

*Collected papers on Southeast Asian
and Pacific languages*

Also in Pacific Linguistics

Bradley, David (ed.), 1997, *Papers in Southeast Asian Linguistics No. 14: Tibeto-Burman Languages of the Himalayas*.

Clark, M. (ed.), 1997, *Papers in Southeast Asian Linguistics No. 16*.

James A. Matisoff, 1997, *Sino-Tibetan numeral systems: Prefixes, protoforms and problems*.

Peiros, Ilia, 1996, *Katuic comparative dictionary*.

Thomas, David (ed.), 1998, *Papers in Southeast Asian Linguistics No. 15: Chamic Studies*.

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

Pacific Linguistics is a publisher specialising in grammars and linguistic descriptions, dictionaries and other materials on languages of the Pacific, the Philippines, Indonesia, East Timor, southeast and south Asia, and Australia.

Pacific Linguistics, established in 1963 through an initial grant from the Hunter Douglas Fund, is associated with the Research School of Pacific and Asian Studies at the Australian National University. The Editorial Board of Pacific Linguistics is made up of the academic staff of the school's Department of Linguistics. The authors and editors of Pacific Linguistics publications are drawn from a wide range of institutions around the world. Publications are refereed by scholars with relevant expertise, who are usually not members of the editorial board.

FOUNDING EDITOR: Stephen A. Wurm
EDITORIAL BOARD: Managing Editors: Malcolm D. Ross, John Bowden and Darrell T. Tryon, I Wayan Arka, Andrew Pawley, Paul Sidwell

Collected papers on Southeast Asian and Pacific languages

edited by
Robert S. Bauer



Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University

Published by Pacific Linguistics
Research School of Pacific and Asian Studies
The Australian National University
Canberra ACT 0200
Australia

Copyright © The authors

First published 2002

National Library of Australia Cataloguing-in-Publication entry:

Collected papers on Southeast Asian and Pacific languages

Bibliography.
ISBN 0 85883 528 2

1. Austronesian languages. 2. Asia, Southeastern – languages.
I. Bauer, Robert S. II. The Australian National University.
Research School of Pacific and Asian Studies. II. Title.
(Pacific Linguistics 530).

495.9

Typeset by Jeanette Coombes
Cover design by Emily Brissenden
Cover photos were taken by Paul Sidwell at Bahnar village, Kontum,
Vietnam Central Highlands, April 2002
Printed and bound by Union Offset Printers, Fyshwick, Canberra

Table of contents

	<i>Notes on contributors</i>	vii
	<i>Introduction</i>	ix
I	<i>Linguistic analysis:</i>	
1	Impact of linguistic borrowing on Meiteilon (Manipuri) A.G. Khan	3
2	Functions of 'give' and 'take' in Lao complex predicates N.J. Enfield	13
3	Vietnamese verbal reduplication Sophana Srichampa	37
II	<i>Language classification:</i>	
4	Nùng An: origin of a species Jerold A. Edmondson	51
5	Morphosyntactic evidence for the position of Chamorro in the Austronesian language family Lawrence A. Reid	63
6	A brief look at thirteen Mon-Khmer languages of Xekong Province, Southern Laos Theraphan L-Thongkum	95
III	<i>Discourse analysis</i>	
7	The tiger mother's child and the cow mother's child: a preliminary look at a Bru epic John Miller and Carolyn Miller	111
8	The temporal movement of the Hlai (Li) origin myth Somsong Burusphat	141

IV Sociolinguistics

- 9** The future of Nyah Kur
Suwilai Premsrirat 155

V Historical linguistics

- 10** A comment on Gedney's proposal for another series of voiced initials
in Proto Tai
Graham Thurgood 169
- 11** The rise and fall and rise and fall of Proto Malayo-Polynesian
Stanley Starosta 185

Notes on contributors

Edmondson, Jerold A. Jerry is a professor in the Program in Linguistics at the University of Texas at Arlington. He is also Honorary Professor of the Central University of Nationalities in Beijing, China. His linguistic interests include historical comparative linguistics and the minority languages of China, Laos, Myanmar, Thailand, and Vietnam.

Enfield, N.J. Nick is Postdoctoral Researcher at the Language and Cognition Group, Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands. His major research interests include the following areas: grammatical description (especially of mainland Southeast Asian languages); semantics, pragmatics, and polysemy; social, cultural, and biological aspects of language (social intelligence, social group dynamics, ground-level mechanisms of language stability and change, 'language contact').

Khan, A.G. Professor Khan is Professor and head of the Department of English at the Manipuri University, India. His interests lie in English language teaching and the phonetics and phonology of Manipuri languages.

Miller, John and Carolyn Miller. Currently based in Chiangmai, Thailand, John and Carolyn are field researchers with the Summer Institute of Linguistics. Carolyn is also the President of SIL International. They have conducted linguistic research in Vietnam, Malaysia, and Thailand, and their research interests include Mon-Khmer and Austronesian languages, language maintenance and shift, and language attitudes.

Reid, Lawrence A. Laurie is Emeritus Professor at the University of Hawai'i in Honolulu, USA. His research interests are in Philippine languages and linguistics, and syntactic change in Austronesian.

Somsong Burusphat. Acaan Somsong is Associate Professor in the Institute of Language and Culture for Rural Development, Mahidol University at Salaya, Thailand. Her research interests are in the area of discourse analysis of Southeast Asian languages, Tai linguistics, and teaching Thai as a second language.

Sophana Srichampa. Acaan Sophana is Assistant Professor in the Institute of Language and Culture for Rural Development, Mahidol University at Salaya, Thailand. Her major research interests are in Vietnamese linguistics, language, and culture.

†**Starosta, Stanley.** Stan was Professor in the Department of Linguistics at the University of Hawai'i, Honolulu, USA. His major research interests were in grammatical theory, computational linguistics, and the structure and prehistory of the languages of East, Southeast, and South Asia, and the Pacific. He passed away in July 2002.

Suwilai Premsrirat. Acaan Suwilai is Associate Professor in the Institute of Language and Culture for Rural Development, Mahidol University at Salaya, Thailand. She also chairs the Institute's PhD Program in Linguistics and Indochinese Studies Committee. She is also one of the chief editors on the Editorial Board of *Mon-Khmer Studies*. Her research interests are in the areas of field linguistics of Mon-Khmer languages, ethnolinguistics, and sociolinguistics.

Theraphan L.-Thongkum is Associate Professor in the Department of Linguistics and Director of the Research Center for Linguistic and Cultural Diversity in the Mae Khong Region, Faculty of Arts, Chulalongkorn University, Bangkok, Thailand. Her major research interests are in phonetics and Southeast Asian linguistics with an emphasis on the Tai-Kadai, Mon-Khmer, and Hmong-Mien languages.

Thurgood, Graham. Graham is Assistant Professor at the California State University at Chico in the USA. His major research interests are in Southeast Asian historical linguistics, particularly Chamic, Tai-Kadai, and Tibeto-Burman.

Introduction

It was a little over three years ago that I received an email from Malcolm Ross, Managing Editor of *Pacific Linguistics*, in which he asked me if I would be interested in editing a volume. It would be up to me to solicit the contributors for the papers. One caveat was that no papers on Chinese could be included because of the fonts problem with the Chinese characters — this more or less eliminated me as a possible contributor since almost everything I have ever worked on has involved Chinese. As things turned out, this proved to be to my advantage because it meant my job was to concentrate solely on editing the papers of the contributors, and so I wouldn't have to fret about my own. Never having edited a book before, I thought this project could be an interesting, challenging, and valuable experience, and so I accepted Malcolm's invitation.

Right away Malcolm sent me one interesting manuscript in search of a volume in which to be published, and so I was in business right from the start. Early in 1998 via email I invited papers from friends and former colleagues in the USA, Europe, and Southeast Asia. I am very happy to say that most people accepted my offer and this was deeply gratifying. Papers began arriving in late 1998 and by the end of 1999 I had received enough manuscripts for publication in the volume.

Well, as already mentioned, I thought editing a volume would be an interesting experience; as it turns out, it indeed has been so and — I should also note — a tremendous amount of work. Thank Heavens for computers, word-processing software, email, etc. Ironically, these hi-tech things have undeniably made my life as editor much easier, but they also introduced their own special, quirky complications. My contributors, bless their hearts, sometimes used their word-processing programs to do some rather 'creative' things with their manuscripts which I was not able to override when I thought a change was in order. There were also a couple of papers (whose authors will go nameless here) who sent me manuscripts that my computer either simply could not read or — what was much worse — that caused it to hang up. I had thought the problem of incompatibility between Macs and PCs had been solved but my experience says otherwise. Oh well, with Malcolm's help we seemed to have overcome these difficulties and headaches, and I finally have been able to finish editing all eleven papers which had been initially contributed.

In introducing the papers that make up this *Pacific Linguistics* volume, I would like to note first that their authors are based in four countries from around the world: India, the Netherlands, Thailand, and the USA (with the majority of authors living in these last two countries). Second, the wide range of languages which have been investigated in these papers represent the following four major language families and one subgroup of mainland and

insular Southeast Asia: (1) *Tibeto-Burman* with Meiteilon (Manipuri) in India; (2) *Mon-Khmer* with Alak, Chatong, Dak Kang, Kaseng, Katu, Laven, Lavi, Nge', Suai, Ta Oi', Tariang, Tariw, and Yaeh in Laos; Bru and Nyah Kur in Thailand; and Vietnamese in Vietnam; (3) *Tai* with Nung An in Vietnam; Lao in Laos; and Hlai in Hainan, China; (4) *Austronesian* with Chamorro in the Marianas Islands; and (5) *Malayo-Polynesian* itself.

I have thought it most appropriate to classify the eleven papers in this volume under five linguistic topics and present them as follows:

I. Linguistic analysis

Three of the papers fall under this topic:

'Impact of linguistic borrowing on Meiteilon (Manipuri)' by **A.G. Khan** examines how the borrowing of foreign words has been accompanied by the borrowing of foreign sounds in Meiteilon, a Kuki-Chin language spoken in northern India. This borrowing has had the effect of introducing new phonemic contrasts into Meiteilon, as well as new rules for word formation.

'Functions of "give" and "take" in Lao complex predicates' by **N.J. Enfield** analyses in meticulous detail the grammatical roles that Lao 'ǎw 'take' and hāj 'give' perform in complex predicate constructions.

'Vietnamese verbal reduplication' by **Sophana Srichampa** describes clearly and exemplifies generously a wide variety of reduplication patterns that occur with Vietnamese verbs and stative verbs. Sophana observes that verbal reduplication expresses such meanings as consecutiveness, emphasis, repetition, augmentation, and diminutiveness. The data presented in this paper were collected during Sophana's several field trips to Hanoi.

II. Language classification

This section also includes three papers:

'Nung An: origin of a species' by **Jerold A. Edmondson** presents original lexical and phonological data on Tai languages he collected in Cao Bang province of Vietnam. Based on his comparison of these data with a Tai language spoken in Guangxi, China, he concludes that they share nearly identical Northern and Central Tai features, and that this is the result of immigration by the Nung An to Cao Bang, Vietnam from some Northern-Tai speaking area.

'Morphosyntactic evidence for the position of Chamorro in the Austronesian family' by **Lawrence A. Reid** examines copious morphosyntactic, lexical, and phonological evidence for language contact and genetic inheritance in order to resolve the thorny theoretical and practical problems that are associated with the classification of Chamorro. He concludes that Chamorro does not subgroup with any Philippine languages but is what he calls a first-order branch of Proto Extra-Formosan.

'A brief look at the thirteen Mon-Khmer languages of Xekong province, southern Laos' by **Theraphan L.-Thongkum** presents lexical and phonological material collected by the author during her seven field trips to southern Laos. From her analysis Theraphan proposes that one group of these languages belongs to the Katuic branch and the other group to the Bahnaric branch of Mon-Khmer.

III. Discourse analysis

There are two papers in this category:

'The tiger mother's child and the cow mother's child: a preliminary look at a Bru epic' by **John Miller and Carolyn Miller** is an in-depth analysis of various grammatical devices employed in an oral narrative text produced by a Bru speaker in central Vietnam; the Millers have focused on various kinds of linking forms, those that advance time progression, deictic expressions, and terms of address.

'The temporal movement of the Hlai (Li) origin myth' by **Somsonge Burusphat** analyses within the framework of Longacre and Dry several kinds of grammatical features found in main clauses which function to advance the movement of time in a legendary story that recounts the origin of the Hlai people who live on Hainan Island that lies off the coast of southern China.

IV. Sociolinguistics

'The future of Nyah Kur' by **Suwilai Premsrirat** takes a close look at an endangered language spoken in central Thailand. Nyah Kur is being learned by fewer and fewer young people who are abandoning it in favor of Thai. The author describes how contact with Thai has been significantly influencing the Nyah Kur phonological system. She proposes that the positive attitudes toward speaking Nyah Kur held by some age groups that make up the Nyah Kur speech community should encourage government officials and linguists to promote the survival of this language.

