Hpuk lai pa rom dui blui gum lo lo-e ceu pwi. (Picture dictionary of three languages (Wa-Thai-English, 71 p.) Chiang Mai: SIL and Payap University Tribal Training Program.

Sarasidhiprakat, Luang. 1961. The tradition of Lawa. *Sinlapakon* 4.5.

Satyawadhna, Cholthira. 1991. The dispossessed: an anthropological reconstruction of Lawa ethnohistroy in the light of their relationship with the Tai. Ph.D. thesis. Australian National University.

Schlatter, Donald. 1965. *Lawa-English dictionary*.

Schlatter, Donald. 1976. Lavua (Lawa, Lua'). In William Smalley (ed) *Phonemes and orthography: language planning in ten minority languages of Thailand*, p. 273-282. Canberra: Pacific Linguistics.

Schliesinger, Joachim. 2000. *Ethnic Groups of Thailand: Non-Tai Speaking Peoples*. Bangkok: White Lotus Press.

Schliesinger, Joachim. 2003a. *Ethnic Groups of Laos: Volume 1. Introduction and Overview*. Bangkok: White Lotus Press.

Schliesinger, Joachim. 2003b. *Ethnic Groups of Laos: Volume 2. Profile of Austro-Asiatic-Speaking Peoples*. Bangkok: White Lotus Press.

Schmidt, P. Wilhelm. 1904. Grundzuge einer Lautlehre der Khasi-Sprache inihren Beziehungen zu derjenigen der Mon-Khmer-Sprachen. Mit einem Anhang: die Palaung-, Wa-, und Riang-Sprachen des mittleren Salwin. *Abhandlungen der bayerischen Akademie der Wissenschaft* 22.3:677-810.

Schmidt, P. Wilhelm. 1906. Die Mon-Khmer Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens. Braunschweig: *Archiv für Anthropologie*, 5, 59-109.

Schrock, Joann L., et al. 1970. *Minority groups in Thailand*. DA pamphlet 550-107, Washington, DC, US Government Printing Office, 1135p.

Scott, James George. 1896. The wild Wa: a head-hunting race. *Asiatic Quarterly Review* 1.1-2:138-152.

Scott, James George. 1901. Wa. *Census of India* vol. I, Ethnographic Appendiz: 214-221.

Scott, James George and J. P. Hardiman. 1900. *Gazetteer of Upper Burma and the Shan State*. Rangoon: Superintendent, Government Printing.

Sebeok, Thomas A. 1942. An examination of the Austroasiatic language family. *Language* 18:206-217.

Seidenfaden, Erik. 1923. The Lawa, additional note. *Journal of the Siam Society* 17 (2).

Seidenfaden, Erik. 1926. The Kha Tong Lu'ang. *Journal of the Siam Society* 20 (1):41-48.

Seidenfaden, Eric. 1940. The Lawa of Umphai and middle Me Ping. *Journal of the Siam Society* 32.1:29-36.

Shafer, Robert. 1952. Etudes sur l'austroasien. *Bulletin de la Societe de Linguistique de Paris* 48:111-158.

Shafer, Robert. 1965. *Studies in Austroasian II*. Helsinki.

Shaorong, Gou. 1990. Si ndah baraog joung gox (Medicines of the Wa nationality of China). Kunming: Yunnan Minorities Press.

Shila, Sarapi. 1993. Palong. *Tribal Research Institute Quarterly* 17.3-4:25-39.

Sila, S. 1993. Palaung. *TRI Quarterly*, 17(3-4):25-39. [in Thai]

Shintani, Tadahiko. 2008. *The Palaung Language, Comparative Lexicon of its Southern Dialects*. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.

Shorto, Harry L. 1960. Word and syllable patterns in Palaung. *Bulletin of the School of Oriental and African Studies* 23:544-567.

Shorto, Harry L. 1963. The structural patterns of northern Mon-Khmer languages. In Harry L. Shorto (ed) *Linguistic Comparison in South East Asia and the Pacific*, p. 45-61. London: School of Oriental and African Studies.

Shorto, Harry L. 1972. The word for "two" in Austroasiatic. In Jacqueline M. C. Thomas and Lucien Bernot (eds.) *Langues et Techniques, Nature et Societe*, p. 233-235. Paris: Klincksieck.

Shorto, Harry L. 1973. Three Mon-Khmer word families. *Bulletin of Oriental and African Studies*, 36.2:374-381.

Shorto, Harry L. 1976. The vocalism of Proto-Mon-Khmer. In Philip N. Jenner, Laurence C. Thompson and Stanley Starosta (eds.) *Austroasiatic studies*, vol. 2:1041-1067. Honolulu: University of Hawaii Press.

Shorto, Harry L. 2006. *A Mon-Khmer Comparative Dictionary*. Canberra, Pacific Linguistics 579.

Shorto, Harry. L. 2013. *Palaung Word List: Based on Materials Collected from Pan Shwe Kya, Namhsan, Sept-Oct, 1957*. Asia-Pacific Linguistics Open Access Monographs, A-PL 004; SEAsian Mainland Languages E-Series (SEAMLES), 003. Canberra, ACT: Asia-Pacific Linguistics, available at http://hdl.handle.net/1885/9782.

Shorto, Harry. L. 2013a. *Riang-Lang Vocabulary: Compiled from the Materials Collected by G. H. Luce*. Asia-Pacific Linguistics Open Access Monographs, A-PL 005; SEAsian Mainland Languages E-Series (SEAMLES), 004. Canberra, ACT: Asia-Pacific Linguistics, available at http://hdl.handle.net/1885/9781.

Shorto, Harry. L. 2013b. *Wa-Praok Vocabulary*. Asia-Pacific Linguistics Open Access Monographs, A-PL 006; SEAsian Mainland Languages E-Series (SEAMLES), 005. Canberra, ACT: Asia-Pacific Linguistics, available at http://hdl.handle.net/1885/9783.

Shorto, Harry L., Jacob, Judith M. and E. H. S. Simmonds. 1963. *Bibliography of Mon-Khmer and Tai linguistics*. London: Oxford University Press.

Sidwell, Paul. 2005. Proto-Katuic phonology and the sub-grouping of Mon-Khmer languages. In Paul Sidwell (ed) *Papers from the 15th annual meeting of the Southeast Asian Linguistics*

Society SEALS XV, p. 193-1204, at http://pacling.anu.edu.au/series/SEALS-PDFs/SEALS-15.pdf. Accessed 13 August 2013.

Sidwell, Paul. 2009a. Classifying the Austroasiatic languages: history and state of the art. Munich, Lincom Europa.

Sidwell, Paul. 2009b. Proto-Mon-Khmer vocalism: moving on from Shorto's 'alternances'. *Journal of the Southeast Asian Linguistics Society*, 1:205-214.

Sidwell, Paul. 2010. The Austroasiatic central riverine hypothesis. *Journal of Language Relationship*, 4:117–134.

Sidwell, Paul. 2011. Proto-Khasian and Khasi-Palaungic. *Journal of Southeast Asian Linguistics Society*, 4.2:144-168.

Sidwell, Paul. 2013. A Mon-Khmer Languages Project. http://www.sealang.net/monkhmer/ (accessed 13 August 2013)

Sidwell, Paul & Roger Blench. 2011. The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis. In N. J. Enfield (ed.). *The Dynamics of Human Diversity: The Case of Mainland Southeast Asia*, p. 317-345. Canberra: Pacific Linguistics.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ (The Rmeet Ethnic Group). ວັນນະສິນ (Vannasin) 5 (May): 8-9.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ (The Rmeet Ethnic Group, continued from previous issue). ວັນນະສິນ (Vannasin) 6 (June): 7-8.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ: ການເອົາຜົວເອົາເມຍສ້າງຄອບຄົວຂອງຄົນຣຶມເມດ (The Rmeet Ethnic Group: Marriage and founding a family, rites and customs). ວັນນະສິນ (Vannasin) 7 (July): 8-9.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ: ການເກີດລູກແລະລ້ຽງລູກ (The Rmeet Ethnic Group: Childbirth and childrearing customs and rituals). ວັນນະສິນ (Vannasin) 8 (August): 10-11.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ: ຄົນຕາຍແລະການສົ່ງສະການ (The Rmeet Ethnic Group: Death and funeral rites). ວັນນະສິນ (Vannasin) 9 (September): 8-9.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ: ຄົນຕາຍແລະການສົ່ງສະການ 2 (The Rmeet Ethnic Group: Death and funeral rites, continued from previous issue). ວັນນະສິນ (Vannasin) 10 (October): 13-14.

Simana, Suksavang. 2001. ຊາວເຜົ່າຣຶມເມດ:ຮີດເຮືອນ (The Rmeet Ethnic Group: House (spirit) rituals). ວັນນະສິນ (Vannasin) 11 (November): 14-16.

Simana, Suksavang. 2002. ຄວາມຄິດແລະຄວາມເຊື່ອ ຊາວເຜົ່າຣຶມເມດ (Rmeet worldview and beliefs). ວັນນະສິນ (Vannasin) 1 (January): 16-18.

Smalley, William A. 1963. Notes on Kraisri's and Bernatizik's word lists. Journal of the Siam Society 51.2:189-201.

Smalley, William A. 1964. The use of non-roman script for new languages. In William A. Smalley, et al. (eds.) *Orthography studies: articles on new writing systems (Helps for translators, prepared under the auspices of the United Bible Society, vol. 6)*. London: North-Holland.

SOAS Wa Dictionary Project. 2013. A Dictionary of the Wa Language with Burmese (Myanmar), Chinese, and English Glosses And Internet Database for Minority Languages of Burma (Myanmar). http://www.humancomp.org/wadict/. Accessed 13 August 2013.

Solomon Yaong Kaung. 1997. Been praix lao ceu pwi vax (Wa folktales). Chiang Mai, Thailand: Payap University, Applied Linguistics Training Program.

Sprenger, Guido. 2004. Encompassment and its discontents: Rmeet and lowland Lao relationships. In Gerd Baumann and André Gingrich (eds.) *Grammars of identity/alterity: a structural approach*, p. 173-191. New York/Oxford: Berghahn.

Sprenger, Guido. 2005. The way of the buffaloes: Trade and sacrifice in Northern Laos. *Ethnology* 44, 4: 291-312.

Sprenger, Guido. 2006a. Bone transfers: Incomplete replacement in Rmeet ritual exchange. *Taiwan*

Journal of Anthropology 4(1): 79-112.

Sprenger, Guido. 2006b. *Die Männer, die den Geldbaum fällten: Konzepte von Austausch und Gesellschaft bei den Rmeet von Takheung, Laos*. Berlin: LitVerlag.

Sprenger, Guido. 2006c. Out of the ashes: Swidden cultivation in Highland Laos. *Anthropology Today* 22(4): 9-13.

Sprenger, Guido. 2006d. Political periphery, cosmological center: the reproduction of Rmeet socio-cosmic order and the Laos-Thailand border. In Alexander Horstmann and Reed Wadley (eds.) *Centering the margin: Agency and Narrative in Southeast Asian borderlands*, p. 67-84. New York/Oxford: Berghahn.

Sprenger, Guido. 2006e. The end of rituals: A dialogue of theory and ethnography in Laos. *Paideuma 52*: 51-72.

Sprenger, Guido. 2007. From kettledrums to coins: Social transformation and the flow of valuables in Northern Laos. In François Robinne and Mandy Sadan (eds.) *Social Dynamics in the Highlands of Southeast Asia: Reconsidering "Political Systems of Highland Burma" by E.R. Leach*, p. 161-186. Leiden: Brill.

Sprenger, Guido. 2008. Do the Rmeet have clans? In Michel Lorrillard and Yves Goudineau (eds.) *Recherches nouvelles sur le Laos*, p. 559-578. Paris: Ècole française de l'Extrême-Orient.

Sprenger, Guido. 2009. Die Pflege der Spaltung: Die Institutionalisierung von Dissoziation in der Moderne und bei den Schamanen der Rmeet, Laos. *Curare* 32(1): 64-77.

Steinmann, Alfred and Sanidh M. C. Rangsit. 1939. Denkmalformen und Opferstätten der Lawa. Zeitschrift für Ethnologie (Berlin) 71:163-174. [Forms and of monuments and places of sacrifice of the Lawa].

Srisook, W. 2002. *The assimilation of Palaung's lifestyle in Huaysaikhao village Maenawang sub-district Mae-Ai district Chiang Mai province*. MA thesis, Rajabhat Chiang Mai University. [in Thai]

Svantesson, Jan-Olof. 1983. The phonetic correlates of register in Va. In A. Cohen and M. P. R. van den Broeck (eds.) *Abstracts of the 10th International Congress of Phonetic Sciences*, p. 691. Dordrecht: Foris.

Svantesson, Jan-Olof. 1988. U. *Linguistics of the Tibeto-Burman Area, 11*(1):64-133.

Svantesson, Jan-Olof. 1989. Tonogenetic mechanisms in Northern Mon-Khmer. *Phonetica* 46:60-79. Also presented at the 21st International Conference on Languages and Linguistics, Lund.

Svantesson, Jan-Olof. 1991. Hu—a language with unorthodox tonogenesis. In Jeremy H. C. S. Davidson (ed) *Austroasiatic languages: essays in honor of H. L. Shorto*, p. 67-79. London: SOAS, University of London.

Svantesson, Jan-Olof. 1993. Phonetic correlates of register in Paraok. *Reports from Uppsala University*: Linguistics 23 (Papers from the 7th Swedish Phonetics Conference, Uppsala), 102-105.

Svantesson, Jan-Olof, Jingliu, Wang, and Chen Xiangmu. 1981. Mon-Khmer languages in Yunnan. *Asie du Sud-Est et Monde Insulindien* 5.1:39-68.

Ta Văn Thông. nd. *Vocabulaire Mang*. Hà Nôi.

Tantiwuthipakorn, Wattana. 1998. A phonological study of Wa at Ban Santisuk Moo 19, Tambol Patung Mae-chan district, Chiengrai province. MA thesis. Mahidol University.

Taylor, L. F. 1956. The general structure of languages spoken in Burma. *Journal of the Burma Research Society* 39.1: 101-120.

Teeranon, Phanintra. 2007. The intrinsic pitch and intrinsic length of high and low vowels in Mon-Khmer language. *The Mon-Khmer Studies Journal*, 37, 41-52.

Teeranon, Phanintra. 2008. Tonal evolution from the influence of initial consonant clusters: Another look from Lavua', Wa, and Plang. In Ronnakiat, N., Saengmanee, A. (Eds.), A Collection of Essays on Southeast Asian Mon-Khmer Studies 40 85 Linguistics: Festschrift for Professor Theraphan Luangthongkum's 60 Birthday. Phikkhanet Printing Center, Bangkok, pp. 30-42. [in Thai]

Than Sein Thit, et al. 1962. *The Wa people*. Rangoon: Yan Aung.