V. Historical linguistics

Two papers under this last category complete the volume:

'A comment on Gedney's proposal for another series of voiced initials in Proto Tai' by **Graham Thurgood** takes another look at a group of fifty-one lexical items for which the Proto-Tai reconstruction of their initial consonants has long presented a conundrum. Graham resolves the problem by marshalling reconstructed forms from other language families of the region which he believes clearly indicate that these items were borrowed from outside Tai-Kadai (a good portion from Chinese).

'The rise and fall and rise and fall of Proto Malayo-Polynesian' by **Stanley Starosta** addresses two questions: Is Malayo-Polynesian a subgroup, and Is it a first-order subgroup? After carefully evaluating the linguistic evidence, Stan concludes that the answer to the first question is 'maybe', while that of the second is 'no'.

As can be readily seen from the above brief description of these papers, this volume of Pacific Linguistics offers a rich diversity of Southeast Asian and Pacific languages and linguistic issues for the consideration of the discerning reader. As editor, I feel confident s/he will find the language data on which the papers are based to be most interesting, and the linguistic analyses stimulating and even provocative in some cases. Whatever the reader's own corresponding views, I predict they are more likely to be challenged than reinforced. With this thought foremost in mind, I submit this volume of papers to the reader with my invitation to explore, study, and savour.

Robert S. Bauer

Department of Chinese and Bilingual Studies
Hong Kong Polytechnic University
Hong Kong

5 June 2002

1 *Impact of linguistic borrowing on Meiteilon (Manipuri)*

A.G. KHAN

1 Introduction

Loan words are a more or less universal linguistic phenomenon, and it is an accepted fact that for all types of linguistic borrowing a sustained contact between two or more speech communities is a pre-condition. Historical analysis of Meiteilon vocabulary reveals that one of the sources of new words in the past has been borrowing from other languages in contact. Meiteilon, a member of the Kuki-Chin sub-group of the Tibeto-Burman language family, has been so hospitable towards Assamese, Bengali, Sanskrit, Hindi, Urdu, Perso-Arabic, English, etc. that the present-day vocabulary contains a large number of loan words. Most of the loan words have found their way into Meiteilon lexicon through Bengali. Such indirect borrowing is common, and sometimes it is difficult to trace accurately the history of a borrowed word. Most of these words were borrowed in the early part of the 18th century. But the borrowing has not been confined to the 18th century, and other languages also have contributed in varying degrees.

The vocabulary of a language is one of the most reliable sources for indicating the socio-cultural advancement of a particular linguistic community and reflecting the socio-cultural changes taking place during a specific period of time. As such, one could almost trace back the history of the Vaishnavite movement by making an in-depth study of the kinds of Sanskrit/Hindi loan words and the point of time when these words entered the Meiteilon lexicon. In earlier literary works like *Poireiton Khunthok*, *Panthoibi Khongul*, *Numit Kappa Nungsamei Puya*, *Naothingthong Phambal Kaba*, etc., the presence of loan words has not been noted, but with the introduction of Vaishnavism, the flow of loan words, mainly from Sanskrit and Hindi, increased to a considerable extent with the passage of time. Recently, the ever increasing popularity of the mass media has also contributed towards the adoption of new loan words from English into Meiteilon lexicon.

This is the second article by the writer on the study of loan words in Meiteilon. An earlier study observed that most loan words have been naturalised so that they remain indistinguishable by phonetic criteria from native words (Khan 1994). However, unassimilated loan words form a new phonological system; when compared with the phonological structure of native words, their separate identity of retaining alien features is

quite evident. This article discusses the impact of these borrowings on the overall phonological system of Meiteilon and briefly considers loan word phonology.

2 Loan word phonology

Meiteilon has borrowed not only loan words from languages with which it has been in contact but it has also accepted new segmental phonemes, their combinatory possibilities and syllable structures operating partly in harmony and partly in conflict with the native words. Some word-initial consonant clusters and a few consonantal phonemes which had no place in the phonological structure in the past have become an integral part of Meiteilon phonology. Present-day Meiteilon has thirty segmental phonemes, twenty-four of these are consonantal phonemes and six are vowel phonemes. Review of ancient Meiteilon literature reveals that in the earlier period there were only fifteen consonantal phonemes, namely, /p, ph, t, th, c, k, kh, m, n, ŋ, s, h, l, j, w/. Later on, five more consonantal phonemes /b, d, ɟ, g, r/ were added. In addition to these voiced stops, the voiced aspirated stops /bh, dh, ɟh, gh/ entered Meiteilon phonology at a later stage. As such, the consonantal phonemes may be divided into two groups: indigenous and borrowed phonemes. As indicated in Table 1 below, the consonantal phonemes included in the indigenous group can be considered to have been inherited directly from the earliest stages of the language. As stated earlier, through centuries of association with Assamese, Bengali, Sanskrit, Hindi, etc., Meiteilon has enriched itself not only by word borrowings, but also by accepting new consonantal phonemes which came with the loan words. Listed in Table 2 are borrowed phonemes /b, d, ɟ, g, bh, dh, ɟh, gh, r/ which have come from other languages.

Table 1: Indigenous consonantal phonemes in Meiteilon

Manner of articulation		Place of articulation								
		Bilabial		Dental	Alveolar		Palatal	Dorso-velar		Glottal
		vls	vd	vls	vls	vd	vls	vls	vd	vls
Plosives	unaspirated	p		t			c		k	
	aspirated	ph		t h					kh	
Fricatives					s					h
Nasals			m			n				ŋ
Lateral						l				
Semi-vowels			w					j		

Table 2: Borrowed consonantal phonemes in Meiteilon

Manner of articulation		Place of articulation				
		Bilabial	Dental	Alveolar	Palatal	Dorso-velar
Plosives	unaspirated	b	d		ɟ	g
	aspirated	bh	dh		ɟh	gh
Plosives						
Trill				r		

2.1 Introduction of new phonemic contrasts

No minimal pairs establish phonemic contrasts between voiceless and voiced stops in modern Meiteilon phonology, if one restricts oneself to indigenous words only. However, with the entry of loan words into Meiteilon, the phonemic contrast can be established with the help of words without making any distinction between indigenous and borrowed words as indicated in the following examples:

Table 3: New phonemic contrasts in Meiteilon

1.	/pəm/	‘pump’	/bəm/	‘two bamboo poles which extend at back of a bullock cart’
2.	/tan/	‘tune’	/dan/	‘charity’
3.	*/kun/	‘twenty’	/gun/	‘quality, merit, etc.’
4.	/phut/	‘foot’	/bhut/	‘ghost’
5.	*/thup/	‘to fold’	/dhup/	‘incense’
6.	*/khot/	‘to scratch’	/ghot/	‘pitcher used during worship, etc.’
7.	*/ləŋ/	‘thread’	/rəŋ/	‘colour’

It is evident from the above minimal pairs that word borrowing not only introduced new phonemes but also succeeded in introducing new phonemic contrasts and their combinatory possibilities which were not available earlier in Meiteilon. The words marked ‘*’ in Table 3 above are indigenous word and the rest are loan words.

2.2 Distributional characteristics of borrowed phonemes

Historical evidence suggests an absence of voiced stops and the trill /r/ in the ancient writings of Meiteilon, such as the *Puyas*, the writings on copper plates, and other ancient books. Borrowed phonemes number nine in all and are voiced consonants. They distinguish themselves from indigenous phonemes by their distributional characteristics. Their distributional characteristics are indicated in the table below; we note that voiced plosives /bh, dh, gh, ʃh/ occur only in loan words.

Table 4: Voiced plosives /bh, dh, gh, ʃh/ occur only in loan words

1. /bh/	
a. Initial occurrence:	
/bhabok/	‘audience, spectator, etc.’
/bhərəm ~ bhrəm/	‘doubt, forgetfulness’
/bhari/	‘too many, plenty of’
b. Medial occurrence:	
/bibhak/	‘department, division, etc.’
/səbhaʃətra/	‘procession’
2. /dh/	
a. Initial occurrence:	
/dhəmkə/	‘threat’
/dhobi/	‘washerman’

- b. Medial occurrence:
 /budhi/ 'intelligence, wisdom, etc.'
 /bjadhi/ 'ailment, disease, etc.'

3. /gh/

- a. Initial occurrence:
 /ghəri/ 'watch, clock'
 /ghot/ 'pitcher used at time of religious rituals'
- b. Medial occurrence:
 /rəghu/ 'proper noun (male)'
 /bighini/ 'obstacle, hindrance, etc.'

4. /ʃh/

- a. Initial occurrence:
 /ʃhuli/ 'a kind of bag'
 /ʃhinjur/ 'necklace, chain'
 /ʃhut/ 'falsehood'

Whenever /bh, dh, gh/ occur in the medial positions of indigenous words, they alternate with their voiceless counterparts /ph, th, kh/, respectively, as in /cəŋbhem ~ cəŋphəm/ 'entrance'; /kandhə ~ kanthə/ 'quilt'; /phəmgħre ~ phəmkħre/ 'has sat'. It is interesting to note that the borrowed phonemes have influenced the distributional pattern of some of the indigenous phonemes. However, the medial occurrence of /bh, dh, gh/, alternative pronunciation with /ph, th, kh/, respectively, is highly restricted and less frequent among native speakers.

Stops /b, d, ʃ, g/ and /r/ occur in word-initial position mostly in loan words; an instance of an indigenous word with /ʃ/ in word-initial position is /ʃgoi/ 'dance'. In the past this word used to alternate with /cəkoi/. The initial occurrence of /b, d, ʃ, g/ possible in a few onomatopoeic words only. As such, when the stops occur in word-initial position, the word is sure to be a loan; thus these phonemes indicate their foreignness and retain the mark of exotic origin, as in /bəndi/ 'prisoner, captive, etc.'; /dai/ 'responsibility'; /guli/ 'tablet, bullet, etc.'; /ʃura/ 'couple, pair'. The voiced stops cannot occur in word-initial position of verbs, except /farikhre/ 'has been promulgated, issued, etc.', which is an example of the verbal hybrid formation (cf. §4.2 below). The voiced stops can occur in word-final position only in loan words; an indigenous word cannot be closed by a voiced stop. In word-final position /b, d, g/ are more often devoiced. Total devoicing of final voiced stops can often be heard, despite representation of voicing in the orthography. The total devoicing of final voiced (unreleased) stops naturalises loan words by fitting them into the Meiteilon phonological system as in the examples below:

Table 5: Devoicing of word-final voiced stops in loan words

/hisab/	=	/hisap/	'account'
/beleb/	=	/belep/	'bulb'
/perised/	=	/periset/	'organisation'
/bibhag/	=	/bibhak/	'department, etc.'
/dag/	=	/dak/	'mark, stain'

In word-medial position, /b, d, ʒ, g/ occur quite frequently in indigenous words as well. But in most cases, they are non-contrastive with their counterparts /p, t, c, k/, respectively. It is quite interesting to note that the adoption of the voiced stop phonemes into Meiteilon has resulted in the voicing through progressive assimilation of /p, t, ck, k/ into /b, d, ʒ, g/ at the morphophonemic level, in the case of indigenous words. This rule of voicing through progressive assimilation operates in most cases and is illustrated below:

Progressive voicing assimilation rule in Meiteilon:

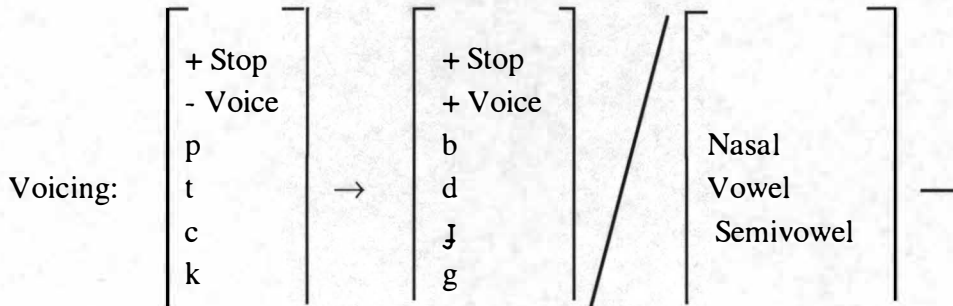


Table 6: Lexical examples of voicing assimilation rule in Meiteilon

/ce + paŋ/	→	/cebaŋ/	‘paper + sheet’	=	‘sheet of paper’
/ləw + pu/	→	/ləwbu/	‘paddy field + owner’	=	‘owner of paddy field’
/sən + pu/	→	/sənbu/	‘cow (cattle) + owner’	=	‘owner of a cow’
/khoi + toŋ/	→	/khoydoŋ/	‘navel + high’	=	‘one having big navel’
/nuŋ + caw/	→	/nuŋʒaw/	‘stone + big’	=	‘big stone’
/phi + cet/	→	/phiʒet/	‘cloth + piece(s)’	=	‘piece(s) of cloth’
/thaŋ + kon/	→	/thaŋgoi/	‘word + curved’	=	‘sickle’
/səŋ + koi/	→	/səŋgoi/	‘house + round’	=	‘house built in front of residential house for other purposes’

There are several exceptions to this rule, especially in the case of compounds. In such words, the voiceless stops /p, t, c, k/ occurring in initial position of the second root do not change even though the first root ends in a voiced sound as indicated in the following examples: /na + pəŋ/ → /napəŋ/ ‘one who is deaf’; /mə + tu/ → /mətu/ ‘his wife’; /i + ca/ → /ica/ ‘my issue’; /u + ku/ → /uku/ ‘bark of tree’. Numerous exceptions to the above rule confirm the fact that the voiced stops entered the Meiteilon lexicon at a later stage, and hence native speakers tend to be unsure about the voicing of consonants in such words.

The morphophonemic analysis of some indigenous words in present-day Meiteilon phonology reveals the tendency of free variation between /p, t, c, k/ and /b, d, ʒ, g/, respectively, as in the following examples: /ləipak ~ ləibak/ ‘country, earth, motherland’; /mətom ~ mədom/ ‘alone’; /tacəw ~ taʒəw/ ‘elder brother’. This alternation between voiceless stops and their voiced counterparts confirms the ‘structural readjustment’ and ‘phonemic redistribution’ of the voiceless stops (indigenous phonemes) in Meiteilon phonology. The distribution of borrowed phonemes is shown below in Table 7.

Table 7: Distribution of borrowed phonemes by word-position in Meiteilon
 [x = occurrence in loan words only; # = occurrence in indigenous words/
 or alternation with voiceless counterparts in certain cases; - = non-occurrence]

Phoneme:	Distribution according to word-position:		
	Initial	Medial	Final
<i>b</i>	x	x #	x
<i>bh</i>	x	x #	-
<i>d</i>	x	x #	x
<i>dh</i>	x	x #	-
<i>ɟ</i>	x	x #	x
<i>ɟh</i>	x	x #	-
<i>g</i>	x	x #	x
<i>gh</i>	x	x #	-
<i>r</i>	x	x #	x

3 Consonant clusters

Word-initial clusters in Meiteilon show very little variety as compared with the theoretically possible occurrences. Consonant clusters /gl-, pl-, dr-, tr-, sk-/ only found in recent loan words from English which are frequently used in day-to-day conversation as in the following examples: /glas ~ gilas/ 'glass'; /plet ~ pilet/ 'plate'; /drəm ~ dərəm/ 'drum'; /trək/ 'truck'. Clusters of /b, bh, d, dh, g/ with /j/ and /r/ occur only in loan words from Hindi or Sanskrit. Clusters with initial voiced stops occur in loan words but they are extremely rare. As such, these are restricted to loan words or onomatopoeic words as in the following examples: /bjadhi/ 'ailment, disease'; /briti/ 'scholarship'; /bhrəm/ 'forgetfulness, doubt'; /gjan/ 'knowledge'; /drəm/ 'oil container'.

4 Hybrid formations

The study of loan words in Meiteilon suggests that the most important and far-reaching impact of borrowing has not been so much on pronunciation as on vocabulary, in particular hybrid formations. The rapidity with which the loan words have been assimilated is shown by the speed with which some of them have become the basis of hybrid formations. Loan words have become quickly assimilated and have entered into an easy and natural fusion with native Meiteilon words. Hybrid formations are attested in nominal as well as verbal formations as discussed below.

4.1 Nominal formations

There are three types of nominal formations in Meiteilon: (1) native + loan elements; (2) loan + native elements; and (3) loan + loan elements. These three types are exemplified in the table below:

Table 8: Types of nominal formations in Meiteilon

1. Native + loan elements:
/sən gari/ from Meiteilon /sən/ + Hindi /gari/ ‘cow, ox’ + ‘cart’ = ‘bullock cart’
/khut moJa/ from Meiteilon /khut/ + Hindi /moJa/ ‘hand’ + ‘stocking’ = ‘gloves’
/selthum mentri/ from Meiteilon /selthum/ + Hindi /mentri/ ‘money, finance’ + ‘minister’ = ‘finance minister’
2. Loan + native elements:
/caməna/ from Hindi /ca/ + Meiteilon /məna/ ‘tea’ + ‘leaf’ = ‘tea’
/mofa phurit/ from Hindi /moJa/ + Meiteilon /phurit/ ‘stocking’ + ‘shirt’ = ‘vest’
/khərip ləwrok/ from Urdu /khərip/ + Meiteilon /ləwrok/ ‘summer’ + ‘crop’ = ‘summer crop’
3. Loan + loan elements:
/mətha khərap/ from Hindi /metha/ + Urdu /khərap/ ‘extremely’ + ‘dull, bad’ = ‘dull-headed fellow’
/rajya sərkar/ from Hindi /rajya/ + Urdu /sərkar/ ‘state’ + ‘government’ = ‘state Governemnt’
/əram cəwkri/ from Urdu /əram/ + Hindi /cəwkri/ ‘comfortable’ + ‘chair’ = ‘easy chair’

4.2 Verbal formations

Verbal formations are formed from loan + native verb elements as indicated by the examples in the table below.

Table 9: Verbal formations in Meiteilon

/Jəban ləwbə/ from Urdu /Jəban/ ‘statement’ + Meiteilon /ləwbə/ ‘to take’ = ‘to record the statement’
/iŋgit təwbə/ from Hindi /iŋgit/ ‘signal’ + Meiteilon /təwbə/ ‘to do’ = ‘to hint’, ‘to point out’
/dərgas thabə ~ dərkhās thabə/ from Urdu /dərkhās/ ‘petition’ + Meiteilon /thabə/ ‘to apply’ = ‘to put up an application’, ‘representation’
/norok tabə/ from Hindi /norok/ ‘hell’ + Meiteilon /tabə/ + ‘to go’ = ‘to go to hell’

The above examples of ‘verbal compounds’ function as verbs only when the loan items are combined with native verbal items. It is interesting to note that in the case of verbal compounds, the loan word always precedes the native verbal item. Verbal formations (with loan element + native verbal forms) are numerous in Meiteilon. A very few loan words take some of the native verbal suffixes as indicated below.

Table 10: Loan words with Meiteilon verbal suffixes

<i>/Jarikhre/</i> from Urdu <i>/Jari/</i> ‘to issue’ + Meiteilon <i>/khre/</i> ‘past perfect marker’ = ‘has been promulgated ordinance, etc.’
<i>/uthaibe/</i> from Hindi-Urdu <i>/uthai/</i> ‘revoke, dissolve’ + Meiteilon <i>/bə/</i> ‘infinitive marker’ = ‘to dissolve the Assembly, etc.’, ‘to revoke the orders’
<i>/cəlaibə/</i> from Hindi <i>/cəlana/</i> ‘to impose, to bring in practice’ + Meiteilon <i>/be/</i> ‘infinitive marker’ = ‘to impose (President Rule) etc.’

Hybrid formations are also attested at the morphological level with plural formations and gender markers, etc.

Table 11: Hybrid formation at the morphological level

Word	Grammatical marker	Gloss
<i>gari</i>	<i>siŋ</i>	‘vehicle’ + ‘plural marker’
<i>dabi</i>	<i>siŋ</i>	‘demand’ + ‘plural marker’
<i>dhəbər</i>	<i>məyam</i>	‘newspaper’ + ‘plural marker’
<i>thika</i>	<i>dar</i>	‘contract’ + ‘doer’ = ‘contractor’
<i>dərgas ~ dərkhas</i>	<i>kari</i>	‘petition’ + ‘writer’
<i>santi</i>	<i>siŋ</i>	‘proper noun’ + ‘male marker’
<i>santi</i>	<i>debi</i>	‘proper noun’ + ‘female marker’
<i>semaŋ</i>	<i>gi</i>	‘society’ + ‘possessive marker’

In regard to morphological changes, borrowing is confined only to the lexical singular forms of words. For their inflectional and syntactical structuring, they are treated like other native words.

5 Proper nouns and place names

After the mass acceptance of Vaishnavism, most names of people were taken either from Sanskrit or Hindi. These names are invariably hybrid forms, retaining only the Meiteilon surname as shown in the table below and the rest of the name is taken from Sanskrit or Hindu. In addition to these names, the extent of borrowing is observed in place names that reflect the influence of donor languages.

Table 12: Personal names in Meiteilon

Meiteilon surname	Sanskrit or Hindi name	Sex marker
<i>thokcom</i>	<i>məni</i>	<i>siŋ</i> (male)
<i>koiŋəm</i>	<i>pjari</i>	<i>debi</i> (female)
<i>səpəm</i>	<i>indira</i>	<i>cənu</i> (female)

6 Loan words used as echo words

Meiteilon has a few examples of echo words involving loans from Hindi, Urdu and other languages, e.g. /hisap pəsap/ 'account', /khorsa pərsa/ 'expenditure', /budu tetu/ 'foolish'. Study of echo words involving the loan words reveals the following: these echo words are not part of the vocabulary of echo words in the source language. The echo words are formed to intensify the meaning of the word concerned, as in /budu tetu/ 'foolish' (very foolish). In addition to the intensifying effect, echo words are also used in the sense of 'etc., etc.' Most echo words are highly restricted in daily usage.

7 Summary and conclusion

It is evident from this study that word borrowing has not only introduced new phonemes /b, bh, d, dh, ʃ, ʒh, g, gh, r/ into Meiteilon, but has also succeeded in introducing new phonemic contrasts and their combinatory possibilities. These borrowed phonemes did not form part of earlier Meiteilon phonology. The borrowed phonemes distinguish themselves from indigenous phonemes by their distributional characteristics. Further, these phonemes have influenced the distributional pattern of some of the indigenous phonemes also. The borrowed words have not only succeeded in adding to the vocabulary of Meiteilon, but they have also introduced several compound words and hybrid formations, proving the fact that these loan words have become an integral part of Meiteilon lexicon. However, most of the loan words have undergone the process of naturalisation.

References

- Khan, A.G., 1994, Voiced plosives in Manipuri (Meiteilon): a linguistic study. In J.S. Rao et al., eds 1994:23–27.
- Rao, J.S. et al., eds, 1994, *Linguistics and linguistics*. Indian Scholar Publications, M.P. India.

2 *Functions of 'give' and 'take' in Lao complex predicates*

N.J. ENFIELD

1 Introduction¹

Lao grammar crucially involves a small set of polyfunctional verbs, performing a range of important grammatical roles (especially in unmarked combination with other verbs). These include *'ăw* 'take', *hàj* 'give', *dâj* 'acquire', *máa* 'come', *pǎj* 'go', *míi* 'have', among a number of others. These items behave like main verbs, as well as performing duties which in other languages might be performed by morphological means. In this paper I examine some roles of *'ăw* 'take' and *hàj* 'give' in Lao grammar, with relation to their similar and related roles in complex predications, especially those involving mechanisms of valency-changing. We also see cases where 'give' and 'take' constructions are appropriated for other purposes. Some preliminary generalisations concerning argument structure and constituent structure are suggested.

2 Clause structure, valency and transitivity in Lao

The Lao language (Southwestern Tai, Laos) is a fairly extreme example of the isolating, analytic type. (See Enfield 1999 on Lao as the national language of the Lao PDR.) Typological features include a large phoneme inventory (very large number of vowel

¹ This paper was written while the writer was on a National Visiting Scholarship at the Research School of Pacific and Asian Studies, ANU, Canberra, 1997; this support is gratefully acknowledged. I am indebted to Tony Diller for his extensive and valuable comments and discussion of this work, as well as his support in my ongoing research on Lao. Grant Evans' assistance in Laos has been invaluable. The fieldwork upon which this paper is based was funded in part by ARC Grant No. A59601467 'Thai-Lao linguistic interaction' and this support is again gratefully acknowledged. This paper has also benefited from valuable comments on the data and analysis from Avery Andrews, Adam Chapman, Bob Dixon, Bill Foley, Nikolaus Himmelmann, Chris Manning, Andy Pawley, Alan Rumsey, and Jane Simpson. None are to blame for shortcomings.

My transcription of Lao is based on International Phonetic Association conventions, except for the following: glottal stop /' /, palatal and velar nasals /ñ, ng/, low central vowel /a/, and high back unrounded vowel /ɯ/. Tones are represented (approximately) as: level (/33/) /ˉ /; low falling (/21/) /ˊ /; high falling (/51/) /ˋ /; low rising (/213/) /ˊˊ /; high rising (/34/) /ˊˋ /.

contrasts), lexical tone, strong left-headedness in constituent structure, ubiquitous multi-verb structures. Like other ‘typical’ isolating/analytical languages, the features of ‘one-morpheme-per-word’, ‘no morphology’ (derivation or inflectional), and so on have been overstated (as has been pointed out for typologically similar Chinese; see DeFrancis 1984; Kratochvil 1968; Norman 1988). For example, there is much productive compounding which is clearly not ‘syntactic’ in nature (that is the semantics of the whole are not straightforwardly computable from the known semantics of constituent morphemes). There is a rich system of reduplicative (usually expressive) derivation, clearly a morpho-phonological process. Grammarians perhaps experience some psychological difficulty in viewing certain items as morphosyntactic markers, since they also function as full lexical items (mostly verbs) elsewhere. However, there is good reason to view them as grammatical morphemes, given their often strict syntactic position and restricted morphosyntactic behaviour when performing more structural or grammatical functions.