Thành Thiên. 1972. Nhóm Kháng [The Kháng]. In *Những nhóm dân tộc thuộc ngữ hệ Nam Á ở Tây Bắc Việt Nam [The ethnic minorities of Austroasiatic origin in Northwest Vietnam]*, 313-367. Hànội: Nhà Xuất bản Khoa học Xã hội.

Theraphan, Thongkum Luang. 1988. Phonation types in Mon-Khmer languages. In Fujimura Osamu (ed.) *Vocal Fold Physiology*. New York: Raven Press.

Thomas, David D. 1964. A survey of Austroasiatic and Mon-Khmer comparative studies. *Mon-Khmer Studies* 1:149-163.

Thomas, David D. 1969. Mon-Khmer in North Vietnam. *Mon-Khmer Studies* 3:74-75.

Thomas, David D. 1973. A note on the branches of Mon-Khmer. *Mon-Khmer Studies* 4:139-141.

Thomas, David and Robert K. Headley. 1970. More on Mon-Khmer sub-groupings. *Lingua* 25:398-418.

Too, Gwat. 1982. *The phonology of Palaung*. Mimeograph 12 pages.

Vương, Hoàng Tuyên. 1963. *Các dân tộc nguồn gốc Nam Á ở miền Bắc Việt-Nam [The ethnic minorities of Austroasiatic origin in North Vietnam]*, 251-207. Hànội: Nhà Xuất bản Khoa học Xã hội.

Wang, You-ming. 1984. My views on using the Wa script to promote education in Wa minority areas. Min-zu Gong-zuo, no. 12.

Watkins, Justin. 2002. *The Phonetics of Wa: Experimental Phonetics, Phonology, Orthography and Sociolinguistics*. Canberra: Pacific Linguistics.

Wenk, K. 1965. Drei Lawa vokabularien aus Nordthailand. *Oriens Extremus* 12:107-127.

Wichiekeeo, Aroonrut. 1998. Ethnic groups in Chiangmai. In Hayashi Yukio (compiler) *Inter-ethnic relations in the making of Mainland Southeast Asia and Southwestern China*, Vol. 1, p. 40-59. CSEAS Kyoto University, Bangkok Office.

Wichiekeeo, Aroonrut. 2000. Lawa (Lua): A study from palm-leaf manuscripts and stone inscriptions. In Hayashi Yukio and Yang Guangyuan (eds.) *Dynamics of ethnic cultures across national boundaries in Southwestern China and mainland Southeast Asia: Relations, societies, and languages*, p. 138-53. Chiang Mai: Ming Muang Publishing House.

Wichiekeeo, Aroonrut. 2002. 'Lua leading dogs, toting chaek, carrying chickens': some comments. In Hayashi Yukio and Aroonrut Wichiekeeo (eds.) *Inter-ethnic relations in the making of Mainlaind Southeast Asia and Southwestern China*, p. 1-22. Bangkok: Amarin Printing and Publishing Company Ltd.

Wu You. 1962. Wo guo dai zu he Bulang zu seng jie ming cheng yu yuan [Etymology of Tai and Plang terms for the ranks of Buddhist monks in China]. *Xiang dai fo xue* 2:48ff.

Wurm, S. A. and Hattori Shirô, eds. 1983. *Language atlas of the Pacific area, Part II: Japan area, Taiwan (Formosa), Philippines, Mainland and Insular South-East Asia*. Canberra: Australia National University.

Xiaohang, Qin. 1998. Some Kam-Tai loan words in Mon-Khmer languages. *Mon-Khmer Studies* 28:85-89.

Xiu, Dingben. 2008. De'angzu jianshi [A brief history of De'ang ethnic group]. Beijing: Minorities Press.

Xueming, Chén. 1986. *Wa language textbook vol.3: Seventh year elementary*. Yunnan: Yunnan Ethnic Publishing House.

Xueming, Chén, et al. 1986. Da-angyu jianzhi. Beijing: National Minorities Press. (A description of the Ta-ang language (Gold Palaung)). [in Chinese]

Young, Gordon. 1962. *The hill tribes of Northern Thailand*. Bangkok: Siam Society.

Young, M. V. 1934. *Lai Yohan, Gospel of John in Wa*. American Baptist Mission Press, Rangoon.

Young, M. V. 1935. *Lai Mahteh, Gospel of Matthew in Wa*. American Baptist Mission Press, Rangoon.

Zhao, Fu-rong. 1989. On the importance of promoting the Wa script. *Thai-Yunnan Project Newsletter* No. 7, December 1989. Canberra: Australian National University.

Zhou, Zhizhi and Yan Qixiang. 1981. *Sketch grammar of Wa* (in Chinese). Beijing: National Publishing House.

Zhou, Zhihi and Yan Qixiang. 1982. *A Brief Overview of the Wa Language*. Beijing: National Publishing House.

Zhou, Zhizhi and Yan Qixiang. 1983a. A brief description of Blang. *Minzu yuwen* [*National minority philology and linguistics*, Beijing], 2:71-81.

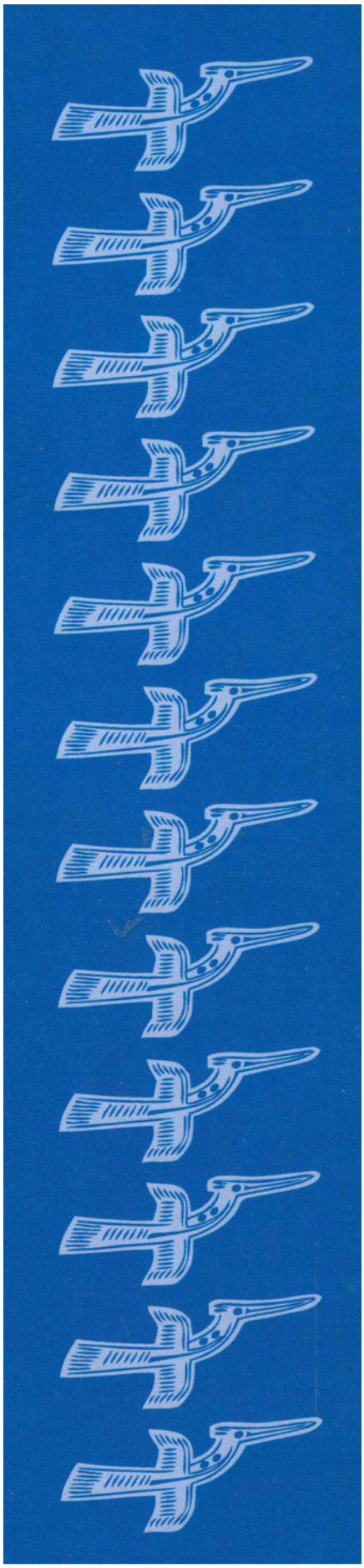
Zhou, Zhizhi and Yan Qixiang. 1983b. Approaching the consonantism of ancient Wa from phonological correspondences among the dialects of contemporary Wa. *Yu-yen Yen-chiu* [*Journal of Linguistic Research*, Beijing, Institute of Nationality Studies, Chinese Academy of Social Sciences], 1:192-211.

Zhou, Zhizhi and Yan Qixiang. 1984. *Wa yu jian zhi* [A concise description of Wa language]. Kunming: Yunnan Ethnic Publishing House.

Zide, Norman H., ed. 1966. *Studies in comparative Austroasiatic linguistics*. The Hague: Mouton.

Mon-Khmer Studies

Volume 42



Author: Nathaniel CHEESEMAN, Jennifer HERINGTON, and Paul SIDWELL
Title: *Bahnaric linguistic bibliography with selected annotations.*
Pages: xxxiv-xlvii
Date received: 31/5/2013
Revised text accepted: 28/10/2013

Volume 42 Editors:
Paul Sidwell
Brian Migliazza

ISSN: 0147-5207
Website: http://mksjournal.org

Published by:

Mahidol University (Thailand)

SIL International (USA)

Bahnaric linguistic bibliography with selected annotations

Nathaniel CHEESEMAN
Linguistics Institute, Payap University, Thailand

Jennifer HERINGTON
Linguistics Institute, Payap University, Thailand

Paul SIDWELL
Australian National University

Abstract
The purpose of this bibliography is to assist Bahnaric linguists, translators, and scholars by indexing and cataloguing Bahnaric reference materials. References are first organized by linguistic domain. Next, references are organized historically via a helpful language index. Lastly, an author index is provided. This work also includes an updated Bahnaric map and a brief description of Bahnaric linguistic features. Readers will be drawn into the debate on Bahnaric classification and the question of which languages need to be deleted or added to the Bahnaric family.
Keywords: Bahnaric, bibliography, reference.
ISO 639-3 language codes: alk, bdg, crw, cua, hld, hal, hre, jeh, jeg, xkk, lmm, rka, kgc, tgr, kta, krv, kxy, kpm, krr, brb, lbo, cma, moo, cmo, mng, mnn, nev, oyb, ren, rmx, spu, sed, skk, sqq, sti, stt, tkz, tdf, tpu, thx, tdr, stg.

Abbreviations

eds editors
ICSTLL International Conferences on Sino-Tibetan Languages and Linguistics
LI Linguistics Institute at Payap University in Chiang Mai Thailand
mimeo mimeograph.
ms manuscript
n.d. no date (date of publication unknown)
s.l. *sine loco* (place of publication unknown)
s.n. *sine nomine* (publisher unknown)
SIL Summer Institute of Linguistics
UND University of North Dakota
USIS United States Information Service
Vol volume

Abbreviations of Journal Titles

ASEMI Asie du Sud-Est et Monde Insulindien
BEFEO Bulletin de l'École française d'Extrême-Orient
BSEI Bulletin de la Société des Etudes Indochinoises, Saigon

1. Introduction

The Bahnaric branch of Austroasiatic is arguably the most diverse of the family, with currently 40 languages identified (Lewis 2013) , spoken by communities in central and southern Vietnam, southern Laos and eastern Cambodia. This bibliography on Bahnaric linguistics is an effort to support linguists working in Mainland Southeast Asia by providing a useful resource to find information on these languages. After a short overview on common Bahnaric linguistic features, the bibliography is organized by linguistic domains. The bibliography concludes with an index of Bahnaric languages, followed by an author index.

CHEESEMAN, Nathaniel, Jennifer HERINGTON and Paul SIDWELL. 2013. Bahnaric linguistic bibliography with selected annotations.
Mon-Khmer Studies (Notes, Reviews, Data-Papers). 42:xxxiv-lxvii

Figure 1: Map of Bahnaric language family by Eva Ujlakyova, 2013.

Some of the first works on the Bahnaric languages were published in the late 1800's. Since the French had influence in Indo-China during this period, much of the early Bahnaric documentation is written in French. During the period between 1957 and 1975, linguists working for SIL and allied organisations compiled large amounts of Bahnaric data, publishing linguistic papers and articles, as well as vernacular publications in these languages. Some Bahnaric languages, such as Sedang, Bahnar, Central Mnong, and Chrau have been researched extensively (e.g. The Sedang language was studied by Ken Smith; Chrau was researched by David Thomas; John and Elizabeth Banker worked in the Bahnar language; Henry and Evangeline Blood researched and published materials on Eastern Mnong.) Other languages, for example Romam, have very little published research. Romam is only briefly mentioned in the Ethnologue, but is not discussed in any known English language sources.

CHEESEMAN, Nathaniel, Jennifer HERINGTON and Paul SIDWELL. 2013. Bahnaric linguistic bibliography with selected annotations.
Mon-Khmer Studies (Notes, Reviews, Data-Papers). 42: xxxiv-lxvii

Today many of these publications are available in Chiang Mai, Thailand, in the library of the Linguistics Institute at Payap University. More Bahnaric resources are also listed online in the SIL Bibliography. The Mon-Khmer Studies Journal (MKS), which was established in 1964, contains numerous detailed scholarly articles about Bahnaric grammar and phonology. SEALANG.net is another site providing online access to many of the Bahnaric resources.

2. Linguistic Classification

The overall structure of the Austroasiatic phylum is not a matter of consensus among concerned scholars; Diffloth and Zide (1992) regard Bahnaric languages as members of an Eastern division of the Mon-Khmer branch of the Austroasiatic language family while Sidwell (2010) classifies Bahnaric on a direct line from proto-Austroasiatic. Much of the literature on the Bahnaric languages—perhaps more so than any other branch—has focused on both historical reconstruction through lexical and phonological comparison and on various lexicostatistical studies. The classification of Bahnaric languages nevertheless remains in a state of flux.

The 17th edition of the Ethnologue reports four main sub-branches within Bahnaric: South, Central, North, and West (Lewis, Simons and Fennig 2013). These four main divisions date back to 1970's. However, there is some debate about the number of subdivisions in the Bahnaric languages. Sidwell (2000) agrees with Adams (1989) on five discrete groups, but acknowledges that others have classified the Bahnaric languages into as few as three and as many as eight subclasses. Comparison of Tables 1 and 2 shows just how much analyses have changed over the past decade:

West	Central	North	East	South
Loven (Jru), Nyaheun, Prou, Ôi, Thre, Laveh, (?)Brao, Krung, Kravet, Sok, Sapuan, Ceng (Jeng) (?)Suq (Sou)	Bahnar, Alak	Rengao, Sedang, Halang, Jeh (Dié), Monom (Bonâm), Hrê (Davak), Todrah (Didrah)	Cua (Kor, Traw), Takua	Stieng, Central Mnong, Southern Mnong, Eastern Mnong, Köho (Sre), Chrau (Jro)

Table 1: Bahnaric subclassification according to Sidwell (2000:4)

While Table 1 indicates the consensus view at the end of the 20th century, which was very much a reflection of historical tendencies to work within national boundaries and traditions. The classification offered a decade later in Table 2 is the outcome of detailed phonological and lexical reconstruction. Still, this more current grouping must be considered provisional, as in all likelihood, it will expand and develop as further work progresses.

West Bahnaric	Central Bahnaric	North Bahnaric	East Bahnaric
Jru' (Laven), Juk, Su' Nyaheun Oi, The, Sok, Sapuan, Cheng Brao, Laveh, Krung, Kravet	Taliang (Kasseng) Alak Central South Tampuon Bahnar South Bahnaric Chrau Sre Stieng Mnong	Halang, Kayong Jeh Kotau Tadrah, Modrah Sedang Hrê Mơnơm (Bơnâm) Rengao Kaco', Ramam	Cua (Kor)

Table 2: Revised Bahnaric subclassification according to Sidwell (2009:203)

Another issue faced by those researching Bahnaric languages is the problematic use of language names. For example Stieng: the Ethnologue (2013) splits Stieng into two varieties, Bulo and Budeh. However, Sidwell (2009) and Adams (1989) would maintain that Stieng is actually one language group. Therefore when Stieng is referenced throughout the bibliography it is sometimes difficult to know which variety is being indicated. Another issue deals with Tareng and Kasseng. The Ethnologue (2013) classifies Tareng and Kasseng as individual Katuic languages. However, Sidwell (n.d.) and Diffloth (1997) hold that Tareng and Kasseng should be merged and considered Bahnaric. Additionally, there has lately been confusion over Kaco and Ramam, as discussed in Edmondson, Gregerson and Sidwell (2011).