The basic Lao clause is (schematically) organised thus:

- (1) (Left Position) Subject [Verb (Object)]²

‘Left position’ (LP) is a kind of ‘topic’ slot, and may contain any nominal, whether it is a core argument of the verb or not. Further, LP may contain phrases, clauses, or even whole sentences. ‘Subject’ is an S/A pivot, but less pervasive in the grammar than, say, subject in English. For instance, Lao subject operates in restricted instances of equi control, but apparently does not figure in mechanisms of relativisation or reflexivity. Virtually any NP anywhere may be ellipsed if reference is contextually retrievable. (Rare exceptions include objects of certain ‘prepositions’ such as *càak* ‘from’, *káp* ‘with’, and *kēē* ‘to’. Indeed, a syntactic requirement for NPs to be explicitly mentioned is less common in Lao than a requirement that they be ellipsed — a number of control constructions, for example, require certain arguments to be omitted when subjects of structurally related clauses are coreferential.) ‘Movement’ of core arguments based on their discourse status is common, particularly fronting (into LP), as well as postposing (for reiteration or afterthought). Classifier phrases, which host the range of nominal modifications including ‘adjectives’, determiners, quantifiers and numerals, may be separated and moved away from the lexical noun in a kind of ‘modifier float’ (see §4.3 below). A good number of verbs are S=O ambitransitive (‘unaccusative’, as in English *break* or *open*). These are typically telic predicates with inherent resulting state, such as *sàng* ‘build/be built’, *mûng* ‘thatch (a roof)/be thatched’, *kǎang* ‘hoist/be hoisted’, but also include (more typically S=A) verbs like ‘eat’ and ‘work’. Demotion of O is effectively achieved in these latter cases by generic noun incorporation, as in, for example *kîn-khàw* [eat-rice] ‘eat (not necessarily rice)’, *àap-nâm*

² Abbreviation conventions: 3P/2P/1P (third/second/first person); CLF (classifier); C.LNK (clause linker); EXPR(essive); HON(orific); LNK (linker, focus particle); NEG(ative); NP (noun phrase); o.bro (older brother); PCL (particle); PFV (perfective); Q(uestion); RCP (reciprocal marker); RDP (reduplication); REL.FUT (relative future); TPC (topic marker); VP (verb phrase); y.sib (younger sibling). A and O are ‘semantic-syntactic categories’ based on grammatical status of typical agents and patients, respectively (Dixon 1994:6). S is not semantically defined — it refers to the single core argument of an intransitive clause. Period between words indicates glossed element is morphologically unanalysable. Proper names are glossed with initial capital and period. Examples are from my own 1997 corpus of spontaneous spoken texts, and references are to page/line number. Unmarked examples are constructed in collaboration with informants.

[bathe-water] 'bathe', *hēi-viāk* [do-work] 'work' (see Durie 1985:51 for the same in Acehnese).

Combination of these three aspects of clausal syntax — movement, ellipsis, and S=O ambitransitivity — can create significant structural ambiguity (see Chao 1968:72 on similar examples in Modern Standard Chinese):

(2) NP V_{ambitr (S=O)}} = [O_{fronted} \emptyset _A V], [A V \emptyset _O], or [S_O V]:

- a. *kāj kīn lēεw*
 chicken eat PFV
 i. 'The chicken, they have eaten.' ('chicken'=O)
 ii. 'The chicken has eaten it.' ('chicken'=A)
 iii. 'The chicken has been eaten.' ('chicken'=S_O)
- b. *khēεw bɔɔ thán míi*
 tooth NEG be.on.time have
 i. 'Teeth, it didn't yet have.' ('teeth'=O, actual reading; 853.7)
 ii. 'The teeth didn't yet have (it/them).' ('teeth'=A, possible reading)
 iii. 'There were not yet any teeth.' ('teeth'=S, possible reading)

There are cases in which it is perhaps impossible to decide whether to assign a structure like (a.i) or one like (a.iii) to a surface string 'NP V'. The structures share overt expression of the patient argument and differ in the level of contextual retrievability of the agent (at zero in (a.iii)). It is thus doubtful whether a clear line can be drawn separating a 'two-place' from a 'one-place' predicate in this case, since the idea of contextual retrievability of an argument is so difficult to characterise in a binary way, given that the context which provides possible referents includes not only what is currently activated in the discourse, but also what is in the inventory of cultural knowledge shared among speakers.

S=A ambitransitive ('unergative') verbs include *pǎj* 'go (somewhere)', *máa* 'come (somewhere)', *'òk* 'emerge (from/into something/somewhere)', *sūaj* 'help (someone)', *púk* 'wake someone up'/'be stimulating' (e.g. strong coffee or tobacco). Another class of S=A verbs has an undergoer subject: for example, *hūn* 'wake up, get a start', *tǎaj* 'die (of something)', and *'ǎj* 'cough (from something)'.

2.1 Preliminary note on three-place predicates in Lao

A certain number of three-place predicates (including *thaam* 'ask', *bòk* 'tell', *tháa* 'apply:on:to') may appear with no 'supporting' verbs, allowing adjacent expression of two objects, postverbally. The theme argument appears in immediately postverbal position:

- (3) *khòj thǎam mǒng láaw*
 1P ask o'clock 3P
 'I asked him/her the time.'
- (4) *láaw tháa sǐ húan lǎng nīi*
 3P apply colour house CLF this
 'S/he painted (i.e. 'applied paint (to)') this house.'

Alternatively, non-theme arguments may be overtly marked as peripheral, as in the following examples — peripheral arguments (underlined> are marked by *nám* ‘with/from’ (a verb-preposition, elsewhere a main verb ‘accompany’), *káp* ‘with’, and *kēē* ‘to’, respectively:

- (5) *lāj nāat-’ăw tōn sīn nám mǎa*
 chase grab-take lump meat with/from dog
 ‘She chased the dog to grab the lump of meat from it.’ (911.5)
- (6) *khōj dāj hāj sǎnǎa káp cāw lêw*
 1P did give promise with 2P PFV
 ‘I did give my promise to you already.’ (857.10)
- (7) *cá’ dāj lāw nūhán pakōp thámmā’ phāa pěn*
 REL.FUT get relate fable.tale comprise dharma in.order.to be
khátí’-khám-sāng-sōn hāj kēē ’anusón-hūn-lǎng
 provision-word-order-teach give to younger.generations
 ‘I’ll tell some dharmic tales as lessons to give to the younger generations.’ (838.4)

Further discussion of three-place predicates below will be more concerned with cases where two predicates conspire to form the necessary means to cooperate in hosting the expression of three arguments.

3 ‘Take’

The following sections describe functions of ‘*’ăw* ‘take’, including its role as a main verb, and, in more detail, as V_1 in a number of complex constructions of the form (schematically) ‘ $NP_1 V_1 NP_2 V_2 (NP_3)$ ’.

3.1 Main verb usage of ‘take’

The predicate ‘*’ăw* ‘take’ may be used as a simple transitive verb to mean ‘take/get’ (with intent/volition) or to mean ‘want (a thing)’:

- (8) *khan si ’ăw pǎa...*
 if REL.FUT take fish
 ‘If you are going to take the fish...’ (915.2)
- (9) *’āaj cāng sī káp-khúwn máa ’ăw nōng*
 o.bro so REL.FUT go.back-return come take y.sib
 ‘So (then) I’ll come back to get you.’ (900.8)
- (10) *’ăw cāk ’ăn*
 want (take) how.many CLF
 ‘How many do you want?’
- (11) *bō ’ăw*
 NEG want (take)
 ‘I don’t want any.’

3.1.1 V- 'take' constructions

There is a productive pattern in which a compound verb is formed with 'ăw 'take' as its second element, and a verb of carrying, gathering, or otherwise coming into possession of something, as its first. Consider the following examples:

- (12) ...lɔ̀k-'ăw nǎng...
 peel.off-take hide
 '...they peeled off the (tiger's) hide...' (944.7)
- (13) nǎng n̄an kɔ̄ɔ lēen-pǎj
 girl that LNK run-go
 cǎp-'ăw ngâaw thūi tók jūu táam dǎən
 grab-take sword which fall be.at along ground
 'The girl ran off, and grabbed the sword which had fallen on the ground.' (892.1)
- (14) hěn mǎa too nǎng khâap-'ăw sàj-kɔ̀k láaw lēen-pǎj lēew
 see dog CLF one carry.in.mouth-take sausage 3P run-go PFV
 'He saw a dog running away, carrying his sausages in its mouth.' (41.10)

3.2 Functions of 'take' in valency-changing mechanisms

As in many verb-serialising languages (see Lord 1993:Ch.5; Durie 1997), 'take' may be used in Lao to introduce an extra argument into the core in the following pattern (with 'ăw 'take' as V₁):

- (15) NP₁ V₁ NP₂ (DIR) [V₂ (NP₃)]

(I will henceforth use the abbreviations NP₁, V₁, NP₂, V₂, NP₃, and DIR to refer to the positions in (15) so marked.)

Note that it is rare for all three NPs in a three-argument clause to be overtly mentioned. Most examples below omit at least one argument, most often the subject. (Where necessary I will indicate 'missing' arguments with \emptyset . Note also that V₂ in (15) is almost always directly preceded by a DIRectional particle *paj* 'go' or *maa* 'come'; see examples below.) What import this has is not yet entirely clear, but it is certain that the effect is not simply predication of motion or direction. Structurally, as may become clearer in discussion below, it appears that the 'go/come' element is not necessarily a preverbal marker of V₂, but may be a complement of the phrase headed by V₁. This conclusion is based on facts about ellipsis of NP₂. Generally, if NP₂ is to be ellipsed (as its discourse status may allow), both V₁ ('ăw 'take') and DIR (*paj* 'go'/*maa* 'come') may remain, but if the entire 'V₁-phrase' (e.g. 'ăw take' and its nominal complement NP₂) is to be ellipsed, it is usually much more natural to (and sometimes impossible not to) also remove the 'go/come' verb which follows NP₂, suggesting it is attached to the V₁-NP₂ phrase.

As V₁, 'ăw 'take' may mark its object NP₂ as an instrument, or as a causee. It may also mark theme arguments with three-place predicates like *hāj* 'give' and *sāj* 'put' (which appear as V₂). The schema is further utilised in a 'pretransitive' construction, as described in Mandarin Chinese and some other serialising languages. These usages will now be discussed, in turn.

3.2.1 'Take' object as instrument

The direct object of 'take' may be an instrument in some action. That NP₂ has a role as 'instrument' can be diagnosed using a test involving semantic entailments:

- (16) If NP₂ in a string S of the form 'NP₁ "take" NP₂ (go/come) VP' is an instrument, then the same sentence with "'take" NP₂ (go/come)' omitted is entailed by S.³

In other words it is NP₁, the *subject* of 'take', which 'does' the action in VP. If the expression were rendered with only two arguments, it is only NP₂ which could be left out. In the following examples, I have put square brackets around omissible material (as usual, any of the NPs are omissible on their own):⁴

- (17) *mán* [*ǎw sǎn máa*] *cám kacĕe fông* *lǎj*
 3P take arrow come ram lock come.apart altogether
 'He broke the lock apart [with an arrow].' (176.17)
 (He took an arrow and rammed the lock; it came apart completely.)
- (18) *...bĕep* \emptyset [*ǎw hǔa-lāan*] *són* *kǎn*
 style take head-bald make.collide RCP
 '...in the manner of butting each other [with bald heads].' (72.6)
- (19) \emptyset [*ǎw néew-visáa māj máa*] *khĕeng.khǎn káp háw naa*
 take manner-plan new come compete with IP PCL
 'They will fight us [with a new strategy], you know.' (150.3)

These constructions can be represented schematically as follows, where the solid lines represent the status of NP₁ as actor with respect to *both* verbs:

- (20) $\left[\left[\text{NP}_{1, \text{AGT}} \right]_{\text{subject}} \left(\left[\text{V}_1 \text{ 'take' } \text{NP}_{2, \text{INSTR}} \right]_{\text{extra argument}} \right) \left[\text{V NP} \right]_{\text{predicate}} \right]_{\text{S}}$

In terms of argument structure, by which I mean a level of grammatical organisation specifying the number and relative prominence of a predicate's arguments (essentially following Manning 1996), this complex predicate could be described as follows:

³ Actually, there is occasionally not true 'entailment', since the VP may be purposive, i.e. merely intended, and perhaps never realised. Optional insertion (in the test) of the relative future marker *cá'* before V₂ might circumvent the problem, and would not render the rule ineffective in distinguishing the instrumental from other constructions.