3. General linguistic features of Bahnaric languages

The Bahnaric languages are phonologically very similar to many Mon-Khmer languages, with the characteristic large vowel inventory. For example, Koho Sre reflects a common pattern with nine vowel qualities with contrastive length plus diphthongs /ia/ and /ua/ (Le 2003). Phonological words in Bahnaric may be mono- or disyllabic, with the latter generally treated as sesquisyllablic (iambic stress). An illustrative example is Smith's maximum word template for Sedang (1979:22):

$$(C_pV_p)(C_m)C_i(C_m)V(G)(N)(C_f)(R)$$

p: presyllable, m: main syllable, i: initial,
G: glide, N: nasalized vowel, f: final, R: register

The vowels in stressed and unstressed syllables differ in their phonological status. Typically only a single non-contrastive vowel, usually schwa or conditioned variant, and a reduced consonant inventory occur in unstressed syllables, while the main syllables carry the full and rather large range of contrastive vowels and consonants. Table 3 shows the full inventory of Sedang consonants.

Manner	Labial	Alveolar	Palatal	Velar	Glottal
stop –voice	p	t	c	k	ʔ
stop +voice	b [ᵐb]	d [ⁿd]	ɟ [ᶮɟ]	g [ᵑg]	
implosive	ɓ	ɗ	(ʄ)		
nasal	M	n	ɲ	ŋ	
nasal glottalised	ˀm	ˀn	ˀɲ	ˀŋ	
nasal -voice	m̥	n̥	ɲ̊	ŋ̊	
fricative		s, ʂ			h
approximant	W	l, r	j		
approximant glottalised	ˀw	ˀl, ˀr	ˀj		
approximant -voice	w̥	l̥, r̥	j̊		

Table 3: Sedang consonants based on Smith (2009)

Like most Austroasiatic languages, the Bahnaric languages are not tonal, but many—such as Sedang—contrast two phonation types (e.g. most if not all North Bahnaric languages contrast breathy versus modal voice, Sedang oddly contrasts creaky versus modal voice).

Bahnaric languages do not have a terribly complex morphology, corresponding to the type characterized by Diffloth and Zide: "[Mon-Khmer] morphology practically never indicates syntactic agreement. This morphology is usually derivational and nonproductive. Its typical function is to change the grammatical class or subclass of the base to which it is attached" (1992:141). For example, in Sedang, Smith (1969) details the use of a causative affix, a reciprocal affix, a nominalizing infix, and several other more minor affixes. Infixation is a common practice

across all Bahnaric languages, with a monophthong, nasal, or liquid following the onset of the word base (Diffloth and Zide 1992). This infixation is often the origin of sesquisyllables in Bahnaric languages.

Syntactically, Bahnaric languages follow a typically Mon-Khmer subject-verb-object pattern. However, when no object is present, it is possible in many languages to have the verb in the first position. This is usually limited to certain constructions, and not an option in all sentence types (Diffloth and Zide 1992). Along with the SVO pattern, Bahnaric languages also place "the possessed after the possessor, the attribute after the noun, and deictics at the end of the noun phrase" (1992:141), a frequent pattern in Southeast Asian languages.

4. Bibliography of Bahnaric linguistics

It is clearly evident from the vast expanse of Bahnaric works that additional references could be added to this bibliography. For further bibliographic research on Bahnaric the following online sources should be consulted:

-The Luce Collection
-SEALANG.net
-Franklin E. Huffman's (1986) Bibliography.
-John F. Embree and Lillian Ota Dotson's 1950 Bibliography.
-Linguistic Institute Library at Payap University (Scriptures and Hymns vernacular works).
-Bibliography of Laos and Ethnically Related Areas by Joel M. Halpern
-French, Vietnamese, Khmer, and Lao language works related to Bahnaric.
-GIAL Library
-ANU Library
-Yale Library
-Cornell Library
-WorldCat
-Other major university library systems, especially those with vibrant Asian linguistics programs

4.1 Comparative Historical Linguistics

Adams, Karen Lee. 1989. *Systems of numeral classification in the Mon-Khmer, Nicobarese and Aslian subfamilies of Austroasiatic*. Pacific Linguistics B, no. 101. Canberra: The Australian National University.

See page 33 for Adams classification of Bahnaric.

Bauer, Christian. 1987. Reanalyzing reanalyses in Katuic and Bahnaric. Mon-Khmer Studies (MKS) 16-17.143–154.

Blood, Henry F. 1966. *A reconstruction of Proto-Mnong*. Grand Forks: Summer Institute of Linguistics, University of North Dakota.

This is Henry Blood's thesis for completion of Master's of Arts at the University of North Dakota. This work contains a one-and-a-half page bibliography and is 118 pages in length.

Diffloth, Gérard. 1991. Tarieng (-Alak), a new branch of Bahnaric. Bangkok-Chiangmai: s.n.

This was a conference paper presented at the 24th International Conference on Sino-Tibetan languages and Linguistics. Referenced by L-Thongkum (1997).

----. 1997. Introduction to comparative Mon-Khmer. Unpublished manuscript. s.l. ms.

Referenced by L-Thongkum (1997). Chapter 4 is written about Bahnaric languages.

Diffloth, Gérard, and Norman Zide. 1992. Austro-Asiatic languages. *International Encyclopedia of Linguistics*, (ed.) by William Bright, 1:137–142. New York: Oxford University Press.

Contains overview on Bahnaric linguistic features.

Edmondson, Jerold A., Kenneth J. Gregerson, and Paul Sidwell. 2009. The North Bahnaric Clade: A Computational Approach. Mahidol University: Research Institute for Language and Cultures of Asia. http://icaal.org/ICAAL-4.2.pdf accessed 2 October 2013.

Very informative tree diagrams displaying North Bahnaric languages comparisons.

Ferlus, Michel. 1971. Simplification des Groupes Consonantiques dans deux Dialectes Austroasiens du Sud-Laos [The Simplification of Consonant Clusters in two Austroasian Dialects of Southern Laos]. *Bulletin de la Société de Linguistique de Paris* 66.1–389.

Laven, Nyaheun, and Brao mentioned.

----. 1974. Review of A Phonological Reconstruction of Proto-North-Bahnaric by Kenneth D. Smith. *Asie du Sud-Est et Monde Insulindien (Mouton)* 5.183–185.

Referenced by Smith (1979).

----. 1974. Carte Des Groups Linguistiques Austroasiatiques Dans Le Centre - Indochinois [A Map of Indochinese Central Austroasiatic Language Groups]. *Asie du Sud-Est et Monde Insulindien* 5.

Includes Bahnaric map.

----. 1975. Vietnamien et Proto-Viet-Muong [Vietnamese and Proto-Viet-Muong]. *Asie du Sud-Est et Monde Insulindien* 6.21–55.

----. 1977. Etude d'une Strate de Changements Phonétiques dans l'ancien Cambodge [A Study of a Stratum of Phonetic Changes in Ancient Cambodia]. *Mon-Khmer Studies* 6.59–67.

Nyaheun, Laven, Brao, Sapuan, Jeng all referenced.

----. 1978. Reconstruction de /TS/ Et /Tš/ en Mon-Khmer. *Mon-Khmer Studies* 7.1–38.

----. 1996. Langues et Peuples Viet-Muong [Viet-Muong Languages and Peoples]. *Mon-Khmer Studies* 26.7–28. http://www.639-tree.org/archives/mks/pdf/26:7-28.pdf accessed 8 October 2013.

The Halang language is mentioned in this work.

----. 2004. The origin of tones in Viet-Muong. *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society*, 297–313. Tempe, Arizona: Arizona State University, Program for Southeast Asian Studies. http://sealang.net/sala/archives/pdf4/ferlus2004origin.pdf accessed 8 October 2013.

----. 2009. Le Mot "Sang" en Austroasiatique [The Word "Blood" in Austroasiatic]. *Mon-Khmer Studies* 38.25–38. http://sealang.net/sala/archives/pdf8/ferlus2008le.pdf accessed 8 October 2013.

----. 2010. The Austroasiatic Vocabulary for Rice: its Origin and Expansion. *Journal of the Southeast Asian Linguistics Society* 3.61–76.

Ghosh, Arun. 1988. *Bibliotheca Austroasiatic A classified and Annotated Bibliography of the Austroasiatic People and Languages*. 1st ed. Netaji Institute for Asian Studies Monograph Series 1. Calcutta: Firma KLM Private Limited.

This work contains a Mon-Khmer language tree, along with many bibliographical references.

Gregerson, Kenneth, Kenneth Smith, and David Thomas. 1976. The Place of Bahnar within Bahnaric. *Austroasiatic Studies*, 371–406. Oceanic linguistics special publication 13. Honolulu: The University Press of Hawaii.

Referenced by Jacq and Sidwell (2000).

Gregerson, Kenneth J. 1990. Dân tộc Rơngao, Bahnar và các dân tộc lân cận. [Ethnicity Rongao, Bahnar and neighboring nations] *Journal of Social Sciences* 5.80–90.

Referenced by the SIL Bibliography (2012). This is a comparative work between Bahnar and Rengao.

Huffman, Franklin E. 1976. The register problem in fifteen Mon-Khmer languages. *Oceanic Linguistics Special Publications* 13.575–589. http://sealang.net/sala/archives/pdf8/huffman1976register.pdf accessed 14 May 2013.

---- 1976. The Relevance of Lexicostatistics to Mon-Khmer Languages. *Austroasiatic Studies Part 1*, ed. by Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, 539–574. Oceanic Linguistics Special Publication 13. Honolulu: The University Press of Hawaii.

---- 1978. On the Centrality of Katuic-Bahnaric to Austroasiatic or What is Katuic-Bahnaric, and why does it have all those ties to the rest of Austroasiatic? Working Paper. s.l., ms. http://people.anu.edu.au/~u9907217/lexico/Huffman1978.pdf accessed 14 May 2013.

Jacq, Pascale, and Paul Sidwell. 1999. *Sapuan (Səpuar)*. München: Lincom Europa

Referenced by Jacq and Sidwell (2000).

----. 2007. Orientation origins: the source of Jru' cardinals. *SEALS XII Papers from the 12th Annual Meeting of the Southeast Asian Linguistics Society 2002*, 101–106. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, The Australian National University.

Keller, Charles. 1992. Roman/Khmer Script Comparative Charts for Brao. Working Paper. s.l., ms. 495.95 W. Bahnaric, Brao folder. Linguistic Institute Library Payap University.

Kingsada, Dr. Thongpheth. 2011. Languages and Ethnic Classification in the Lao PDR. *Social Sciences Journal* 4.40–53. http://www.lass.gov.la/attachments/article/81/Dr%20Thongphet.pdf06.pdf accessed 14 May 2013.

L-Thongkum, Theraphan. 2001. *Languages of the tribes in Xekong province Southern Laos: a foundation for research and development*. Bangkok: Chulalongkorn University Press.

This work is written in Thai; it contains 670 pages. The latter half of the work, beginning in Chapter 5, is dedicated to Bahnaric languages. L-Thongkum gives descriptions of various Bahnaric classification theories, as well as comparative historical reconstructions of proto-Bahnaric. L-Thongkum, Jacq & Sidwell, Smith, and again Sidwell's reconstructions are all compared in this section. Chapter 6 covers sound changes in Bahnaric languages. The book concludes with a large detailed map of Xekong province in Laos.

----. 1997. The place of Lawi Harak and Tariang within Bahnaric. Mon-Khmer Studies (MKS) 27.109–117.

----. 2001. Is the Northwest Bahnaric sub-branch a fact or fiction? Mahidol University.

Referenced by http://ling.arts.chula.ac.th/contents/File/thecvthai09.pdf, and accessed 14 May 2013.

Lewis, M. Paul, Gary F. Simons, and Charles D. Fennig (eds.), 2013. *Ethnologue: Languages of the World, Seventeenth edition.* Dallas, Texas: SIL International. Online version: http://www.ethnologue.com accessed 14 May 2013.

The Ethnologue is the world's authority on language references. It references forty Bahnaric languages. Not all Bahnaric languages mentioned by researchers are reported in the Ethnologue.

Mann, Noel, Wendy Smith, and Eva Ujlakyova. 2009. Linguistic Clusters of Mainland Southeast Asia: a Description of the Clusters. Chiang Mai: Linguistics Institute Payap University. http://li.payap.ac.th/images/stories/survey/Linguistic%20Clusters%20of%20Mainland%20Southeast%20Asia%20A%20Description%20of%20the%20Clusters.pdf accessed 14 May 2013.

In this report, pages 7-10 are specifically written to give a general overview of Bahnaric languages. Thirty-nine Bahnaric languages are mentioned along with the general literacy situation for this language cluster. A simple map is included to show the general area where Bahnaric languages can be located. There are several research questions in this paper which identify specific questions related to Bahnaric languages. There is also a helpful one page selected bibliography.

Parkin, Robert. 1991. A Guide to Austroasiatic Speakers and Their Languages. Oceanic Linguistics Special Publication 23. Honolulu: University of Hawaii Press.

Pages 68-82 focus specifically on the Bahnaric languages. He briefly mentions the Bout ethnic group under the Bahnaric section (page 79).

Schmidt, Willhelm. 1905. Grundzuge einer lautlehre der Mon-Khmer Sprachen. *Denkschr Akad. Wiss.* 51.1–233.

Referenced by Arun Ghosh (1988). This work compares presumed cognates in four Mon-Khmer languages includeing Bahnar and Stieng.

Shelden, Howard. 1979. A Reconstruction of Proto-South Bahnaric. Class Paper. University of Texas Arlington, ms.

Resource can be located at the Linguistics Institute Library at Payap University in Chiang Mai, Thailand.

Sidwell, Paul. 1994. Proto-Jeh-Halang and Proto-North-Bahnaric revisited. La Trobe University Working Papers in Linguistics 7.133–142. http://arrow.latrobe.edu.au:8080/vital/access/manager/Repository/latrobe:33054;jsessionid=AEBF207CBB40B4C6A1FF0761AC765CDD accessed 1 October 2013.

----. 1998. A reconstruction of Proto-Bahnaric. Melbourne: Department of Linguistics and Applied Linguistics, University of Melbourne PhD diss.

Referenced by Google Scholar.

----. 1999. A report on the status and distribution of Bahnaric languages and speakers on the Boloven Plateau, Lao PDR. History of Language 5.10–18.

Referenced by Wikipedia accessed 14 May 2013. Perhaps this is similar in nature to Barbara Wall's 1975 work written in French?