⁴ There is a syntactic test for clause coordination in Lao (involving the possibility or not of insertion of the clause linker *lekaa*), which gives some important results here. *Lekaa* is acceptable before and after *máa* 'come' in (17), but much better after it. Arguably in pre-*máa* position the result is not equivalent to the original string, since it would entail literally 'coming' (i.e. from the place where the subject 'took' the arrow, to the place where s/he rammed the lock). In (18) and (19), however, *lekaa*-insertion is *not* acceptable, since it forces separate clauses, and therefore a literal reading for 'ǎw 'take', which could not apply where the 'ǎw-object is a nominal which cannot be literally 'taken', like *hǔa-lāan* 'bald head' or *néew visáa māj* 'new strategy'.

- (21) $V_1 < \overbrace{NP_1, NP_2} \text{---} V_2 < \text{---}, NP_3 >>$

Thus, in these instrumental expressions, two verbs, each with their own argument structure, are combined, whereby NP_1 is an argument of both verbs, *and* is the most prominent argument with respect to both verbs. NP_2 (instrument) does not appear in the (embedded) argument structure of V_2 .

3.2.2 'Take' object as causee

The object of 'take' in (15) may also be a causee. Here, in contrast to the examples in §3.2.1, NP_2 (the *object* of 'take'), *not* NP_1 , 'does' the action of the following VP. These examples fail the instrumental entailment test in (16). The following examples are causatives (at least semantically), as evidenced by the generally felicitous substitutability of *hāj* 'give' (as a causative verb) for 'ǎw 'take':

- (22) \emptyset 'ǎw *sían-miàng* *máa* *sūaj* (\emptyset)
 take S.M. come help
 '(He would) get Siang-Miang to (come and) help (him).' (93.16)
- (23) \emptyset 'ǎw *khón* *pāj* *khút-hēt* *khóng.mǎang*
 take people go dig-do/make canal
 'They got the people to dig the canals.' (267.9)
- (24) \emptyset 'ǎw *pasáasón* *pāj* *hían* *jūu* *vāt* *nǎa*
 take common.person go study be.at temple PCL
lekāa 'ǎw *khón* *pāj* *sǎn*
 CLNK take person go teach
 'They got the common people to (go and) study at the temples, you know, and they got people to (go and) teach them.' (255.1)
- (25) *...tēē* \emptyset 'ǎw *pasáasón* *pāj* *hēt* *náa.séeng*
 but take common.person go do/make irrigated.rice.field
 '...but they got the common people to make the irrigated rice fields.' (270.8)

Note that by simply replacing NP_2 , the object of 'ǎw 'take', in (25) with a (semantically) typical instrument, an instrumental rather than causative meaning (with the accompanying differences in entailments) emerges:

- (26) *tēē* \emptyset 'ǎw *khāng-cák* *pāj* *hēt* *náa.séeng*
 but take apparatus-engine go do/make irrigated.rice.field
 '...but they used machinery to make the irrigated rice fields.'

Thus, in (25) it is NP_2 (*pasáasón* 'common people') which makes the fields (not NP_1), while in (26) it is NP_1 that makes the fields, *not* NP_2 . It is important to note that despite the apparent structural similarity of the instrumental and causative 'ǎw-constructions, the two cannot be collapsed into a single construction, since it can be demonstrated that their semantic entailments differ.

The causative analysis for the string 'NP₁ 'take' NP₂ (go/come) VP' may be summarised as follows: where the correspondence lines indicate that NP₁ and NP₂ are actors (or, at least, most prominent arguments) with respect to 'take', and VP, respectively (see (21), above):

- (27) $[[\text{NP}_{1, \text{CAUSER}}]_{\text{subject}} [\text{V}_1 \text{'take'} \text{NP}_{2, \text{CAUSEE}}] [\text{V NP}]_{\text{predicate}}]_S$
-

I suggest the following argument structure corresponding to this (see (21), above), revealing that the most prominent (or 'highest') arguments of V₁ and V₂ are separate NPs (namely, the higher and lower arguments, respectively, of V₁):

- (28) $V_1 < \text{NP}_1, \text{NP}_2, V_2 < - (, \text{NP}_3) >>$
-

3.2.3 [*'Take' NP 'come'*] as *'take NP and come' versus 'cause NP to come'*

In the 'ǎw-causatives we have seen so far, it is certainly the causee which performs the action in V₂, but there is also a strong degree of responsibility on the part of the causer. Now, there are further examples whose proper interpretation (if interpreted as causatives) is complicated by the 'causer's' direct involvement in the V₂ action itself. Consider the following example:

- (29) *véeláa 'ǎw 'ǎa máa...*
 time take aunty come
 'When you bring your auntie (here)...' (194.27)

In the context in which it appears, (29) cannot mean 'when you "cause" your aunty to come here', nor can 'ǎa 'aunty' take an instrumental role in the action.

It is common for serialising languages to use a pattern 'take' + 'come/go' to mean 'bring/take' (Durie 1997; Lord 1993). However, different interpretations of this observation have been offered. Compare, for example, Lord's paraphrase of 'take + come' (i.e. 'bring') as 'cause to come' (Lord 1993:Ch.5), to Durie's 'verb-by-verb rendering' of a Yorùbá construction 'he took book come' as 'He takes a book; he comes' (Durie 1997:290). Lord's paraphrase suggests that the object of 'take' is the logical subject of 'come' (as in a causative, see (27), (28), above), while Durie's has a same-subject interpretation (parallel to the instrumental argument structure arrangement, (20), (21), above). Durie's description is perhaps closer to the truth, given that a causative interpretation is (semantically, at least) not really plausible here. It is hardly possible to paraphrase *John brought a cake* as 'John caused a cake to come'. Even if we accept this paraphrase, there remains an important distinction between *John brings a cake* and *John causes a cake to come*, namely that in the former case John must also come, while in the latter (if indeed a plausible context can be found), John need not come at all. Also, the 'coming' of John and the cake are hardly alike, since John's participation involves volition and agency, while the cake is presumably participating as a theme (or even an undergoer).

I am thus inclined to treat a construction like 'take' + 'come' in Lao as being basically idiomatic (for 'bring'), or perhaps as a special case of instrumental construction.

3.3 Role of 'take' with three-place predicates

Examples (30)–(33), below, show three-place predicates *vāj* 'put/place/fix', *sōng* 'send', *hàj* 'give', and *sāj* 'put/put in' appearing as V_2 , with the 'take'-object NP_2 the theme. NP_3 is obligatory here (i.e. not in the sense that it must be explicitly expressed, but that there must be a contextually retrievable referent for it). Unlike the valency-changing operations above, however, it is not possible to say in these cases that NP_1 or NP_2 'VP-ed' without reference to the other — all three are core arguments:

- (30) ϕ 'ăw *kīaw vāj pót*
 take cutter put post
 'She put the cutter on the post.' (929.1)
- (31) ϕ 'ăw *véen-tǎa máa sōng cék khúun*
 take mirror-eye ('spectacles') come send Chinaman return
 'He sent the spectacles back to the Chinaman.' (57.8)
- (32) ϕ 'ăw *ngāaw máa hàj 'āaj nēε*
 take sword come give o.bro PCL ('please')
 'Please give me the sword.' (891.15)
- (33) *tamláa_i khǎw kɔɔ 'ăw ϕ máa sāj thǒng-sǎa*
 recipe 3P LNK take come put bag-shirt
 'The recipe, he put in his shirt pocket.' (40.10)

Example (33) shows fronting of the theme *tamlaa* 'recipe'. The following version is fine, where the theme appears in the NP_2 slot (marked as ϕ_i in (33)):

- (34) *khǎw kɔɔ 'ăw tamláa máa sāj thǒng-sǎa*
 3P LNK take recipe come put bag-shirt
 'He put the recipe in his shirt pocket.'

Note that it is not only 'ăw 'take' which may mark the theme argument of a three-place predicate in this way. Other verbs of handling, such as *nǒk* 'lift' in the first clause of the following example, may also be used, where semantically appropriate. In this example, 'ăw 'take' marks the theme in the subsequent clauses 'put bamboo shoots in' and 'put water in':

- (35) ϕ *dǎng fáj lekāa nǒk ... mǎɔ-kěeng nǎj ...*
 light fire C.LNK lift pot-soup big

sāj tâw-fáj lekāa 'ăw nɔɔ.máj sāj ϕ / 'ăw nâm sāj ϕ
 put stove-fire C.LNK take bamboo.shoots put take water put
 '(He) lit the fire, and then put the big soup pot on the stove, and then put bamboo shoots in (it), and put water in (it).' (925.7)

Note that any of the three-place predicates may be expressed as an apparently simple transitive verb (i.e. without 'ăw 'take' and its object), as long as the identity of the three arguments is clearly understood from the context (the relevant clause is underlined):

- (36) *cāw jàak dāj nǎng néew-dǎj khǎɔj míi / khǎɔj hàj cāw*
 2P want acquire what type-which 1P have 1P give 2P
 'Whatever I have that you want to get, I'll give it to you.' (408.5)

Consider now the (notionally) three-place predicate *fǎng* 'bury':

- (37) *hǎw kɔɔ khút khǔm / 'ǎw sǎw fǎng ø mēen bɔɔ /*
 IP then dig hole take post bury be.so PCL(Q)
lǎng-càak 'ǎw sǎw fǎng ø lêew...
 back-from take post bury PFV
 'Then we dig a hole, and plant the post (in it), right? (Then,) after we've
 planted the post...' (21.13)

Speakers generally agree that the default referent of \emptyset in (37) is *dǐn* 'earth, ground',⁵ and thus the following sentence is acceptable:

- (38) \emptyset 'ǎw sǎw fǎng *dǐn*
 take post bury earth
 'S/he buried the post in the ground.'

Now, the following two strings are also acceptable, given appropriate discourse status of the relevant nominals:

- (39) *fǎng sǎw*
 bury post
 'S/he buried the post (in the ground).'

- (40) *fǎng dǐn*
 bury ground
 'S/he buried (it) in the ground.'

Note here that *fǎng* 'bury' cannot appear as a three-place predicate without a 'supporting' verb such as 'ǎw 'take' (except by using LP to accommodate a non-subject NP; see (59), below):

- (41) **fǎng sǎw dǐn*
 bury post ground
 (S/he buried the post the ground.)

Thus, as long as semantic roles of nominals are clear, three-place predicates such as *fǎng* 'bury' and *hàj* 'give' can be, and often are, treated as simple transitive verbs (i.e. two-place predicates).

3.3.1 'Effected object' construction

A subtype of the construction discussed in the section above is the 'effected object' construction, in which the two lower arguments refer essentially to the same entity, but in states before and after some process (predicated by V_2). Compare the English 'effected double object' construction in the translation of this example:

⁵ One might think from this example that the referent of \emptyset could be *khǔm* 'hole'. However, *khǔm* 'hole' cannot appear as direct object of *fǎng* 'bury'. Apparently, a direct object of *fǎng* 'bury' must refer to the *substance* in which something is buried, not to the empty space which provides a place for the thing to be buried.

- (42) *cõn nân kɔɔ 'ǎw ø pǎj hēt mía*
 bandit that LNK take go make wife
 'The bandit then made her his wife.' (893.4)

3.4 'Pretransitive' function of 'take'

In the examples considered so far, the object of 'ǎw 'take' (i.e. NP₂) has had a relation to the following verb phrase as either agent (as in causative constructions), or a secondary core argument in a three-place predication (either a theme, or an instrument). (The specific semantic/role relation of the instrumental argument to the following verb is not storable in terms of either 'agent' or 'patient' — see examples (17)–(19), above. It is simply an 'instrument', a long recognised basic case role (Fillmore 1968).) In the various examples above, we have seen 'take' serving as a mechanism to introduce a third core argument (conceptually, a participant 'midstream' on the 'action chain', neither an original 'energy source', nor a terminal 'energy sink' (Langacker 1991:Ch.7.1)).

In the following common construction, structurally equivalent to (15), 'ǎw 'take' performs no valency-change function whatsoever (i.e. no argument is 'added' or 'subtracted' from the core):

- (43) NP₁ 'take' NP₂ (go/come) V_{tr}

In (43), NP₁ and NP₂ are logical subject and object, respectively, of V_{tr}. Example (43) is notionally equivalent to a simple transitive clause in that it predicates a transitive event, and specifies two participant arguments (see the 'pretransitive' (*bǎ*) construction in Mandarin Chinese: Chao 1968:342ff.; Li & Thompson 1981:Ch.15; Lord 1993:114ff., inter alia; see also Jagacinski 1987 on the same construction in Tai Lue). Here are some examples:

- (44) *ø_i sī 'ǎw ø_j pǎj khàa sân bɔ'*
 RELFUT take go kill thus PCL(Q)
 'So they're going to kill us, are they?' (674.1)
- (45) *ø_i khwāt-ǎw ø_j lekkaa 'ǎw ø_j máa tīi*
 carve-take C.LNK take come beat
 'They'd carve the drums, and then beat them.' (262.9)
- (46) *phān kɔɔ 'ǎw too-nūi pǎj hían khúu-kǎn*
 3.HON LNK take CLF-this go study same-RCP
 'They also did study this.' (270.6)
- (47) *ø_i kɔɔ 'ǎw ø_j máa vāw tɔ-tɔ kǎn pǎj*
 LNK take come say connect-RDP RCP go
 'So, I tell (the stories), passing them on.' (33.20)

This kind of construction is not limited to simple transitive expressions. Consider the following examples, first showing *kǎn* 'eat', a simple transitive verb, and second a causative construction, with a third argument added to the clause, using the additional verb *hàj* 'give':

- (48) *lūuk khanɔɔj kǎn nóm*
 child 1P consume milk
 'My child drank milk.'

- (49) \emptyset *hàj lûuk khanôj kîn nóm*
 give child IP consume milk
 'I'll feed milk to my child.'

Now, NP₃ in this 'give'-causative construction (see §4, below) may be 'raised' in a pretransitive construction:

- (50) \emptyset 'ăw *nóm hàj lûuk khanôj kîn*
 take milk give child IP consume
 'I'll feed my child.' (845.6)

3.4.1 Range of use of the pretransitive construction

The precise semantic or functional import of the pretransitive construction in Lao is not yet entirely clear. Li and Thompson have noted that the equivalent construction in Mandarin Chinese may be used either when 'something happens to' the 'take'-marked NP, or when it is 'definite, specific, or generic' (Li & Thompson 1981:483). In general, presence of these conditions corresponds to increased transitivity (Hopper & Thompson 1980). In examples from the previous section, the 'ăw-marked NPs in (44) and (45) are clearly affected ('killed' and 'beaten' respectively), while those in (46) and (47) are, in contrast, atypical 'patients', being 'studied' and 'told' respectively, and thus hardly 'affected' in any literal sense. They are, however, referential and specific, in terms of their discourse status.

The following examples apparently display cases in which it would be almost unacceptable *not* to use the pretransitive construction:

- (51) *sàj-kòk nî ... caw 'ăw pǎj cǎwn*
 sausage this ... 2P take go fry
 'These sausages...you go and fry.' (39.10)
- (52) *?caw (pǎj) cǎwn sàj-kòk nî*
 2P (go) fry sausage this
 (You (go and) fry these sausages.)
- (53) *khǎw cá' 'ăw khón pǎj dát-sàng*
 3P REL.FUT take person go modify-build
 'They were going to take people for re-education.' (644.9)
- (54) *?khǎw cá' (pǎj) dát-sàng khón*
 3P REL.FUT (go) modify-build person
 (They were going to re-educate people.)

Why this is so requires further consideration, but some points can be made here. The verbs *cǎwn* 'fry' and *dát-sàng* 're-educate' both contain a strong notion of affectedness (specifically, a resultant change of state) of the patient. Also relevant is the discourse status of the O arguments *sàj-kòk nî* 'these sausages' and *khón* 'people'. Example (52) shows a referential and specific argument 'these sausages' appearing after the verb, a position strongly preferred for overt expression of new/non-referential arguments. Postverbal expression of any NP with this discourse status (i.e. referential and specific) is generally avoided. The following example lends support to this hypothesis, by showing a good

occurrence of 'sausages' in postverbal position, when the discourse status is *non-referential/non-specific*:

- (55) *cāw cǎwn sàj-kòk hàj khòj dēε*
 2P fry sausage give IP PCL(please)
 'Fry some sausages for me, please.'

In (53), the 'ǎw-object *khón* 'people' is generic, and as such may behave in similar fashion to 'definites' (Givón 1984:Ch.11; Langacker 1991:Ch.3).

The interaction between syntax and the discourse status of NPs in Lao is a fascinating and important area for further research, and one which is clearly central to many mechanisms of Lao grammar.

3.4.2 Purposive reading of pretransitive constructions

Pretransitive constructions may often be construed as purposive, with the lower predicate intended rather than asserted, and pragmatically defeasible. Consider the following example, with two possible readings:

- (56) *'ǎw ø máa 'āan*
 take come read
 i. 'He read it.'
 ii. 'He took it to read.' (56.10)

The actual reading in context is (56ii), as revealed by the following line of the text, in which *'āan* 'read' is explicitly negated:

- (57) *'òj lóot 'āan bōw dāj*
 oh so.then read NEG can
 'Oh! He couldn't read it.' (56.11)

3.4.3 Argument structure of the pretransitive construction

On the basis of the above discussion, I suggest the following complex argument structure for the pretransitive construction (see (21), (28), above):

- (58) $V_1 < \overbrace{NP_1, NP_2} \underbrace{V_2 <-, ->} >$

Here, both arguments of V_1 are arguments of V_2 , with the same respective prominence relations in the structure of each of the two (structurally combined) verbs.

4 'Give'

4.1 Main verb usage of 'give'

As a main verb meaning 'give', *hàj* may host its three arguments (donor, recipient, and gift) in a number of ways.

4.1.1 'Give' as a lone verb with the gift in Left Position

Donor and recipient NPs may appear as A and O, respectively, with the gift NP fronted, in LP:

(59) NP_{GIFT} NP_{DONOR} 'give' NP_{RECIPIENT}

(60) *pām hũa nân khòj hậj cầw*
 book CLF that 1P give 2P
 'That book, I gave you.'

Of other logically possible NP orderings, only [recipient-donor-'give'-gift] works, and is certainly marked in comparison to (59).

4.1.2 'Give' as a lone verb with the gift in postverbal position

The gift may appear in postverbal position in what looks like a double object construction:

(61) NP_{DONOR} 'give' NP_{GIFT} NP_{RECIPIENT}

This construction is best analysed as a case of noun incorporation, due to the strongly constrained range of nominals that may appear in the NP_{GIFT} slot in (61) (i.e. only non-referential/non-specific arguments are possible). Consider the following examples:

(62) *mēē dậj hậj sảnả đả phảnả đả sả đả*
 mother did give promise king-tiger fix.in.place
 'The mother did give the tiger king a promise.' (851.4)

(63) *cầw hậj ngớn khòj*
 2P give money 1P
 'You gave me money.'

Now, there are examples which appear to suggest that both gift-recipient and recipient-gift postverbal orderings are possible. The following example, with recipient preceding gift is fine, although perhaps less common (no examples appear in my texts):

(64) NP_{DONOR} 'give' NP_{RECIPIENT} NP_{GIFT}

(65) *cầw hậj khòj hầ đả kíp*
 2P give 1P five-hundred kip
 'You gave me 500 kip.'

Consider, however, the following unacceptable example, with the same constituent order as (65), but with the simple noun *ngớn* 'money' in the NP_{GIFT} position of (64):

(66) **cầw hậj khòj ngớn*
 2P give 1P money
 (You gave me money.)

It appears that the ordering in (64) and (65) results from a combination of zero anaphora and floating nominal modification or 'NP split'. The phrase *hầ đả kíp* 'five hundred kip' is a classifier phrase which quantifies *ngớn* 'money'. Example (65) may thus be analysed as having a 'zero' in the postverbal 'gift' slot, where the modifying classifier phrase *hầ đả kíp*

kìip 'five hundred kip' has 'floated' to sentence-final position, as made explicit in (67). The full structure, with the postverbal 'gift' slot filled, is shown in (68) (see (63) and (65), above):

- (67) *câw hàj ∅ khòj hà-lòj kìip*
 2P give 1P five-hundred kip
 'You gave me 500 kip.'
- (68) *câw hàj ngán khòj hà-lòj kìip*
 2P give money 1P five-hundred kip
 'You gave me 500 kip (of money).'

The float of nominal modification to final position results from a restriction (relating to discourse status of the gift argument) inherent in the noun-incorporating 'double object construction'. The following example, with the fully elaborated NP in postverbal 'gift' position is unacceptable (see (63)):

- (69) **câw hàj ngán hà-lòj kìip khòj*
 2P give money five-hundred kip 1P
 (You gave five hundred kip to me.)

Now, consider the acceptability of the following example, where the whole 'gift' NP is intact, with 'give'-recipient-gift order:

- (70) *câw hàj khòj ngán hà-lòj kìip*
 2P give 1P money five-hundred kip
 'You gave me five hundred kip.'

It appears that here the whole 'gift' NP appears in 'afterthought' position, and the structural 'gift' object slot (between 'give' and recipient) contains zero. The specificity of the overall argument is presumably what disallows it from appearing in postverbal position.

4.1.3 'Give' as V_2 in the 'take'-construction

A third, and more common way to use *hàj* 'give', is in a serial construction headed by *ǎw* 'take' (see §3.2, above):

- (71) NP_{DONOR} 'take' NP_{GIFT} 'give' NP_{RECIPIENT}
- (72) *háw ǎw ngán hàj mēε-thàw*
 1P take money give mother-old
 'I gave money to my mother-in-law.' (388.5)

This structure allows a complex NP like *ngán hà-lòj kìip* 'five hundred kip (of money)' to be expressed in full, without being split by modifier float, or moved to an outer position (cf. examples (63)–(70), §4.1.2, above):

- (73) *câw ǎw ngán hà-lòj kìip hàj khòj*
 2P take money five-hundred kip give 1P
 'You gave me 500 kip.'

A couple of points can be noted here. First, the combination of movement and nominal ellipsis can create further possible constituent orders, such as the following example (schematically, 'NP_{GIFT} 'take' 'give' NP_{RECIPIENT}'):

- (74) [lùuk-fáj-săaj nîi]ᵢ ϕⱼ 'ăw ϕᵢ hàj mán
child-fire-project ('torch batteries') TPC ('this') take give 3P
'Torch batteries, we, gave (to) them.' (412.6)

Second, verbs more specific than 'ăw 'take' may be used as V_1 in this context, where semantically appropriate (as noted above with other 'handling' verbs for transfer expressions; cf. *n̄ōk* 'lift' as V_1 with the three-place transfer predicate *sāj* 'place', (35), §3.3, above). Here, *nám* 'to lead, guide' is used in the V_1 slot:

- (75) cǎ' tōng nám sǎn nîi hàj sěnáa.ămáat
REL.FUT must lead official.letter this give military.forces
'We'll have to take this official letter to the military forces.' (89.11)

The fact that 'ăw 'take' is most common as V_1 in these constructions is due to its maximally abstract semantics as a verb of 'handling'.

4.1.4 Argument structure of 'give'

Based on the discussion so far, the argument structure of *hàj* 'give' as a transfer verb is similar to that for the pretransitive constructions (§3.4, above), in that the prominence relations of the two arguments of V_1 are preserved for V_2 . Compare (58), repeated here, with (76), a structure specific to transfer verbs like 'give', and made more explicit in (77), using the nominal arguments from example (72), above:

- (58) V_1 < NP₁, NP₂, V_2 <- , ->>
-

- (76) V_1 < NP₁, NP₂, V_2 <- , -, NP >>
-

- (77) 'take' < 'I', 'money', 'give' <- , -, 'mother-in-law' >>
-

Examples (58) and (76) differ only in that an extra argument appears under V_2 in the latter case. The correspondence of argument prominence relations across V_1 and V_2 remains the same (i.e. the first and second most prominent arguments of V_1 are the first and second most prominent arguments, respectively, of V_2).