----. 1999. Proto South Bahnaric: a reconstruction of a Mon-Khmer language of Indo-China. Pacific Linguistics B-116. Canberra: Pacific Linguistics Publishers.

----. 1999. The Austroasiatic Numerals 1 to 10 from a Historical and Typological Perspective. *Numeral Types and Changes Worldwide*, 253–271. Berlin: John Benjamins Publishing Company.

Rengao and Chrau mentioned. In the hyperlink above pages 253 and 254 were not scanned and made available online.

----. 1999. The Role of Historiography in Evaluating the Results of Comparative Linguistic Work-A-Case Study. *The Emergence of the Modern Language Sciences: Studies on the transition from historical-comparative to Structural Linguistics in Honour of E.F.K. Koerner*, 597–608. Amsterdam: John Benjamins Publishing Company.

First four pages available online above through Google Books.

----. 2002. Genetic Classification of the Bahnaric Languages: A Comprehensive Review. Mon-Khmer Studies (MKS) 32.1–24.

Sidwell reviews the different linguists who have attempted to classify Bahnaric languages. Sidwell gives evidence that there are three main strands in the Bahnaric language family—Western, Central, and Northern dialects. Sidwell later amends his three branch theory in MKS 39, where he cites Cua as a distinct fourth branch of Bahnaric.

----. 2002. The Mon-Khmer Substrate in Chamic: Chamic, Bahnaric & Katuic Contact. *SEALS XII*, 113–128. Canberra: Research School of Pacific and Asian Studies, Australian National University. http://pacling.anu.edu.au/series/SEALS-PDFs/SEALS-12.pdf#page=123. accessed 1 October 2013.

----. 2003. A note on the reconstruction of proto West Bahnaric and investigation of early West Bahnaric-Katuic contact. Mon-Khmer Studies (MKS) 33.159–166. http://sealang.net/archives/mks/pdf/33:159-166.pdf accessed 8 October 2013.

----. 2005. Proto Katuic Phonology and Implications for Sub-grouping, E1:193–204. Canberra: Pacific Linguistics. http://sealang.net/sala/archives/pdf8/sidwell2005proto.pdf. accessed 1 October 2013.

Mentions Bahnar and Bahnaric. Helpful classification table from Diffloth referenced.

----. 2007. Comparative Mon-Khmer Linguistics in the 20th Century: Where From, Where To? *Austro-Asiatic Linguistics: In memory of R. Elangaiyan*, 38–104. Mysore: Central Institute of Indian Languages. http://www.sealang.net/monkhmer/sidwell2007comparative.pdf accessed 1October 2013.
Overview of research done on Bahnaric languages p.9-14

----. 2008. Issues in the morphological reconstruction of Proto-Mon-Khmer. *Morphology and Language History*, 251–265. Amsterdam Studies in the Theory and History of Linguistic Science 4. Amsterdam: John Benjamins B.V.

Brief mention of Bahnar.

----. 2009. The Austroasiatic central riverine hypothesis. Ho Chi Minh City: s.n. http://jolr.ru/files/(51)jlr2010-4(117-134).pdf accessed 14 May 2013.

----. 2009. *Classifying the Austroasiatic languages: History and state of the art*. Vol. 76. Muenchen: LINCOM Europa.

Pages 73-86 focus on Bahnaric classification.

----. 2009. How many branches in a tree? Cua and East (North) Bahnaric. *Discovering History through Language: Papers in Honour of Malcolm Ross*, 193–204. Canberra: Pacific Linguistics.

----. 2010. The Austroasiatic central riverine hypothesis. Journal of Language Relationship 4.117–134.

----. 2011. Katuic-Bahnaric: Austroasiatic sub-family or convergence area? *Austroasiatic Studies: Papers from the ICAAL4: Mon-Khmer Studies Journal Special Issue No. 3*, 100–111.

Dallas; Canbera; Salaya: SIL International; Pacific Linguistics; Mahidol University. http://icaal.org/ICAAL-4.2.pdf. accessed 2 October 2013.

----. 2012. Should we Reconstruct Decimal or Non-Decimal Counting for Proto-Austroasiatic? Power Point Presentation. Paper presented at the SEALS XXII, Bangkok. http://jseals.org/seals22/sidwell2012shouldp.pdf. accessed 2 October 2013.
Brief mention of Bahnaric.

Sidwell, Paul, and Roger Blench. 2011. The Austroasiatic Urheimat: the Southeastern Riverine Hypothesis. *The Dynamics of Human Diversity*, 315–344. Canberra: Pacific Linguistics. http://rb.rowbory.co.uk/Archaeology/SE%20Asia/SR09/Sidwell%20Blench%20offprint.pdf accessed 1 October 2013.

Sidwell, Paul, and Pascale Jacq. 2003. *A handbook of Comparative Bahnaric*. Vol. 1. Pacific Linguistics 551. Canberra: Research School of Pacific and Asian Studies, Australian National University.
Covers West Bahnaric languages including the Juk language.

Smith, Kenneth D. 1967. Sedang Dialects. Bulletin de la Société des Études Indochinoises 42. Nouvelle Série.196–255.

Master of Arts thesis from the University of North Dakota.

----. 1978. An automatic typewriter assist to comparative linguists, with application to 30 Mon-Khmer languages. Mysore, India: s.n.

----. 1978. The Payoff from Rhyming Dictionaries: Distribution and Frequency of Word-Final -VCs and a Bridge between Bahnaric and Katuic. Mysore: s.n.

----. 1981. A Lexico-statistical study of 45 Mon-Khmer languages. Linguistics across Continents, ed. by Andre Gonzalez and David Thomas, 180–205. LSP Monograph Series 2. Manila: SIL.

Twenty-four of the forty-five languages Smith compares are Bahnaric languages. In this work, Smith reaches some interesting conclusions about Bahnaric languages. One conclusion is Cua should be seen as an East Bahnaric language. Sidwell in his phonological research (2010), makes the same conclusion. Cua acts as a bit of a "wild card" language within the Bahnaric family; over the decades, its sub-classification has been uncertain. As a follow-up to this work, see Sidwell (2002 and 2010) in MKS.

----. 1974. A Computer Analysis of Vietnam Language Relationships. Work Papers, 18:99–113. Grand Forks: Summer Institute of Linguistics University of North Dakota.

----. 1992. The -VC rhyme link between Bahnaric and Katuic. Mon-Khmer Studies (MKS) 18-19.106–159.

Thomas, David D. n.d.(a) Communicatives, Existives, and Statives in Proto-South Bahnaric. Ms. Mahidol University, Salaya, Thailand, ms. Linguistic Institute Library Payap University, Chiang Mai, Thailand.

This work mentions the following languages: Central Mnong, Eastern Mnong, Chrau, Stieng, and Koho. This work was also published in Austroasiatic Languages Essays in honour of H.L. Shorto.

----. 1959. Language Relationships in Kontum Province. s.l., ms. 495.96 General North Bahnaric Folder. Linguistic Institute Library Payap University, Chiang Mai Thailand.

The languages compared briefly in this paper are Bahnar, Rengao, Halang, Sedang, and Jeh.

----. 1960. Các ngừtộc trong tinh Kontum. Van-hóa A-chau 3.58–60.

This work mentions the following languages: Bahnar, Rengao, Halang, Sedang, and Jeh.

----. 1964. A survey of Austroasiatic and Mon-Khmer comparative studies. Mon-Khmer Studies (MKS) 1.149–163.

----. 1966. Mon-Khmer subgroupings in Vietnam. Studies in comparative Austroasiatic linguistics, ed. by N. Zide, 194–202. The Hague: Mouton. http://sealang.net/sala/archives/pdf8/thomas1966mon.pdf accessed 14 May 2013

----. 1968. ----. Letter.

This is a letter David Thomas wrote to Rich, (last name unknown). In the letter Thomas compares Kayong, Rengao, and Sedang 2, with Sedang 1, Bahnar, Jeh, Bonam, Hre, and Cua. This is probably one of the early comparisons of North Bahnaric languages. This item is archived at the Linguistic Institute Library Payap University, Chiang Mai, Thailand. 495.96 North Bahnaric Folder.

----. 1973. A Note on the Branches of Mon-Khmer. Mon-Khmer Studies (MKS) 4.139–141.

----. 1979. The place of Alak, Tampuan, and West Bahnaric. Mon-Khmer Studies (MKS) 8.171–186.

----. 1987. On early Monic, Vietic and Bahnaric relations. Mon-Khmer Studies (MKS) 16-17.177–179.

Thomas, David D., and Marilyn Smith. 1967. Proto-Jeh-Halang. Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung 20.157–175.

Referenced by Jacq and Sidwell (2000).

Thomas, David, and Robert K. Headley. 1970. More on Mon-Khmer subgroupings. Lingua 25.398–418.

Wilson, Ruth S. 1966. A comparison of Mu'o'ng with some Mon-Khmer languages. Studies in Comparative Austroasiatic Linguistics. 203–213. http://sealang.net/sala/ accessed 14 May 2013.

4.2 Grammatical Descriptions

Banker, Elizabeth. 1964. Bahnar reduplication. Mon-Khmer Studies (MKS) 1.119–134.

Banker, John. 1964. Transformational paradigms of Bahnar clauses. Mon-Khmer Studies (MKS) 1.7–39.

Blood, David L. 1974. Review of: Chrau grammar, by David D. Thomas. *Lingua* 34.283–285. http://www.sciencedirect.com/science/article/pii/0024384174900266 accessed 8 October 2013.

Article can be purchased for $19.95.

Blood, Henry, and Evangeline Blood. 1966. The Pronoun System of Uon Njun Mnong Ro+lo+m. Mon-Khmer Studies (MKS) 2.103–111.

Blood, Evangeline. 1966. Clause and Sentence Types in Mnong Rơlơm. Papers on Four Vietnamese Languages, 23–27. Te Reo Reprints 2. Auckland: Linguistic Society of New Zealand.

Blood, Henry F. 1963. The Vowel System of Uon Njuñ Mnong Rơlơm. Văn-hóa Nguyệt-san 12.951–965.

Referenced by Blood, Henry. (1976).

Bon, Noëllie. 2012. Numeral classifiers of Stieng : A typological and areal approach. Agay. http://jseals.org/seals22/bon2012numeralh.pdf accessed 14 May 2013.

Burton, Eva. 1969. A brief sketch of Cua clause structure. Mon-Khmer Studies (MKS) 3.5–8.

Cohen, Patrick D. 1976. The Noun Phrase in Jeh. MKS 5.139–152.

---- 1966. Presyllables and reduplication in Jeh. Mon-Khmer Studies (MKS) 2.31–53.

Cohen, Nancy. 1976. Some interclausal relations in Jeh. Mon-Khmer Studies (MKS) 5.153–164.

Condominas, George. 1965. Two brief notes concerning Mnong Gar. *Lingua* 15.48–52.

Referenced by Arun Ghosh (1988). Mnong Gar is a dialect of Eastern Mnong.

Cooper, James S. 1966. Halăng Verb Phrase. Papers on Four Vietnamese Languages, 28–34. Te Reo Reprints 2. Auckland: Linguistic Society of New Zealand.

Davis, John J. 1973. Notes on Nyaheun Grammar. Mon-Khmer Studies (MKS) 4.69–75.

Ferlus, Michel. 1980. Formation des Registres et Mutations Consonantiques dans les Langues Mon-Khmer [Formation of Registers and Consonant Changes in the Mon-Khmer languages]. *Mon-Khmer Studies* 8.1–76.
Mentions Bahnar, Hre and early Sedang.

Gradin, Dwight. 1976. Word Affixation in Jeh. MKS 5.25–42.

----. 1976. The Verb in Jeh. MKS 5.43–75.

Gregerson, Kenneth J. 1979. Predicate and argument in Rengao grammar. Summer Institute of Linguistics Publications in Linguistics 61. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington. http://www.sil.org/acpub/repository/15816.pdf accessed 16 May 2013.

This work was written for Gregerson's PhD at the University of Washington.

Gregerson, Kenneth, Jacqueline Jordi, and Charles Keller. 2011. Merge right: iambic effects, grammaticalization and cliticization in Brao grammar. MKS Special Issue No. 3: Austroasiatic Studies: Papers from ICAAL 4: Vol 2. 162–177.

Haupers, Ralph. 1962. Word-final syllables in Stieng. *Van-hóa Nguyêt-san* 11.846–848.

Referenced by Arun Ghosh (1988).

----. 1964. Stieng Papers. Grand Forks, ms. Copied from manuscript at SIL-UND library. Linguistic Institute Library Payap University, Chiang Mai, Thailand.

As noted on the first page of the work, this work includes: "Nouns and Noun constructions, syntax, proto Cambodian Stieng, proto Bahnar Stieng, Analysis of a text, and texts."

Jacq, Pascale. 2006. The who and how of Nyaheun /hi/. Mon-Khmer Studies (MKS) 36.1–19.

Keller, Charles. 1982. Tampuan preliminary grammar questionnaire. s.l., ms. Copied manuscript at Linguistic Institute Library, Payap University, Chiang Mai, Thailand.

----. 1976. A grammatical sketch of Brao, a Mon-Khmer language. Grand Forks: Summer Institute of Linguistics, University of North Dakota.

This work is a part of the Summer Institute of Linguistics series called Work Papers. Volume 20, supplement 1. One page bibliography included.

Kommala, Chinda. 1978. Preliminary Grammar Questionnaire: Oi. s.n. Folder 495.956. Copy held at Summer Institute of Linguistics Library Chiang Mai.

Referenced by Jacq and Sidwell (2000). Oi is an alternate name for the Oy language of Laos.

----. 1978. Preliminary Grammar Questionnaire: Nyaheun. s.n. Folder 495.956. Copy held at Summer Institute of Linguistics Library Chiang Mai.

Referenced by Jacq and Sidwell (2000).

Lai, Phạm Hũu. 1964. Âm – hệ Ngôn – ngũ' Bahnar. Dalat: s.n. Book at Linguistic Institute Library, Payap University, Chiang Mai, Thailand.

This is a term paper from a class Kenneth Smith taught. This paper is 33 pages in length.

Leitch, Myles. 1981. A Study of Rengao in Proto-North Bahnaric. Class Paper. s.l., ms. 495.95 Bahnaric Folder. Linguistic Institute Library Payap University.

This is a term paper from a class Dr. Ken Smith taught. This paper is 33 pages in length.

Manley, Timothy M. 1972. Outline of Sre Structure. Oceanic Linguistics Special Publication 12. Honolulu: University of Hawaii Press.

Referenced by Jacq and Sidwell (2000).

Miller, Vera Grace. 1976. An Overview of Stiêng Grammar. Supplement 3. Vol. XX. Work Papers. Summer Institute of Linguistics University of North Dakota Sessions.

Master of Arts thesis, 64 pages with a 2 and a half page bibliography. This work is six chapters is length. Includes Noun and Verb Phrases, as well as clause and sentence structure.

Phillips, Richard L. n.d. Central Mnong Linguistic Articles. Vietnam Data Microfiche Series VP14-16. s.l. s.n.

Referenced by Phillips (1963).

----. 1973. SIL Bibliography: A Mnong pedagogical grammar: The verb phrase and constructions with two or more verbs. Mon-Khmer Studies (MKS) 4.129–138.

Eastern Mnong.

----. 1973. Mnong vowel variations with initial stops. Mon-Khmer Studies (MKS) 4.119–127.

Central Mnong.

Smith, Kenneth D. 1969. Sedang affixation. Mon-Khmer Studies (MKS) 3.108–129.

----. 1974. Homonyms in Sedang Kinship Terminology. Typescript. s.l., ms.

Referenced by Smith (1979).

----. 1976. Sedang pronoun reference. Mon-Khmer Studies (MKS) 5.165–178.

----. 1976. North Bahnaric Numeral Systems. Linguistics 174.61–63.

Referenced by Smith (1979).

----. 1979. *Sedang* Grammar. Pacific Linguistics B-50. Canberra: Australian National University.

There is a twelve page bibliography included in this work. The Linguistic Institute Library at Payap University, Chiang Mai, Thailand, holds a corrected draft from 1975.

Thomas, David. n.d. Chrau Sentences. ms.

Referenced by Thomas, Dorothy (1980).

----. 1971. *Chrau Grammar*. Oceanic Linguistics Special Publication 7. Honolulu: University of Hawaii Press.
http://www.jstor.org/discover/10.2307/20019129?uid=3739136&uid=2&uid=4&sid=21102716329831 accessed 8 October 2013.
Availible online for free, downloadable for $14.

----. 1976. A Chrau noun phrase battery. Mon-Khmer Studies (MKS) 5.135–138.

----. 1976. South Bahnaric and other numeral systems. *Linguistics* 174.65–80.

Copy held at the Linguistics Institute library at Payap University, Chiang Mai, Thailand,.

----. 1986. Some Proto-South-Bahnaric clause grammar. Mon-Khmer Studies (MKS) 15.111–124. http://rilca.mahidol.ac.th/e-resources/documents/05-MonKhmer/15/thomas1986proto.pdf accessed 8 October 2013.

This work mentions the following languages: Chrau, Koho, Central Mnong, Eastern Mnong, and Stieng.

Thomas, David, and Dorothy Thomas. 1959. Chrau Sentence Components (Chrau Sentences Revisited). s.l., ms.

This paper is based on the work of Thomas and Thomas between 1959-1975.

Thomas, Dorothy. 1969. Chrau affixes. *MKS* 3.90–107.

----. 1980. The Paragraph Level in Chrau. Mon-Khmer Studies (MKS) 8.187–220.

Vogel, Sylvain. 2001. Esquisse d'une grammaire de la langue Phnong [Outline of a grammar of the language Bunong]. Unpublished manuscript. s.l., ms.

Referenced by Bequette, Rebecca Lee Elaine (2008). Central Bunong.

Vogel, Sylvain, and Jean-Michel Filippi. 2002. Eléments de langue phnong. [Elements of Bunong language] Phnom Penh: European Commission - Cambodia Malaria Control Project.

Vocabulary and grammar for Central Bunong.

4.3 Phonetics and Phonology

Banker, John E. 1961. Bahnar phonology. Ms. Saigon, ms.

Referenced by Gregerson, Kenneth J., and Kenneth D. Smith. 1973.

Bequette, Rebecca. 2006. Phonology of Bunong. Unpublished manuscript. s.l., ms.

Blood, Henry. 1976. The phonemes of Uon Njuñ Mong Rơlơm. Mon-Khmer Studies (MKS) 5.4–23.

Mong Rơlơm is now known in the Ethnologue as a dialect of Eastern Mnong.

Bon, Noëllie. 2010. Phonology of the Stieng Language: a Rime Study 9. Journal of the Southeast Asian Linguistics Society 3.22–48. http://chl.anu.edu.au/pacling/series/SEALS-PDFs/JSEALS-3-2.pdf#page=27 accessed 16 May 2013.

Butler, Becky. n.d. Registrogenesis in Bunong: The Interaction of Phonetics and Phonology. JSEALS.

Forthcoming in JSEALS.

Cooper, James, and Cooper, Nancy. 1966. Halăng phonemes. Mon-Khmer Studies (MKS) 2.87–98.

Crowley, James Dale. 2000. Tampuan phonology. Mon-Khmer Studies (MKS) 30.1–21.

Davis, John J. 1968. Nyaheun Phonemes. Overseas Missionary Fellowship (Mission Evangelique Paxse, Laos).

This work is located at Payap University in the Linguistics Institute library, Chiang Mai, Thailand.

Ефимов А. Ю. [Efimov, A.Yu.]. 1987. Историческая Фонология Южноъахнарических Яэыков [Historical Phonology of South Bahnaric Languages]. Москва [Moscow]: Академия Наук СССР Орлена Трулового Красного Энамени Институт Востоковеления [Academy of Sciences of the SSR Oriental Institute].

This work is written in Russian, and only a few sections of the front page had been translated into English. This work contains a four page bibliography in which Efimov cites several other Russian works. Can be located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand

Filippi, Jean-Michel. 2000. Le parler phnong de Mondulkiri. Unpublished. 495.954 fil. Copy held at Summer Institute of Linguistics Library Chiang Mai.

This work is written in French and covers phonetics and phonology in the Central Mnong language.

Gradin, Dwight. 1966. Consonantal tone in Jeh phonemics. Mon-Khmer Studies (MKS) 2.41–53.

Gregerson, Kenneth J., and Kenneth D. Smith. 1973. The Development of To'drah Register. Mon-Khmer Studies (MKS) 4.143–184.

Phonological information on the Todrah language.

---- 1976. Tongue-root and register in Mon-Khmer. *Austroasiatic Studies Part 1*, ed. by Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, 323–369. Oceanic Linguistics Special Publication 13. Honolulu: The University Press of Hawaii.

This work references the following languages: Hre, Jeh, Sedang, Halang, Rengao, and Central Mnong.

Haupers, Ralph. 1969. Stieng phonemes. Mon-Khmer Studies (MKS) 3.131–137.

Hoan, Nguyễn Văn. 1974. Affixation in Kơho. Saigon: Faculty of Letters, University of Saigon m.a. thesis.

Jacq, Pascale. 2001. A Description of Jruq (Loven): a Mon-Khmer language of the Lao PDR. Canberra: Australian National University master's thesis. https://digitalcollections.anu.edu.au/handle/1885/49408 accessed 1 October 2013.

This is Pascale Jacq's thesis for a Master of Philosophy from Australian National University. This thesis is 562 pages long, including a twelve-page bibliography.

----. 2002. Phonetic realisations of /ʔC/ and /HC/ word initial sequences in Jruq (Loven). Mon-Khmer Studies (MKS) 32.25–53. http://sealang.net/sala/ accessed 16 May 2013.

Jruq is listed as an alternate name for the Laven language in the Ethnologue.

Keller, Charles E. 2001. Brao-Krung phonology. Mon-Khmer Studies (MKS) 31.1–13.

Keller, Charles, Jacqueline Jordi, Kenneth Gregerson, and Ian G. Baird. 2008. Brao dialects: lexical and phonological variations. Revue de l'Institut de la Langue Nationale de l'Academie Royale du Cambodge.87–152.

Le, Duong Tan. 2003. A phonological comparison of Koho and Maa varieties. Chiang mai: Payap University Master of Arts in linguistics. *http://dcms.thailis.or.th/dcms/browse.php?option=show&browse_type=title&titleid=257043&display=list_subject&q=M* accessed 16 May 2013.

Master of Arts thesis, 228 pages. This work contains a 12 page bibliography.

Maier, Jacqueline G. 1969. Cua phonemes. Mon-Khmer Studies (MKS) 3.9–19.

Manley, T.M. 1976. Pharyngeal Expansion: Its Use in Sre Vowels and Its Place in Phonological Theory. *Austroasiatic Studies Part 2*, ed. by Philip N. Jenner, Laurence C. Thompson, and Stanley Starosta, 833–841. Oceanic Linguistics Special Publication 13. Honolulu: The University Press of Hawaii.

Nguyễn Thị Ánh Miểu.{XE "Nguyễn Thị Ánh Miểu"} 1985. Hiện Tượng Láy Trong Tiếng M'nông [Reduplicative Forms in Mnong]. Ho Chi Minh: Trường Đại Học Tổng Hợp Thành Phồ Ho Chí Minh [The City University of Ho Chi Minh] graduate essay.

This work is about the Southern Mnong language.

Phillips, Richard L. 1971. Phonological Reconstruction of Proto-Bahnaric. s.l., ms.

Referenced by Smith (1979).

----. 1973. Vowel distribution in Hrê. Mon-Khmer Studies (MKS) 4.63–68.

Sidwell, Paul J. 2000. Proto South Bahnaric: a reconstruction of a Mon-Khmer language of Indo-China. Pacific Linguistics 501. Canberra: Pacific Linguistics.

----. 2010. Cua (Kor) historical phonology and classification. Mon-Khmer Studies (MKS) 39.105–122.

In-depth study about Cua. In his research Sidwell is able to draw some conclusions on how Cua fits into the Bahnaric family. Includes a one page bibliography. Combining Sidwell (2010) and Smith (1981) gives a good understanding of how Cua is classified within Bahnaric.

Smalley, William A. 1954. Srê phonemes and syllables. JAOS 74.217–222.

Referenced by Gregerson, Kenneth J., and Kenneth D. Smith. (1973).

Smith, Kenneth D. 1967. A Phonological Reconstruction of Proto Central North Bahnaric. University of North Dakota SIL 1967 University of North Dakota SIL Workpapers. s.l.: s.n.

Referenced by Jacq and Sidwell (2000).

----. 1968. Laryngealization and de-laryngealization in Sedang phonemics. Linguistics 38.52–69.

Referenced by Wikipedia accessed May 16 th 2013.

----. 1969. The Phonology of Sedang Personal Names. Anthropological Linguistics 11.187–198.

----. 1969. Sedang Ethnodialects. Anthropological Linguistics 11.143–147.

----. 1972. A phonological reconstruction of Proto-North-Bahnaric. Language Data, Asian-Pacific Series 2. Santa Ana, CA: Summer Institute of Linguistics.

----. 1973. Denasolaryngealization in Sedang. Mon-Khmer Studies (MKS) 4.53–62.

----. 1973. More on Sedang Ethnodialects. Mon-Khmer Studies (MKS) 4.43–51.

In this work Smith compares phonological features of several Sedang ethnodialects.

----. 1973. Eastern North Bahnaric: Cua and Kơtua. Mon-Khmer Studies (MKS) 4.113–118.

----. 1973. Proto-North Bahnaric, 18. s.l.: s.n.

Phonological Reconstruction, presentation for the First International Conference on Austroasiatic Linguistics.

----. 1975. Phonology and Syntax of Sedang, A Vietnam Mon-Khmer Language. University of Pennsylvania doctoral thesis.

Includes a thirteen page bibliography, and an additional page and a half of vernacular publications.

----. 1975. The Velar Animal Prefix Relic in Vietnam Languages. Ms. s.l., ms.

Thomas, David. 1962. Remarques sur la phonologie du Chrau. BSLP 57.175–191.

Referenced by Blood, Henry (1976)

----. 1968. Some Mon-Khmer Vowel System, 10. Christ Church, N.Z.: s.n.

Conference paper referenced by Smith (1973) in MKS 4.

Thomas, Dorothy. 1966. Chrau intonation. Mon-Khmer Studies (MKS) 2.1–13.

Watson, R. 1969. A note on Ta-oi, Nge', and Nyaheun. Mon-Khmer Studies (MKS) 3.130–130.

Vowels in Nyaheun.

4.4 Dictionaries and Word lists

Unknown Author. n.d. Vietnam Word List. Mimeographed. s.l., ms.

This work contains 30 different language word lists, each from Southeast Asia. This work was produced by SIL's Vietnam Branch. Each word list is six pages in length. Referenced by Smith, Kenneth D. (1975).

Unknown Author. n.d. Bahnar Pronunciation Grammar Notes Vocabulary. Kontum: s.n.

This work is 57 pages in length and is mostly a vocabulary.

Unknown Author. 1959. *Bahnar*. Huẽ: Học Tiéng Thượng.

Vietnamese and Bahnar phrase book. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1959. *Ko'ho*. Huẽ: s.n.

Vietnamese and Ko'ho phrase book. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1960. Hrê. Học Tiếng Thượng [study of minority languages]. Huẽ: s.n.

Vietnamese and Hrê phrase book and vocabulary.

Unknown Author. 1960. Sédang. Học Tiếng Thượng [study of minority languages]. s.l.: s.n.

Vietnamese and Sédang phrase book.

Unknown Author. 1966. Comparative 100 word list of English, Jeh, Loven, Brou, Nyaheun, Chrau, Oi, Loveh. s.n. Folder 495.956. Copy held at Summer Institute of Linguistics Library Chiang Mai.

Referenced by Jacq and Sidwell (2000). Oi is an alternate name for the Oy language of Laos.

Unknown Author. 1983. TỪ' ĐIÊ'N VIỆT - KƠHO. Ho Chi Minh: Sơ Văn Hơa và Thông Tin [Department of Culture and Information].

185 page dictionary written in Vietnamese and Koho.

Unknown Author. 1986. Từ Điển Stiêng - Việt. [Ho Chi Minh: s.n.]

Stiêng and Vietnamese Dictionary. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Azémar, H. 1886. Dictionnaire Stieng: Recueil De 2,500 Mots Fait A Bro'lam En 1865 [Stieng Dictionary: 2,500 words of the Brolam area in 1865]. Excursions et Reconnaissances 12.99–146, 251–344. http://www.odsas.net/scan_sets.php?set_id=68&doc=16414&step=0 accessed 13 June 2013.

Referenced by Owen (2002) and Shorto (2006).

Banker, John, and Betty Banker. 1961. *Bahnar Dictionary*. s.l.: Summer Institute of Linguistics.

Bahnar, Vietnamese, and English dictionary. About one hundred pages in length.

Banker, John, Elizabeth Banker, and Mơ. 1979. Bahnar dictionary: Plei BongMang Yang dialect Ngữ-Vựng. Huntington Beach, CA: Summer Institute of Linguistics. http://archive.org/details/rosettaproject_bdq_ortho-1 accessed 16 May 2013.

'Bi, Điêu, and Ralph Haupers. n.d. Nói tiếng Sơđiêng = Stieng phrase book (Sơđiêng - Việt - Anh). Saigon: Summer Institute of Linguistics.

Referenced by Haupers (1991).

Blood, Evangeline. 1976. *Eastern Mnong Vocabulary*. Vietnam Data Microfiche Series VD15-15. Huntington Beach, CA: Summer Institute of Linguistics.