4.2 Role of 'give' in complex expressions with other 'giving' verbs

The word *hàj* 'give' may appear as V_2 in complex expressions with other giving verbs, such as *m̄ōp* 'hand over', *sōng* 'send', and *thavăaj* 'present, offer'. In the examples with 'give' as a main verb which we have just seen, the main verb *hàj* 'give' is in V_2 position, and

the ('non-main') verb in V_1 position is 'ăw 'take', or another semantically appropriate handling verb. In the following cases, however, the 'main' verb is apparently V_1 , and *hàj* 'give' in V_2 position now plays a more structural role (in bringing a third argument into the core). As above, the theme argument is direct object of V_1 :

- (78) *phôn thūi-sút* \emptyset *kɔɔ mɔɔp* *múang* *hàj* *sĩnsáj*
 result at-extreme LNK hand.over kingdom give S.
 '(As) the final result, he handed over his kingdom to Sinxay.' (205.10)

As we have seen above, there are other possible surface orders due to movement and ellipsis — the following examples show postposing, and fronting, respectively, of the theme argument (direct object of V_2):

- (79) *háv cá'* *mɔɔp* \emptyset *hàj* \emptyset [*sáang-máa-ngúa-khwáaj-*
 IP REL.FUT hand.over give elephant-horse-cow-buffalo-
-sĩng-khɔɔng-păanakăan-kêew-věen-ngón-khám]_i
 -things-stuff-of.various.kinds-crystal-rings-silver-gold
 'I'll hand over livestock, goods, and many precious items.' (88.3)

- (80) [*thūk-sĩng-thuk-jāang kīaw.káp lūang nī*]_i
 each-thing-each-kind about matter this
'áaj *mɔɔp* \emptyset _i *hàj* *nɔɔng* *děj*
 o.bro hand.over give y.sib PCL
 'Everything concerning this matter, I hand over to you.' (94.12)

4.3 Further note on 'modifier float'

As mentioned above, a noun phrase in Lao may 'split' where the lexical noun and the classifier phrase (containing various quantifiers, determiners and modifiers) are separated, with the classifier phrase postposed to sentence-final position. Here is a typical example, in which the 'discontinuous' NP is underlined:

- (81) *táw dɔɔk-máj* *hàn* *têek* *sá'* *săam táw*
 vase flower-plant.suffix that break EXPR three vase
 'Those vases smashed, three (of them).' (63.12)

The following example reveals a further complication to those discussed above, whereby it appears that the (underlined) nominals in NP_2 and NP_3 slots (see the schema in (71)) refer to one and the same argument, namely the gift:

- (82) *phān* 'ăw *ngón* *hàj* *cét* *kìip*
 3.HON take money give seven kip
 'He gave me seven kip.' (332.3)
- (83) *phān* *dáj* 'ăw *khɔɔng.khwǎn* *hàj* 'ăn *nāng*
 3.HON did take gift give CLF one
 'He did give him a (certain) present.' (875.2)

It appears that the gift argument is (overtly) the direct object of two separate verbs in the structure. But I assume there is an empty argument slot immediately after *hàj* 'give' in these

cases, since in both cases a nominal referring to the recipient can felicitously be inserted immediately after *hàj* ‘give’. Thus, it is only a ‘fortuitous’ case of the gift argument appearing (only apparently) as the object of *hàj* ‘give’, resulting from a combination of movement (postposing of the gift argument’s classifier phrase), and ellipsis (of the post-V₂ recipient argument). Thus, the underlined nominals in these examples arguably form a ‘discontinuous’ noun phrase, as described in (81), above. (Alternatively, the utterance-final nominals in these examples could be analysed as performing an adverbial role.)

4.4 Peripheral argument marking

The word *hàj* ‘give’ may mark a peripheral beneficiary argument:

- (84) *khǒn-dǐn-khǒn-sáaj hàj khǎw hanǎa*
 dig.up-earth-dig.up-sand give 3P PCL
 ‘I dug up earth and sand for them, you know.’ (350.5)
- (85) *phùu-nân kɔɔ aan ø hàj láaw*
 person-that LNK read give 3P
 ‘That fellow read it for him.’ (54.18)

That the *hàj*-marked nominals are peripheral arguments here is revealed first by their potential for ellipsis (together with *hàj*, *not* requiring contextual retrievability, i.e. not specified by the core argument structure of the verb), and second by entailment relations with sentences whose *hàj*-phrase is omitted. The entailment which diagnoses peripheral status of a *hàj*-marked nominal is as follows: if NP₃ in a string S of the form ‘NP₁ V NP₂ *hàj* NP₃’ is a peripheral (typically benefactive) argument, then the same sentence with ‘*hàj* NP’ omitted is entailed by S.⁶ Thus, (85) entails the following:

- (86) *phùu nân kɔɔ 'āan ø*
 person that LNK read
 ‘That fellow read it.’

On the other hand, since the object of *hàj* ‘give’ in (72) (repeated here from above) is a core argument, (72) does *not* entail (87) (and thus fails the test for peripheral argument status):

- (72) *hǎw 'ǎw ngǎn hàj mēε-thàw*
 1P take money give mother-old
 ‘I gave money to my mother-in-law.’ (388.5)
- (87) *hǎw 'ǎw ngǎn*
 1P take money
 ‘I took money.’

Finally, note that while the effect of using *hàj* ‘give’ to bring in a peripheral argument is typically *benefactive*, the following example shows that this is not necessarily the case:

⁶ This may appear slightly problematic, since this entailment apparently holds for examples like (78). The crucial factor, however, which would rule out such examples, is that the ellipsed argument must *not require* contextual retrievability.

- (88) *khɔ̃ɔj kɔ̃ɔ bɔ̃ɔ hēt nǎng hāj mūu cāw*
 IP LNK NEG do what/anything give group 2P
 'I won't do anything to you people.' (411.4)

Out of context, (88) would normally be taken as benefactive (and would be translated 'I won't do anything *for* you people'), but it is clear from the context (the speaker is making a deal with a neighbouring group of people, that the two groups are to help each other by not obstructing each other's activities) that the *hāj*-phrase cannot be interpreted as benefactive.

4.4.1 Structural ambiguity in 'give' constructions

As already noted, there is possible ambiguity between *hāj* 'give' as a main verb with three core arguments, and as a marker of a peripheral (typically benefactive) argument. Another possibility exists, where *hāj* 'give' and its object may be read as the main predicate of a separate clause, in a purposive complement, or as in a clause chain:

- (89) *hǎw nǎng dāj hēt hāan hāj mán*
 IP still have.to ('get') make platform give 3P
 i. 'I still have to make a platform for them.'
 ii. 'I still have to make a platform to give them.' (26.13)

- (90) *cót tǎmláa hāj khɔ̃ɔj dēε*
 jot recipe give IP PCL
 i. 'Please write down the recipe for me.'
 ii. 'Please write down the recipe and give it to me.' (39.7)

Thus, possible readings for a string 'NP₁ V NP₂ 'give' NP₃' could be:

- (91) i. '[NP₁ V NP₂] for NP₃.' (e.g. (85))
 ii. 'NP₁ gives NP₂ to NP₃.' (e.g. (72))
 iii. '[NP₁ V NP₂] and then give(s) it (NP₂) to NP₃.' (e.g. (90))
 iv. '[NP₁ V NP₂] in order to give it (NP₂) to NP₃.' (e.g. (90))

It is perhaps the case that a distinction between (91 iii) and (91 iv) is impossible to make in cases like (89) and (90), given that they are situated in the future, and so in both cases the act of giving is unrealised. A test which would bring out the distinction involves negation of 'give' in a subsequent clause, forcing the purposive complement reading (see (90)):

- (92) *láaw cót tǎmláa hāj khɔ̃ɔj tēε láaw hēt tǎmláa hàn sǎ*
 3P jot recipe give IP but 3P do recipe that be.lost
lǎj bɔ̃ɔ dāj sōng hāj khɔ̃ɔj lēεw
 so.then NEG did send/present give IP PFV
 i. 'S/he noted the recipe for me but lost it, and never gave it to me.'
 ii. 'S/he noted the recipe to give me, but lost it, and never gave it to me.'
 iii. (*S/he noted the recipe and gave it to me, but lost it and never gave it to me.)

In practice, however, this apparent vagueness is unproblematic, as interpretation is easily resolved with reference to context.

4.5 Referential disjunct or ‘switch-reference’ marking

A semantically weakened, structurally functional role of *hàj* ‘give’ is as a referential disjunct (or ‘switch-reference’) marker in control constructions. (This is related to valency-change operations, since a signal of switched subject indicates that a new argument is introduced in the subject position of the second clause.) Consider these two examples of the complement-taking predicate *jàak* ‘want to’, which demands that the following verb have a zero subject coreferential with the main subject:

- (93) \emptyset *jàak sòn dee-nô*
 want fight PCL-PCL
 ‘He wants to fight, don’t you think?’ (122.9)
- (94) \emptyset *jàak 'òk pǎj sóm sǎn pǎj 'àap-nâm*
 want exit go enjoy garden go bathe-water
 ‘She wanted to go out and enjoy the garden, and bathe.’ (159.12)

Now, in the following, *hàj* ‘give’ appears immediately after *jàak* ‘want (to)’, marking reference of the following subject as non-coreferential with the main subject, and also leaving ellipsis of the lower subject optional:

- (95) *háv_i jáak hàj \emptyset _j sòn*
 IP want give fight
 ‘I want them to fight.’ (142.13)
- (96) *mēε kɔɔ bɔɔ jáak hàj lúuk tǎj*
 mother LNK neg want give child die
lúuk kɔɔ bɔɔ jáak hàj mēε tǎj
 child LNK NEG want give mother die
 ‘The mother didn’t want her child to die, and the child didn’t want its mother to die.’ (864.9)

4.6 Role of ‘give’ in causative constructions

The verb *hàj* ‘give’ may appear as either a main causative verb, or a secondary verb in expressions with other causatives.

4.6.1 ‘Give’ as causative verb

The verb *hàj* ‘give’ serves as a general interpersonal causative, loosely equivalent in various contexts to *let*, *have*, *get*.

- (97) *phān kɔɔ bɔɔ hàj \emptyset pǎj*
 3.HON LNK NEG give go
 ‘He wouldn’t let me go.’ (332.2)
- (98) *khán vāa sán mǎn-'hǎn sāv dǎd / hàj phān máa phōp \emptyset*
 if say thus tomorrow morning PCL give 3.HON come meet
 ‘Then in that case, tomorrow morning, y’hear! Have them come and meet us.’ (79.1)

- (99) *khɔ̀ɔj bɔ̀ɔ hət̚j thūk hət̚j n̄âak n̄ăng*
 1P NEG give wretched give difficult anything
 'I won't let you be poor or have any difficulties.' (868.11)
- (100) *ø hət̚j nɔ̀ɔj pǎj s̄u ləw d̄ē*
 give N. go buy liquor PCL ('please')
 'Get Noy to buy some liquor, please.'

Here, we may point to yet another case of possible structural ambiguity, relating to the referential disjunction function of *hət̚j* 'give'. Consider (95), repeated here from above:

- (95) *háv_i jàak hət̚j ø_j sɔ̀n*
 1P want give fight
 'I want them to fight.' (142.13)

In the translation here (construed given the actual context), *hət̚j* 'give' performs a referential disjunction function, reversing control of the main complement-taking predicate *jàak* 'want' (conceptually, predicating a handover of control; see Newman 1996). This reading would see *sɔ̀n* 'fight' as the primary lower verb, subordinate to *jàak* 'want', while *hət̚j* 'give' plays a structural function of switching reference. An alternative reading, however, would see *hət̚j* 'give' with a causative function, as the main lower verb, with *sɔ̀n* 'fight', further embedded in subordination to it. Thus, (95) could mean 'I want to get them to fight'.

The following example shows the same kind of ambiguity:

- (101) *bɔ̀ɔ jàak hət̚j nɔ̀ɔng 'ɔ̀k càak vǎng*
 NEG want give y.sib exit from palace
 i. 'He didn't want his sister to leave the palace.'
 ii. 'He didn't want to let his sister leave the palace.' (160.9)

Thus, where we find the combination [*jàak* 'want' + *hət̚j* 'give'], it may not always be possible to clearly distinguish the two interpretations of *hət̚j* 'give', as a 'switch-reference' marker, or as a (subordinated) primary causative verb.

4.6.2 'Give' as secondary causative verb

The verb *hət̚j* 'give' may 'mark' other causative verbs, basically as V_2 in what appears to be a V-V compound:

- (102) *khɔ̀ɔj s̄ang-hət̚j khǎw pǎj*
 1P order-give 3P go
 'I ordered them to go.'
- (103) *láv khɔ̀ɔ-hət̚j khɔ̀ɔj k̄n-khaw*
 3P request-give 1P eat-rice
 'S/he requested that I eat.'

The following rephrasings of (102) and (103) reveal complications relating to the possibility (or not) of the causee appearing between the two verbs:

- (104) *khɔ̀ɔj s̄ang khǎw hət̚j paj*
 1P order 3P give go
 'I ordered them to go.'

- (105) **láaw khǎj khǎj həj kǐn-khàw*
 3P request IP give eat-rice
 (S/he requested that I eat.)

(Note, however, that (105) is acceptable with the meaning ‘S/he begged me, that I let her/him eat.’)

5 Discussion: structure of complex predicates

In this paper I have described some important grammatical roles of Lao verbs *ǎw* ‘take’ and *həj* ‘give’ in complex predicate constructions. Lao grammar exploits these basic predicates for certain structural functions relating to valency-increasing operations (i.e. addition of an extra argument entailed by causative, instrumental, or benefactive expressions), or syntactic permutations related to discourse status of arguments (as in the ‘pretransitive’ construction). I have not discussed theoretical issues relating to these kinds of structures, but I hope my data and discussion may contribute to current research in syntactic theory on argument structure, complex predicates, and other areas of interest in syntactic research (see Manning 1996; Alsina et al. eds 1997; Andrews & Manning 1998; and references therein).