This work is a dictionary and thesaurus. This work is written in Eastern Mnong, Vietnamese and English. Copy held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Blood, Henry, and Evangeline Blood. 1966. Mnong Dictionary. s.l.: Summer Institute of Linguistics Vietnam Branch.

Blood, Henry, Evangeline Blood, and Y Tang Hmok. 1976. Mnong Ralăm Dictionary and Thesaurus. Huntington Beach, CA: Summer Institute of Linguistics.

Eastern Mnong dictionary, 422 pages.

----. 1976. Mnong Ralăm Thesaurus. s.l.: s.n.

This is a lengthy work, probably similar in nature to the above entry.

Bochet, Gilbert, and Jacques Dournes. 1953. Lexique Polyglotte (Francais, Kơho, Rơglai, Vietnamien). Saigon, France, Asia: s.n.

Referenced by Blood (1966).

Boulbet, J. 1972. Dialogue Lyrique Des Cau Maa' (Tam pöt maa') [Dialogue Lyrique from Cau Maa' (Tam Can Maa')]. Vol. 85. Publications De L'école Française D'extrême-Orient. Paris: Ecole Française D'extrême-Orient.

This work is in French, and contains a 17-page glossary. Boulbet also has written other works on the Maa language see Huffman's 1986 bibliography.

Bul, Ha. 1976. *Koho Vocabulary*. Vietnam Data Microfiche VD12-72. Huntington Beach, CA: Summer Institute of Linguistics.

This work is a Koho dictionary/thesaurus written in Kơho, Vietnamese and English. Copy held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Burton, Eva, and Jacqueline G. Maier. 1966. Cua Dictionary. Quảng Ngãi Province, ms.

Copy held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand. Referenced by Sidwell in MKS 39.

----. 1971. Vietnam word list (revised): Cua. *http://www.sil.org/acpub/repository/51389.pdf* accessed 16 May 2013.

Cabaton, Antoine. 1905. Dix dialectes Indochinois recueillis par Prosper Odend'hal: etude linguistique par Antoine Cabaton [Ten Indochinese dialects collected by Prosper Odend'hal: linguistic study by Antoine Cabaton]. Journal Asiatique (JA) 5.265–344.

Referenced by Smith (1979), and Katuic Bibliography (2009), accessed 16 May 2013.

Cheon, Jean, and A. Mougeot. 1890. Essai de dictionnaire de la langue Chrau. [A Chrau Test Language Dictionary]. Bulletin de la Société des Études Indochinoises 2.5–106.

Referenced by Sidwell (2000).

Cohen, Patrick D. 1979. Chù chih dŏ tơtayh Jeh = Ngữ-vưng Jeh = Jeh vocabulary. Manila: Summer Institute of Linguistics.

Referenced by Haupers (1991).

Cooper, James, and Nancy Cooper. 1972. Halang. Wordlist. s.l., ms.

Referenced by Prachakij-karacak (1995).

----. 1974. Halang Vocabulary. s.l., ms. 495.964 coo. Copy held at Payap University, Linguistic Institute, Library.

Crowley, James Dale, Vay Tieng, and Wain Churk. 2007. Tampuan Khmer English dictionary: with English Khmer Tampuan glossary. Cambodia: EMU International & National Language Institute of the Royal Academy of Cambodia.

Referenced by: http://en.wikipedia.org/wiki/Tampuan_people accessed May 16 th 2013.

Dourisboure, P.X. 1889. *Dictionnaire bahnar-français*. Hong Kong: Societe des mission et rangers.

Bahnar and French dictionary. Copy at the National Library of Thailand.

Dournes, Jacques. 1950. Dictionnaire Srê (Köhö)-Français. Saigon: des Missions Etrangères de Paris.

Referenced by Shorto (2006).

Drouin, S, and K'Naĭ. 1962. *Dictionaire Français-Montagnard* [Dictionary French-Montagnard]. 4 vols. Dalat: s.n.

Four volume Koho dictionary written in French. The total four volumes contain well over a thousand pages. Sidwell references this in Proto South Bahnaric: "mimeographed: Was on film 2359, Cornell University. Watson Collection".

Ferlus, Michel. 1970. Lexique Laven. Manuscript. s.l., ms.

----. 1998. Nhaheun-French-English Lexicon. (Ed.) Pascale Jacq and Paul Sidwell. München: Lincom Europa.

Gradin, Dwight, and Pat Cohen. 1964. Jeh Dictionary. Kontum, ms. Copy held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

----. 1969. Jeh Rhyming Dictionary. s.l.: s.n.

Over five hundred pages.

Gerber, Th. 1937. *Lexique Franc-Stieng*. Saigon: impr. du Théâtre.

Gregerson, Kenneth, Marilyn Gregerson, and André Ir. 1969. Rơngao Dictionary. Kontum: Summer Institute of Linguistics.

Gregerson, Kenneth J., and Marilyn Gregerson. 1977. Ngữ-vựng Rơngao = Rengao vocabulary. Tủ sách ngôn-ngữ dân-tộc thiểu-số Việt-Nam 19. Manila: Summer Institute of Linguistics.

Gregerson, Kenneth, and Timothy Friberg. 1973. Tampuan. Wordlist. s.l., ms.

Guilleminet, Paul, and Alberty, R.P. Jules. 1959. *Dictionnire Bahnar*-Français. 2 vols. Publications de l'Ecôle Française d'Extrême-Orient 40. Paris: Publications de l'Ecôle Française d'Extrême-Orient.

Bahnar and French dictionary. Volume 1: A-K Volume 2: L-Z.

Haupers, Lorraine, and Ralph Haupers. 1991. *Stieng* - English *Dictionary*. Manila: Summer Institute of Linguistics Thailand Group.

263 pages, contains a brief introduction to the Stieng alphabet. Draft copy of the Stieng Dictionary dated 1980 and still another earlier version dated 1968 are held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Irwin, George. 2009. Vietnam word list (revised): Kơho (Sre). Wordlist Draft. s.l., ms. http://www.sil.org/resources/archives/35754 accessed 16 May 2013.

According to the SIL website this was created in 1970.

Jacq, Pascale, and Paul Sidwell. 1999. Loven (Jruq) consolidated lexicon. Vol. 23. München: LINCOM Europa.

Referenced by Jacq and Sidwell (2000).

----. 2000. A Comparative West Bahnaric Dictionary. Vol. 21. Languages of the World/Dictionaries. Muenchen: Lincom Europa.

Contains a four-and-a-half page bibliography.

Johnston, Beulah, Janice Saul, and David Thomas. 1978. Alak. Wordlist. s.l., ms.

Referenced by Prachakij-karacak (1995).

Keller, Charles. compiler 2012. Vietnam word list (revised): Brao. s.n.

Referenced by the SIL Bibliography. Contains 283 lexical items. Brao is an alternate name for the Lave language. The original wordlist was taken in 1976 by Keller in Ratanakiri Province in Cambodia. A copy is held at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Keller, Charles, and Sally Keller. 1977. Brao Language (Cambodia), Brao Rhymed Vocabulary. Dallas: Summer Institute of Linguistics.

Referenced by Jacq and Sidwell (2000).

Kommala, Chinda. 1978. Southeast Asia Word List (revised): Loveh. s.n. Folder 495.956. Copy held at Summer Institute of Linguistics Library Chiang mai.

Referenced by Jacq and Sidwell (2000). Loveh probably refers to the language called Lave in Laos. Laveh is an alternate name for Lave.

Leger, Daniel. 1974. Vocabulaire compare et recherche du vocabulaire dentaire Bahnar-Jolong [Comparative Vocabulary and Dental Vocabulary Research in Bahnar-Jolong]. ASEMI 5.123–132.

Luce, Gordon H. n.d. Comparative lexicon: Mi - Xong - Chong - Nañang - Sue - Tareng - Hin - Kontu - Alak - Churu - Kaseng - Ka - Laveh - Sedang - Halang - Kui - Samre - Por - KuyDek - Chrau - Phnong - Proons - Bahnar - Stieng - Boloven. Luce Collection, MS 6574-7, pp. 049a-b. National Library of Australia.

Referenced by SEALANG SALA accessed 16 May 2013.

Maier, Jacqueline, and Eva Burton. 1966. Cua Lessons and Dictionary. s.l.: s.n.

Maier, Jacqueline, and Van Cau Finh. 1976. Cua Vocabulary: Cua-Vietnamese-English Thesaurus (a computer print-out). Summer Institute of Linguistics.

Referenced by Jacq and Sidwell (2000).

Manson, Francis. 1854. The Talaing Language. Journal of American Oriental Society 4.277–288. http://www.jstor.org/stable/592280?seq=1 accessed 16 May 2013.

Matras-Troubetzkoy, Jacqueline. 1974. Elements De Vocabulaire Brou. ASEMI 5.137–140.

This work is written about the the Brao (Lave) language.

Matras, Jacqueline, and Marie Martin. 1972. Contribution A L'ethnobotanique Des Brou (Cambodge - Province de Ratanakiri). Journal D'Agriculture Tropicale Et De Botanique Appliquée XIX.1–97.

Bru botanical terms.

Miller, Carolyn. 1988. Southeast Asia Word List: Trieng. s.n. Copy held at Summer Institute of Linguistics Library Chiang Mai, Thailand.

Referenced by Jacq and Sidwell (2000).

Phan-Thanh-Văn, L.M. 1958. Học Tiếng Việt Việtnamese – Bahnar Diglot. Kontum: s.n.

Vietnamese - Bahnar gloss, 63 pages.

Phillips, Richard L. 1961. Hre Dictionary. Banmêthuột, ms.

Copy held at Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

----. 1962. Mnong Dictionary. Vietnam Data Microfiche Series VD14-73. s.l.: s.n.

Written about Central Mnong (Bunâr).

Phillips, Richard L., and Y Kem Kpơr. 1966. *Mnong Thesaurus*. Vietnam Data Microfiche Series VD14-74. Banmêthuột: Hội Tin Lành.

Referenced by Phillips (1963). Central Mnong (Bunâr).

Prachakij-karacak, Phraya. 1995. *Some Languages of Siam*. (Trans.) David Thomas and Sophana Srichampa. Bangkok: Institute of Language and Culture for Rural Development Mahidol University.

Written in Thai and English. Page 145 contains a wordlist comparison of five non-west Bahnaric languages (Alak, Alak 2, Tampuan, Salang [Halang], Phanong [Mnong, Central). (788 words are compared).On page 198 five West Bahnaric languages are compared (Yaheun [Naheun], Boriwen [Laven], Su' [Sou], Ta-oy, Ceng [Jeng]). (Ta-oy is now classified as Katuic). Five more West Bahnaric languages are compared on page 239: Sork [Sok], Sapuan, Rawe [Lave], Palau [Lave], Kraseng [Kasseng].

Ratanakul, Suriya, David Thomas, and Sophana Srichampa. 1991. Some Late 19th Century West Bahnaric Word Lists: A Preliminary Report. Mahidol University: s.n.

This work mentions the following languages: Sok, Nyaheun, Loven, Sou, Jeng, Sapuan, and Kasseng.

Phạm - Xuân – Tín, Rev.. 1962. Bahnar Gơlar Ở Pleiku Bahnar Gơlar Vocabulary. s.l.: Summer Institute of Linguistics University of North Dakota.

Sang Lục, Thô, and David D. Thomas. 1966. Ngũ-vụnh Chrau = Chrau vocabulary = Chrau - Việt - English. Saigon.

Basic vocabulary in English, Vietnamese, and Chrau, based on topics. 128 pages. Located at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Shorto, Harry. 2006. A Mon-Khmer Comparative Dictionary. (Ed.) Paul Sidwell, Doug Cooper, and Christian Bauer. Canberra: Pacific Linguistics.

Includes a seven page bibliography.

Smith, Kenneth D. 1967. *Ngữ-vựng Sedang = Sedang vocabulary*. Saigon: Bô Giáo-duc (Department of Education).

Sedang, Vietnamese, and English vocabulary, 128 pages.

Smith, Kenneth D. 2000. Sedang Dictionary with English, Vietnamese, and French Glossaries. Special Volume No. 1. MKS. Dallas: Summer Institute of Linguistics.

Taupin, J. 1888. Vocabulaire Braou. Bulletin de la Société des Études Indochinoises 12.59–64.

Referenced by Jacq and Sidwell (2000). This work is in French.

Thomas, David D. n.d. *Chrau Concordance*. Vol. 1. 5 vols. s.l.: s.n.

These 5 volumes are located at the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Thomas, David. 1978. Cheng. Wordlist. s.l., ms.

Cheng is an alternate name for the Jeng language group. Referenced by Prachakij-karacak (1995).

----. 1969. Vietnam minority languages. Mimeo. Saigon, ms.

Referenced by Gregerson, Kenneth J., and Kenneth D. Smith (1973).

Thomas, David, and Dorothy Thomas. 1974. Chrau Dictionary. s.l.: Summer Institute of Linguistics Vietnam Branch.

A very lengthy, work hundreds of pages.

Thomas, David, and Richard Johnston. 1978. Southeast Asia Word List: Nyaheun. s.n. Copy held at Summer Institute of Linguistics Library, Chiang Mai, Thailand..

Referenced by Jacq and Sidwell (2000).

Thomas, Dorothy M. 1978. Southeast Asia Word List (revised): Oi. Copy held at Summer Institute of Linguistics Library Chiang mai.

Referenced by Jacq and Sidwell (2000). Oi is an alternate name for the Oy language in Laos.

Thomas, Dorothy M., and D. Andrianoff. 1978. Southeast Asia Word List (revised): Lawen. Copy held at Summer Institute of Linguistics Library, Chiang Mai, Thailand.

Referenced by Sidwell (2000). The Ethnologue records no such language as Lawen, this might be a misspelling of Laven.

Trebilco, Oliver, compiler. 1971. Vietnam word list (revised): Hrê. s.l., ms.

Referenced by SIL Bibliography accessed May 16th 2013.

Trebilco, Oliver, Joyce Trebilco, and Đinh Ngiah. 1961. Từ Điển Việt - Hrê - Anh [Dictionary English - Hrê - Vietnamese]. Vol. 1. 3 vols. s.l.: s.n.

Contains 3762 entries.

----. 1961. Từ Điển Anh - Hrê - Việt [Dictionary English - Hrê - Vietnamese]. Vol. 2. 3 vols. s.l.: s.n.

Contains 3711 entries.

----. 1961. Từ Điển Anh - Hrê - Việt [Dictionary English - Hrê - Vietnamese]. Vol. 3. 3 vols. s.l.: s.n.

Contains 3899 entries.

4.5 Sociolinguistics (includes language planning, survey, ethnolinguistics.)

Unknown Author. 1907-1908. Les Peuples Mon-Khmers. BEFEO 7.213–263; 8.1-35.

Referenced by Blood (1966).

Unknown Author. 1973. Ethnic Minorities of South Vietnam. MKS 4.ix.

This is a 1971 language map of South Vietnam produced by SIL.

Unknown Author. 1998. The Peoples of Vietnam. Chiang Mai: Asian Minorities Outreach.

Profiles of the Bahnaric languages spoken in Vietnam.

Unknown Author. 1999. Faces of the Unreached in Laos Southeast Asia's Forgotten Nation. Chiang Mai: Asian Minorities Outreach.

Profiles of the Bahnaric languages spoken in Laos. Contains a two-and-a-half page bibliography. Pages 139 and 140 are particularly interesting since they note other possible ethnic and or language groups in Laos.

Unknown Author. 2007. The Peoples of Cambodia. Cambodia: Cambodia Research Network - CRN.

Profiles on Bahnaric language groups. Contains a two and a half page bibliography.

Unknown Author. 2013. Montagnard (Vietnam). Wikipedia. http://en.wikipedia.org/wiki/Montagnard_(Vietnam) accessed 16 May 2013.

Baudesson, Henry. 1997. Indo-China and its Primitive People. Bangkok: White Lotus Press.

Includes a three page bibliography.

Bequette, Rebecca Lee Elaine. 2008. Participant Reference, Deixis, and Anaphora in Bunong Narrative Discourse. Dallas: Graduate Institute of Applied Linguistics Master of Arts.

Includes a three page bibliography.

Bequette, Todd. 2005. Report on a visit to Kaoh Nheaek district. Unpublished manuscript. s.l., ms.

Kaoh Nheaek district, located in Mondulkiri province in Cambodia, is a Bahnaric speaking area.

Bondet de la Bernardie, J. 1949. Le dialecte des kha Boloven. *BSEI* 24.57–78.

Bradley, David. 2007. Language Endangerment in China and Mainland Southeast Asia. Language Diversity Endangered, ed. by Matthias Brenzinger, 181:278–302. Trends in linguistics. Studies and monographs. Berlin: Mouton de Gruyter.

Chazée, Laurent. 1995. Atlas des ethnies et des sous-ethnies du Laos. Bangkok: Privately printed.

Referenced by Schliesinger (2003). This work references several ethnic groups such as Kate, Solas, and Tahang. These ethnic groups are Mon-Khmer, but it's uncertain to their relationship with Bahnaric. The work Faces of the Unreached in Laos mentions that the group Tahang could be related to Talieng.

----. 2002. The Peoples of Laos: Rural and Ethnic Diversities. Bangkok: White Lotus Co. Ltd.

Referenced by Schliesinger (2003). Chazée references a Bahnaric group on page p.95 that is called Lawi (Saveung). Lawi and Saveung do not appear in the Ethnologue.

Coedès, George. 1962. Les Peuples de la Péninsule Indochinoise: Histoire, Civilizations. [Peoples of the Indochinese Peninsula: History, Civilizations.]. Paris: Dunod.

Referenced by Smith (1979).

Cooper, James S. 1973. An ethnography of Halăng rhymes. Mon-Khmer Studies (MKS) 4.33–41.

Dang Nghiem Van, Chu Thai Son, and Luu Hung. 1993. The Ethnic Minorities in Vietnam. Hanoi: The Gioi Publishers.

Referenced by Schliesinger (2003).

Davis, John Personal letter 1966. ----. Personal letter. 495.95 W. Bahnaric (Brao), folder. Linguistic Institute Library Payap University.

This is a letter from John J. Davis to David Thomas he briefly mentions the Nyaheun language.

Davis, J., and R. Smith. 1973. Tribes of Southern Laos. MKS 4.xi.

Language map of southern Laos dated 1970.

Devereux, George. 1947. The Potential Contributions of the Moi to the Cultural Landscape of Indochina. Far Eastern Quarterly 6.390–395.

Referenced by Smith (1979). The Moi are also known as the Hre language group.

Diffloth, Gérard. 1993. The indigenous languages of Cambodia and the elections, Report for the Electoral Component of UNTAC. Phnom Penh.

Referenced by ICC (2006).

----. 1992. Indigenous languages of Cambodia, 4. Berkeley: s.n.

This work contains three maps of Cambodia.

Dournes, Jacques. 1974. Une documentation sur des parlers Koho [A Documentation of Koho dialect]. ASEMI 5.161–170.

Referenced by MKS 5.

Fraisse, Andre. 1951. Les villages du plateau des Bolovens. BSEI 26.52–72.

Referenced by Lebar (1964). Fraisse was able to gather some information from the Oy language group in Laos.

Gregerson, Marilyn, and Dorothy Thomas (eds.) 1980. Notes from Indochina on Ethnic Minority Cultures. SIL Museum of Anthropology.

Includes information about the Stieng, Sedang, Chrau, Jeh, and Mnong Lăm language groups. There are also a number of small bibliographical sections throughout the book. Contains an ethnogeography on Chrau.

Guilleminet, Paul. Languages spéciaux utilisés dan la tribu Bahnar du Kontum (Sud Viet-Nam-Indochine). *BEFEO* L.

Harmand, F.J. 1997. Laos and the hilltribes of Indochina: Journeys to the Boloven Plateau, from Bassac to Hue through Laos, and to the origins of the Thai. Bangkok: White Lotus Press.

Referenced by Schliesinger (2003).

Hayashi, Yukio. 1995. Notes on the Inter-Ethnic Relations in History: With Special Reference to Mon-Khmer Peoples in Southern Laos. Sakon Nakhon.

Referenced by a Korean bibliography online and by Schliesinger (2003).

Hickey, Gerald Cannon. 1964. The Major Ethnic Groups of the South Vietnamese Highlands. Santa Monica: The Rand Corporation.

Referenced by Schliesinger (1997).

Hoffet, J. 1933. "Les Mois de la chaine annamitique," Terre, air, mer [The Months of the Annamite Chain: Land, Air, and Sea]. La géographie 59.1–43.

Referenced by Lebar (1964). Hoffet was able to gather some information from the Oy language group in Laos.

International Cooperation Cambodia (ICC). 2006. Bahnaric Language Cluster Survey. Survey. Linguistic Institute Library Payap University.

This work is about a language survey conducted in Mondulkiri and Kratie Providences of Cambodia.

Lafont, Pierre-Bernard. 1962. Personal Notes. (Specific notes from previous field experiences in Laos and Vietnam.). s.l., ms.

Referenced by Lebar (1964). Lafont was able to get field information on the Oy of Laos, (see page 145 of Lebar 1964).

Lebar, Frank K., Gerald C. Hickey, and John K. Musgrave. 1964. Ethnic Groups of Mainland Southeast Asia. New Haven: Human Relations Area Files, Inc.

288 pages. Contains articles about different Bahnaric groups. Twelve-and-a-half page bibliography.

Lieurade, Médecin-Colonel. 1951. Généralités sur les populations montagnardes du Sud indochinois. *Bulletin de la Société des Études Indochinoises* 26.

This work contains a language map of southern Vietnam, Cambodia, and southern Laos.

Magaspag, Chitse E. 2009. Language use and attitudes of Kachok speakers: towards an assessment of the Kachok language vitality. Manila: Philippine Normal University m.a. thesis. http://www.sil.org/resources/publications/entry/50817 accessed 16 May 2013.

Kachok is another name for the Kaco' language.

McKinstry, John. 1960. *Bibliography of Laos and Ethnically Related Areas*. (Ed.) Joel M. Halpern. Laos Paper 22. Amherst: University of Massecusetts.

The references in this bibliography are mostly in French, and cover a wide range of topics related to Loas. While most references are not related to Bahanaric, there are a few older referneces pertain to the people groups in Loas.

Nuttle, David A. 1961. The Montagnards of South Vietnam Highlands. Saigon: USIS.

Referenced by Schliesinger (1997).

Owen II, James Edmond. 2002. A Discourse Analysis of Two Stieng Narratives. Arlington: The University of Texas at Arlington.

Master of Arts thesis. This work contains a three page bibliography. First discourse analysis on two Stieng narratives.

Peters, K.M. 1964. Tribes of South Vietnam. Language Map. Saigon: SIL.

Can be found in "Papers on Four Vietnamese Languages" edited by Thomas 1966, (see page ii).

Schliesinger, Joachim. 1997 and 1998. Hill Tribes of Vietnam. 2 vols. Bangkok: White Lotus Co. Ltd.

Contains short sections on various ethnic minority groups in Vietnam. There is an extensive bibliography in volume one.

----. 2003. Ethnic Groups of Laos - Profile of Austro-Asiatic-Speaking Peoples. Vol. 2. Bangkok: White Lotus Press.

Contains people profiles on different Austro-Asiatic people groups. This work includes many Bahnaric groups.

Schrock Joann L. (et.al.) 1966. Minority Groups in the Republic of Vietnam. Ethnographic Study Series Pamphlet No. 550-105. Washington D.C.: Department of the Army Headquarters.

An extensive work (over one thousand pages) that gives in-depth detail of the Bahnar, Cua, Halang, Hre, Jeh, Koho, Maa, Rengao, and Sedang peoples, as well as the Stieng and M'nong ethnic groups.

Smith, Kenneth D. 1974. Sociolinguistics and the Bahnar Pronoun System. Typescript. s.l., ms.

Thomas, David and Dorothy. 1960. Dave and Dottie Thomas's Method for Collecting and Arranging Their Language Material in Chrau. Work Papers of the Summer Institute of Linguistics University of North Dakota, 4:109–112. Work Papers of the Summer Institute of Linguistics, University of North Dakota. Grand Forks: SIL/UND.

The authors describe their methodology for language data collection.

Thomas, Dorothy. 1978. The Discourse Level in Chrau. Mon-Khmer Studies (MKS) 7.233–295.

----. 1966. Chrau Zoology: an ethnolinguistic study. Te Reo 7.1–14.

Referenced by Ronald L. Smith in MKS 4.

Vogel, Sylvain. 2006. Introduction à la langue et aux dits traditionnels des Phnong de Mondulikri [An Introduction to the Language and Traditional Sayings of Bunong of Mondulikri]. Phnom Penh: Editions Funan.

Referenced by Bequette, Rebecca Lee Elaine (2008). This is written about Central Bunong.

U.S. Information Service. 1962. Montagnards of South Vietnam Highlands. Saigon: U.S.I.S.

Referenced by Schliesinger (2003).

Wall, Barbara. 1975. Les Nya Hön: etude ethnographique d'une population du plateau des Bolovens (sud-Laos). Paris: Vithagna.

Referenced by Sidwell (2000). Ethnographic study of the peoples of the Boloven Plateau in southern Laos. Specifically the Nyaheun language.

4.6 Anthropology

Unknown Author. 2013. 54 Ethnic Groups of Vietnam. http://www.offroadvietnam.com/eng/13-45.php#32.

Azémar, H. 1886. *The Stiengs of Brolam*. Lai Thieu: s.n.

This work contains cultural information about many aspects of Stieng life. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Baird, Ian George. 2008. Various forms of colonialism : the social and spatial reorganisation of the Brao in southern Laos and northeastern Cambodia. Vancouver: University of British Columbia doctor of philosophy - PhD. https://circle.ubc.ca/handle/2429/1337 accessed 16 May 2013.

Embree, John F., and Lillian Ota Dotson. 1950. Bibliography of the Peoples and Cultures of Mainland Southeast Asia. New Haven: Yale University, Southeast Asia Studies.

Referenced by Schliesinger (1997).

Gerber, T. 1951. Coutumier Stieng [Stieng Customs]. BEFEO 45.228–269.

Referenced by Gregerson, Marilyn, and Dorothy Thomas (eds.) (1980).

Guilleminet, Paul. 1952. Coutumier de la tribu Bahnar, des Sedang et des Jarai de la province de Kontum [Customs of the Bahnar, Sedang, and Jarai tribes of the Kontum province]. Paris: E. de Boccard.

Referenced by Smith (1979).

Hickey, Gerald Cannon. 1967 The Highland People of South Vietnam: Social and Economic Development. Santa Monica: Advanced Research Projects Agency.

Referenced by Schliesinger (1997).

----. 1967. Some Aspects of Hill Tribe Life in South Vietnam. Southeast Asian Tribes, Minorities, and Nations, ed. by Peter Kunstadter, 2:745–770. Princeton: Princeton University Press.

----. 1982. Sons of the Mountains: Ethnohistory of the Vietnamese Central Highlands to 1954. New Haven: Yale University Press.

----. 1982. Free in the Forest: Ethnohistory of the Vietnamese Central Highlands, 1954-1976. Binghamton, NY: Vail-Ballou Press.

Mallow, Kreg P. 2002. Perceptions of social change among the Krung hilltribe of northeast Cambodia. Wheaton: Wheaton College Graduate School m.a. thesis.

Matras-Troubetzkoy, Jacqueline. 1974. L'essartage chez les Brou du Cambodge Organisation collective et autonomie familiale [The Clearing Among the Brou people of Cambodia: Collective Organization and Family Autonomy]. Études rurales.421–437.

----. 1980. Sacrifice et possession chez les Brou: une revanche des femmes [Sacrifice and Possession Among the Brou People: A Revenge by Women?]. ASEMI 11.415–429.

Mole, Robert L{ XE "Mole, Robert L." }. 1970. *The Montagnards of South Vietnam A Study of Nine Tribes*. Tokyo: Charles E. Tuttle Company.

Chapters 6-8 cover the following three languages Cua, Hre, and Jeh. Page 3 has a nice map of the Ethno-Linguistic groups of South Vietnam.

Nguyễn Duy Quý, and Nguyễn Duy Chiếm (eds.) 1998. *The Sedang of Vietnam*. National Centre For Social Sciences And Humanities of Vietnam.

This book contains many colorful pictures of Sedang life and culture.

Raulin, Henri P. 1946. L'Evolution des Stieng de la Delegation de Hon-quan [[The Evolution of Stieng from among the Delegation of Hon-quan]. BSEI 21.67–71.

Referenced by Gregerson, Marilyn, and Dorothy Thomas (eds.) (1980).

----. 1947. Les Technique de la Percussion et de la Production du Feu chez les Stieng.[The technique and percussion and production of fire in Stieng]. BSEI 22.111–121.

Referenced by Gregerson, Marilyn, and Dorothy Thomas (eds.) (1980).

Smith, Kenneth D. 1976. Sedang Animal Folk Taxonomy. MKS 5.179–194.

Soulié, Maurice. 1927. Merie ler, roi des Sedang, [Marie I, King of the Sedang] 1888-1890. Paris: Marpon.

This work is written in French. Marie I established himself as king of the Sedang people in 1888, (See Wikipedia Link accessed May 16th 2013). Referenced by Smith (1979).

Thông Tấn Xã Việt Nam Vietnam News Agency. 2006. *Việt Nam Hình Ảnh Cộng Dồng 54 Dân Tộc Vietnam Image of the Community of 54 Ethnic Groups*. Hanoi: The Vna Publishing House.

This book contains profiles on 54 ethnic groups in Vietnam which includes: Churu, Hre, Maa, Eastern Mnong, Central Mnong, Romam. Many colourful ethnic and cultural photos included.

Xu Man, and Từ Chi. 1986. Ca'c Dân Tộc Giarai_Bana [Art Culture of the Giarai (Jarai) and Bahnar people]. Kontum: Gialai and Kontum Culture and Information Office.

Different cloth patterns of the Jarai and Bahnar people.

4.7 Vernacular Publications and Christian Resources

The Linguistic Institute library at Payap University, Chiang Mai, carries nearly one hundred Bahnaric works, primarily Christian scriptures, or hymn books written in the various Bahnaric languages such as Halang, Jeh, Charu, Bahnar. Many of these works were not added to the vernacular publications section of this bibliography.

Unknown Author n.d. Calêu Hrê. Quảng Ngãi, Vietnam: s.n.

Hre hymnbook. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand where it is listed as Hre Hymns #1.

Unknown Author. n.d. Nau Brah Ndu Ngơi Nau Tâm Rnglăp Mhe. s.l.: United Bible Societies.

Central Bnong New Testament. Because of script issues, only Central Mnong on the Vietnamese side of the Vietnamese-Cambodian border could understand it.

Unknown Author. 1958. Sră Đơs-Chò Dalat.

This is a Koho hymnbook containing 107 hymns. It also contains church liturgy. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1960. Pơnuaĭ pô lơngì (Đạo Dức Chúa Trời). Hội – Thánh Tin – Lành, Vietnam: s.n.

This work is written in Chrau and contains songs, teachings of God, and Catechisms. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1966. Tech Hêêl Kasay Pâk Kool [Life of Christ]. trial edition. s.l.: s.n.

This work is in the Cua language; it covers selected New Testament Bible stories from the life of Christ. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1972. Mô-Se Păng Kuon Pơtauv. Saigon: s.n.

This work is written in Halang and is the story of Moses and the Princess. This work is held at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1976. Luka Tal 15 [Luke 15]. s.l.: s.n.

This is Luke 15 written in the Chrau language. Copy held at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1976. *Rài Hơdiŭ Jôsep*. Vietnam: s.n.

Old Testament story of Joseph in the Chrau language. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand. At the library there can also be found a 1974 copy of this story.

Unknown Author. 1978. Kơy Jêm Chin Thơ Ka Oh Ay Jo-Thayq Luka 15. Vietnam: s.n.

This work is the book of James and Luke 15 written in the Cua language. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1983. Kơy Mak Taboon Kadrôôk ----- Varauq Ku Oh Ay Saphok Dơop. Quảng Ngãi, Vietnam: s.n.

This is the Gospel of Mark and the book of Acts translated into Cua. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1983. Mar-Kô - Dơdruong Dăng Pơthô. Vietnam: s.n.

This work is Mark and the book of Acts written in the Halang language. This work is held at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1984. Dơdruong Bojiêng Plìng Tơ'neh Jong Rô-ma. Kontum: s.n.

This is Genesis 1-12, John, and Romans written in the Halang language. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1988. *TRUYỆN CÔ' CO' HO*. Hanoi: Nhà xuất bản văn hoá dân tộc.

Koho folktales in Vietnamese, 173 pages.

Unknown Author. 1989. Hla Bơar Hơri. s.l.: s.n.

This work is a Bahnar hymnal with 337 hymns as well as church liturgy. This work is located at the Linguistic Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 2011. s.n. Vơn gŭq yăh khõe. s.l.

Chrau health and hygiene booklet.

Unknown Author. 2013. Bible Jesus Movie In Bahnar clip1. http://www.youtube.com/watch?v=gdsSFHjDNZM accessed 16 May 2013.

Blood, Henry, and Evangeline Blood (trans.) 1969. The Origin of Dak Nue. MKS 3.61–63.

Legend from Eastern Mnong translated into English.

Cohen, Patrick. 1966. Jeh Computer Concordance. s.l.: s.n.

Referenced by Cohen, Patrick D. (1976) in MKS 5. This work was "Produced under the auspices of the Summer Institute of Linguistics through the University of Oklahoma under National Science foundation grant No. RS 00307." (page 152, MKS 5).

Mơ, Siu, and John Banker. 1962. Bahnar Texts. Grand Forks: Summer Institute of Linguistics.

This work contains 30 Bahnar stories. Bound in 2003.

Phillips, Richard L. n.d. Mnong Bunâr Health, Science, Ethics, and Arithmetic. Vietnam Data Microfiche Series VE14-12.. s.l.: s.n.

Referenced by Phillips (1963). Central Mnong Bunâr.

Smith, Kenneth D. 1973. Sedang Song-poetics. Typescript. s.l., ms.

Thomas, David, and Dorothy Thomas. 1982. Sirăq Cô Voq De Kinh-thánh Tân-ước. s.l.: New York International Bible Society.

Chrau New Testament. Copy held at the Linguistic Institute library at Payap University, Chiang Mai, Thailand.

Võ Thủ Lễ. 1988. Piăm. s.l.: Nhà Xuất Bản Đà Nẳng [Da Nang Publishing House].

This is the story of Piăm written in Bahnar.

4.8 Language Learning Materials

Unknown Author. 1960. *Stieng*. Huẽ: s.n.

Language learning phrasebook written in Vietnamese and Stieng. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Unknown Author. 1961. *Tiêng Bahnar*. Kontum: s.n.

Vietnamese – Bahnar language learning booklet. Located in the Linguistics Institute Library at Payap University, Chiang Mai, Thailand.

Banker, Elizabeth, Mơ, and Sip. 1973. Tơdrong pơhrăm nởr Pơma = Bài học tiếng Bahnar = Bahnar language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuốn 20. Saigon: Department of Education.

Banker, John E., and Yup. 1974. ‘Bai pơhrăm nởr Bahnar Kontum = Bài học tiếng Bahnar Kontum = Kontum Bahnar language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuốn 14. Saigon: Department of Education.

Cuốn 14 probably stands for book 14 in a series or volume 14.

Cohen, Patrick D., Dwight Gradin, and Thŭng. 1976. ‘Bai pơsèm hŏk tơtayh Jeh = Bài học tiếng Jeh Jeh language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuốn 15 Phần 2. Manila: Summer Institute of Linguistics.

Cooper, James S. 1971. ‘Bai hŏk nởr Halăng = Bài học tiếng Halăng = Halăng language lessons. ủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuốn 6 Phần. Saigon: Department of Education.

Đồ, Đinh, Jacqueline G. Maier, and Đinh Mốc. 1974. Bay hok pok Kool Kua = Bài học tiếng Cua = Cua language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam 10. Saigon: Department of Education.

Referenced by Haupers (1991).

Evans, Helen, and Peggy Bowen. 1963. *Kơho Language Course*. Dalat: Christian & Missionary Alliance.

Koho language learning book, part dictionary and part phrase book. Also descriptions of Koho grammar are found throughout the book. Thomas, David n.d.(a) mentions this work as copyright 1965 and on mimeograph.

Gradin, Dwight and Pat Cohen. 1970. Jeh Dictionary and Jeh Basic Lesson. s.l., ms. Copy held at Payap University, Linguistic Institute, Library.

There is also a single copy of Jeh Basic Lessons 1963, in the LI library at Payap University.

Gregerson, Marilyn, and Paul Neo. 1974. ‘Bai pơchrâm nâr Rơngao = Bài học tiếng Rơngao = Rongao language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuốn 19. Saigon: Department of Education.

Keller, Charles. 1977. Brao Vocabulary Language Lessons, Miscellaneous Field Notes. Dallas: Summer Institute of Linguistics.

Referenced by Jacq and Sidwell (2000).

Kpơr, Y. Kem, and Richard L. Phillips. 1974. Nti ngơi nau Bu Nong = Bài học tiếng Mnông = Central Mnong language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam 11. Saigon: Department of Education.

Brief introduction about the Mnong people and introduction to Central Mnong. There is a major script difference between Central Mnong spoken in Vietnam and Cambodia. This work follows the script and variety of Central Mnong found in Vietnam. 30 pages.

Phillips, Richard L. 1963. Mnong Language Course. Vietnam Data Microfiche Series VD14-78. s.l.: s.n.

Grammar, phrase book, and language learning book for Central Mnong (Bunâr).

Smith, Kenneth D. 1967. ´Bái Hòk Tơpui Rơtéang Sedang Language Lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam Cuôn 2 Phân 2. Saigon: Summer Institute of Linguistics.

Thomas, David, and Dajao Jaken (Thời). 1981. Minsăm trong Chrau = Bài học tiếng Chrau = Chrau conversation lessons. Vol. 2. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam 1. Saigon: Department of Education.

This work gives simple sentences and phrases that are written side-by-side in English, Vietnamese and Chrau. This work would be useful for anyone who wanted to learn simple Chrau. This work is located at the Linguistic Institute library at Payap University.

Trebilco, Oliver, Joyce Trebilco, and Đinh Ngiah. 1974. 'Boi Hõc Bà Hrê Hrey Language Lessons. Tủ sách ngôn-ngữ dân-tộc thiểu-số Việt-Nam Cuốn 12 Phần 2. Saigon: Trung-Tâm Học-Liệu Bộ Giáo-Dục Xuất-Bản.

Hre language.

Y Tang Hmok. 1976. Pơp riêm ngơi pơp Mnông Lam = Bài học tiếng Mnông Lăm = Mnong Lam language lessons. Tủ sách ngôn-ngữ dân tộc thiểu số Việt Nam 23. Manila: Summer Institute of Linguistics.

Simple side-by-side language lessons written in English, Vietnamese, and Eastern Mnong.

4.9 Literacy Materials (includes orthography)

For additional Bahnaric literacy materials contact the Linguistic Institute library at Payap University, Chiang Mai.

Unknown Author. 2013. Oh hok sarooy pok kool Kua: Puq 1, phaloot sarop (Em học vần tiếng Cua: Quyển 1, lớp vỡ-lòng). s.l.: SIL.

This work is a Cua primer #1. This website contains more information on this work: http://www.sil.org/resources/archives/30754 accessed 16 May 2013.

Unknown Author. 2013. Oh hok sarooy pok kool Kua: Puq 2, phaloot sarop (Em học vần tiếng Cua: Quyển 2, lớp vỡ-lòng). s.l.: SIL.

This work is a Cua primer #2. This website contains more information on this work: http://www.sil.org/resources/archives/30889 accessed 16 May 2013.

Unknown Author. 2013. Oh hok sarooy pok kool Kua: Puq 3, phaloot sarop (Em học vần tiếng Cua: Quyển 3, lớp vỡ-lòng). s.l.: SIL.

This work is a Cua primer #3. This website contains more information on this work: http://www.sil.org/resources/archives/30850 accessed 16 May 2013.

Gregerson, Marilyn. 2009. Learning to read in Ratanakiri: a case study for northeastern Cambodia. International Journal of Bilingual Education and Bilingualism 12. http://www.tandfonline.com/doi/pdf/10.1080/13670050902935789#.UlNo09LXCNI accessed 8 October 2013.

To access this work, one must pay a fee of $37.

Hiett, Steven. 2003. An Assessment of Khmer Language Skills and Literacy Levels Within the Adult Hilltribe Population of Mondulkiri Province, the Kingdom of Cambodia. Phnom Penh: International Cooperation Cambodia.

Referenced by ICC (2006). Mondulkiri Province is a primarily Central Mnong speaking area.

Jacq, Pascale. 2004. The development of a Lao-based orthography for Jruˀ. Mon-Khmer Studies (MKS) 34.97–112. http://sealang.net/sala/archives/pdf8/jacq2004development.pdf accessed 8 October 2013.

Maier, Jacqueline, and Eva Burton. 1975. Cua Rhyming Dictionary and Primer Glossary. Linguistic Institute Library Payap University.

This work was found in a binder and is probably 100 pages in length.

Phillips, Richard L. n.d. Mnong Bunâr Culture-Folklore Reader. Vietnam Data Microfiche Series VE14-16. s.l.: s.n.

Referenced by Phillips (1963). Central Mnong Bunâr.

----. n.d. Mnong Bunâr Primers and Guide. Vietnam Data Microfiche Series VD14-01.. s.l.: s.n.

Referenced by Phillips (1963). Central Mnong Bunâr.

Généralités sur les populations montagnardes do Sud indochinois

5. Bahnaric languages Index

Language names are hyperlinked to the Ethnologue

Language Name	EGIDS	ISO Code	Remarks
Alak	6a	alk	
Bahnar	5	bdg	
Bout	-	No code	See Parkin, Robert (1991)
Chrau	6b	crw	Researched by David D. Thomas
Cua	6b	cua	
Halang Doan	6b	hld	
Halang	6b	hal	
Hre	5	hre	
Jeh	5	jeh	
Jeng	6a	jeg	
Kaco' or Lamam	6b/6a	xkk or lmm	
Karol	6b	rka	
Kasseng or Tareng	6b/6a	kgc and tgr	Sidwell recommends a four-way language merger of Kasseng/Tareng/Talieng/Trieng
Katua	6a	Kta	
Kavet	6b	krv	
Kayong	6a	kxy	
Khaonh	-	[no code]	Referenced in Bahnaric Language Cluster Survey by ICC
Koho	5	kpm	
Kru'ng 2	6b	krr	
Lave (Laos) or Brao (Cambodia)	5	brb	
Laven	6a	lbo	
Lawi or Lavi (Saveung)	-	[no code]	Chazée mentions this ethnic group. Bradley calls this group Swoeng
Maa	6a	cma	
Mel	-	[no code]	Referenced in Bahnaric Language Cluster Survey by ICC
Monom	6a	moo	
Mnong, Central	5	cmo	
Mnong, Eastern	6a	mng	
Mnong, Southern	6a	mnn	
Nyaheun	6a	nev	
Oy	6a	oyb	Sidwell recommends Oy and The should be merged.
Ro'ang	-	[no code]	Referenced in Bahnaric Language Cluster Survey by ICC
Rengao	6b	ren	
Romam	6a	rmx	
Sapuan	6a	spu	
Sedang	6a	sed	Researched by Kenneth D. Smith
Sok	6a	skk	
Sou or Su'	6a	sqq	
Stieng Bulo	6a	sti	There is debate if the Stieng languages should be

			merged.
Stieng, Budeh	6a	stt	
Suai or Juk	-	No code	Sidwell (2003)
Takua	6a	tkz	
Talieng	6a	tdf	
Tampuan	6b	tpu	
The	6a	thx	Sidwell recommends Oy and The should be merged.
Thmon, Thmoan, or Thmaun	-	No code	Referenced in Bahnaric Language Cluster Survey by ICC
Todrah	6a	tdr	
Trieng	6a	stg	