We may now briefly review the three basic argument structure arrangements suggested above for the various *ǎw* ‘take’ and *həj* ‘give’ constructions in Lao (repeated with original numbers from above):

- (21)
$$\overbrace{V_1 < NP_1, NP_2, V_2 < -, NP_3 >>}$$

- (28)
$$V_1 < NP_1, NP_2, V_2 < -, (NP_3) >>$$

- (58)
$$\overbrace{V_1 < NP_1, NP_2, V_2 < -, - >>}$$

In each case two verbs (V_1 and V_2) combine to form a single clause (or complex predicate), and their respective argument structures (in the sense of Manning 1996) merge, whereby at least one argument is shared between the two, and whereby the shared argument may or may not have the same prominence with respect to each of the two verbs. Structures (21, 28, 58) are logical possibilities. The schema in (21) covers instrumental constructions (§3.2.1), the quasi-instrumental constructions discussed in §3.2.3, and purposive constructions (see §3.4.2), where the most prominent argument of V_1 is also the most prominent argument of V_2 , and the object of V_1 is not a core argument of V_2 .

The schema in (28) covers causatives (*ǎw*-causatives, §3.2.2; *həj*-causatives), where the less prominent argument of V_1 is the most prominent argument of V_2 .

The schema in (58) covers pretransitives (§3.4), as well as various three-place predicates such as 'give' and 'put' (§3.3), in which the prominence relations of the arguments are preserved across both V_1 and V_2 (see also the schema in (76), above).

These preliminary observations leave open a number of questions which remain to be resolved in settling on an analysis of the argument structure of these and other (both complex and simple) predicates in Lao. Further work needs to be done on the role of preverbal directional particles *pǎj* 'go' and *máa* 'come' (so often appearing in 'ǎw 'take' constructions, especially), as well as on the semantics of the constructions, and the discourse conditions governing their usage. The question of constituent structure (probably involving conjoined VPs under a higher VP node) is another area for further research. It may also be worth considering an analysis which posits (ready-made) construction types rather than strings which speakers assemble in novel ways, given the high level of idiomaticity of these constructions in spoken Lao.

References

- Alsina, Alex, Joan Bresnan and Peter Sells, eds, 1997, *Complex predicates*. Stanford: CSLI Publications.
- Andrews, Avery and Christopher D. Manning, 1998, *Complex predicates and information spreading in LFG*. Stanford: CSLI Publications.
- Chao, Yuen Ren, 1968, *A grammar of spoken Chinese*. Berkeley: University of California Press.
- DeFrancis, John, 1984, *The Chinese language: fact and fantasy*. Honolulu: University of Hawai'i Press.
- Dixon, R.M.W., 1994, *Ergativity*. Cambridge: Cambridge University Press.
- Durie, Mark, 1985, *A grammar of Acehnese, on the basis of a dialect of North Aceh*. Dordrecht: Foris.
- 1997, Grammatical structures in verb serialization. In Alsina et al., eds 1997:289–354.
- Enfield, N.J., 1999, Lao as a national language. In Evans, ed. 1999:258–290.
- Evans, Grant, ed., 1999, *Laos: culture and society*. Chiang Mai: Silkworm Press.
- Fillmore, Charles, 1968, The case for case. In Emmon Bach and Robert Harms, eds *Universals in linguistic theory*, 1–90. New York: Holt, Rinehart and Wilson.
- Givón, Talmy, 1984, *Syntax: a functional-typological introduction*. Amsterdam/ Philadelphia: John Benjamins.
- Hopper, Paul and Sandra Thompson, 1980, Transitivity in grammar and discourse. *Language* 56:251–299.
- Jagacinski, Ngampit, 1987, The Tai Lue of Xipsongbanna in China's Yunnan Province: description and a study of the OV order in the 'AU construction. PhD thesis, Ohio State University. Ann Arbor: University Microfilms International.
- Kratochvil, Paul, 1968, *The Chinese language today: features of an emerging standard*. London: Hutchison University Library.
- Langacker, Ronald W., 1991, *Foundations of cognitive grammar*. (Volume 2: *Descriptive application*). Palo Alto: Stanford University Press.
- Li, Charles N. and Sandra A. Thompson, 1981, *Mandarin Chinese: a functional reference grammar*. Berkeley: University of California Press.

- Lord, Carol, 1993, *Historical change in serial verb constructions*. Amsterdam/Philadelphia: John Benjamins.
- Manning, Christopher D., 1996, *Ergativity: argument structure and grammatical relations*. Stanford: CSLI Publications.
- Newman, John, 1996, *Give: a cognitive linguistic study*. Berlin/New York: Mouton de Gruyter.
- Norman, Jerry, 1988, *Chinese*. Cambridge: Cambridge University Press.

3 Vietnamese verbal reduplication

SOPHANA SRICHAMPA

1 Introduction

Reduplication is a morphological process in which all or part of a syllable is repeated. It can be regarded as a general process of affixation in which the entire base word or some part of it is affixed to itself to form one or a series of new words. These new words usually have the same syntactic classification as the base; their predicatable meanings are related to but deviate from the meaning of the base. Reduplication is widely found among the languages of East and Southeast Asia, such as Mandarin (Li & Thompson 1981), Cantonese (Bauer & Benedict 1997), Thai (Laksaneeyanawin 1982) and Vietnamese.¹ In this paper I describe the major patterns of reduplication that are associated with Vietnamese verbs (including stative verbs).

Previous studies on Vietnamese reduplication have generally described the possible reduplicated classes but they did not pinpoint any specific class of the language. Due to the fact that the verb occupies the major position in the structure of the language, I will describe its behaviour with respect to its reduplication. Vietnamese verbal reduplication expresses a range of meanings, including consecutive, emphatic, repetitive, augmentative and diminutive meanings. In this study Vietnamese reduplication has been divided into two types according to the form of the original morpheme, namely, full reduplication and partial reduplication.

2 Vietnamese syllable and phonology

Since the Vietnamese syllable is both a phonological and morphological unit of analysis, the structure of the syllable is a matter of relevance to the study of reduplication. The structure of the Vietnamese syllable can be represented in the following formula:

$$\begin{array}{c} T \\ (C_1) (w) + V + (C_2) \end{array}$$

¹ For the Vietnamese data I use the following sources: Hoàng Văn Hành (1994); Nhan Ngô Thanh (1984) and Thompson (1967).

The syllable begins with the onset or initial consonant C_1 which is an optional element. The remainder of the syllable is the final or rhyme which comprises the obligatory main vowel V and tone T . The rhyme may be optionally preceded by the medial labial glide /w/ and closed by an optional final consonant C_2 .

The following four tables list the phonemes of the northern or Hanoi variety of Vietnamese that may occur in the various slots of the syllable: Table 1 lists the initial consonants; Table 2 lists the final consonants; Table 3 lists the vowels; and Table 4 lists the tones. The sound segments have been transcribed in IPA (in bold) and the corresponding Vietnamese orthography (in italics). The Vietnamese names of the six Vietnamese tones and the phonetic descriptions of their tone contours are included in Table 4.

In all the examples of Vietnamese reduplication listed and discussed below, the Vietnamese orthographic forms are given first and are then followed by the corresponding phonemic transcription. In the phonemically transcribed material the raised numbers 1 to 6 at the end of the syllable indicate the tone categories of the syllables and these numbers correspond to the numbered tones listed in Table 4.

Table 1: Vietnamese initial consonant phonemes (**bold**) and corresponding orthographic symbols (*italics*)

	Bilabial	Labio-dental	Alveolar	Palatal	Velar	Glottal
Stops			t t th th	c ch tr	k c k	
Implosives	ɓ b		ɗ ɗ			
Fricatives		v v f ph	z r d s s x		ɣ g gh x kh	h h
Nasals	m m		n n	ɲ nh	ŋ ng ngh	
Approximants	w w		l l	j r		

Table 2: Vietnamese final consonant phonemes (**bold**) and corresponding orthographic symbols (*italics*)

	Bilabial	Alveolar	Palatal	Velar
Stops	p p	t t	c ch	k c
Nasals	m m	n n	ɲ nh	ŋ ng
Approximants	w u o		j i y	

Table 3: Vietnamese vowel phonemes (**bold**) and corresponding orthographic symbols (*italics*).

	Front	Back	Diphthongs
High	i i	ɯ ʊ u u	<i>iə (ia, lê-) uə (ɯa, ɯə-) uə (ua, uô-)</i>
Higher mid	e ê	ə ɔ o ô	
Lower mid	ɛ e	ʌ â ɔ o	
Low	a ă	a: a	

Table 4: Names of Vietnamese tones and phonetic descriptions of their tone contours

Vietnamese tone name	Phonetic description of tone contour
1. <i>ngang</i>	Mid-high level
2. <i>huyền</i>	Low falling
3. <i>hỏi</i>	Mid falling-rising
4. <i>ngã</i>	Mid to high rising with glottal stop
5. <i>sắc</i>	Mid-high rising
6. <i>nặng</i>	Low falling with glottal stop

3 Full reduplication

The entire phonological form of a stem is repeated in the reduplicated affix. We may recognise two types of full reduplication: monosyllabic and polysyllabic.

3.1 Monosyllabic full reduplication

The entire syllable is repeated in full. Many reduplications express repetitive or diminutive meanings as indicated in the following examples.

3.1.1 Repetitive meanings

- (1) *gằm* /ɣam²/ 'wear a sullen look' => *gằm gằm* /ɣam ɣam/
- (2) *bùng* /buŋ²/ 'flare up' => *bùng bùng* /buŋ² buŋ²/
- (3) *khum* /xum¹/ 'bend, bow' => *khum khum* /xum¹ xum¹/
- (4) *mờ* /mə²/ 'grow dim, dull' => *mờ mờ* /mə¹ mə¹/
- (5) *nghiêng* /ŋiəŋ¹/ 'be inclined (to)' => *nghiêng nghiêng* /ŋiəŋ¹ ŋiəŋ¹/
- (6) *nhông* /ŋoŋ¹/ 'wander, roam' => *nhông nhông* /ŋoŋ¹ ŋoŋ¹/
- (7) *ọc* /ɔk⁶/ 'spurt, flow' => *ọc ọc* /ɔk⁶ ɔk⁶/
- (8) *rành* /za:ŋ²/ 'know well' => *rành rành* /za:ŋ² za:ŋ²/

3.1.2 Diminutive meanings

- (9) *quèn* /kwən²/ 'indifferent'
=> *quèn quèn* /kwən² kwən²/ 'rather indifferent, somewhat poor'
- (10) *rợn* /zən⁶/ 'shiver with fear'
=> *rợn rợn* /zən⁶ zən⁶/ 'to shiver slightly with fear'

3.2 Disyllabic full reduplication

This type of reduplication also mainly expresses repetition. Reduplication of the two-syllable word takes the form AABB as follows:

- (11) *biếng nhác* /biəŋŋ ɲa:k⁵/ ‘lazy, idle’
=> *biếng biếng nhác nhác* /biəŋŋ⁵ biəŋŋ⁵ ɲa:k⁵ ɲa:k⁵/
- (12) *đi lại* / ‘to come’
=> *đi đi lại lại* /đi¹ đi¹ lai⁶ lai⁶/ ‘to walk back and forth’
- (13) *hăm hở* /ham¹ hə³/ ‘be zealous and enthusiastic’
=> *hăm hăm hở hở* /ham¹ ham¹ hə³ hə³/
- (14) *lầm lì* /lɒm² li²/ ‘be taciturn, say little’
=> *lầm lầm lì lì* /lɒm² lɒm² li² li²/

4 Partial reduplication

This type of reduplication takes two forms: (1) reduplication of the initial consonant as in *ho he* /hə¹ he¹/ ‘act against’ and (2) reduplication of the rhyme as in *bịn rịn* /bin⁶ zin⁶/ ‘be loath to part with’. In the disyllabic form the tone changes on both the first and the second syllables from low falling tone to mid rising tone, as in:

- (15) *bần thần* /bɒn² thɒn²/ ‘haggard, worried’
=> *bần **thần** bần thần* /bɒn³ thɒn³ bɒn² thɒn²/

4.1 Monosyllabic partial reduplication

There are various possibilities of monosyllabic partial reduplication relating to the syllable structure — six patterns can be grouped into three subtypes: (i) to (iii) rhyming (same vowel, same final and sometimes same tone); (iv) chiming (same initial, but different vowel, final, or tone); and (v) rhyming and chiming (with different tone). The reduplication may either precede or follow the stem.

a. Rhyming

i. Reduplication of rhyme with loss of initial consonant.

The initial consonant of the stem is lost and the rhyme is reduplicated as in:

- (16) *ăn **năn*** /an¹ nan¹/ ‘show repentance’
(17) *áy **náy*** /a⁵ ɲa⁵/ ‘uneasy’

The reduplicated syllable comes first and the original stem comes second. This type occurs with relatively few verbs.

ii. Change of initial consonant

The initial consonant in the reduplicated form is different from that of the stem (in bold) as in:

- (18) *rủn* /zun³/ ‘be flabby, be faint’
=> *bủn **rủn*** /bun³ zun³/ ‘flagging, flaccid’

- (19) *nhăn* /ɲan²/ ‘eat with one’s front teeth’
 => *căn nhăn* /kan² ɲan²/ ‘grumble, grunt’

The reduplicated syllable can occur in either the first or second position as in:

- | FIRST | SECOND |
|--|--|
| (20) <i>bủn rủn</i> /bun ³ zun ³ /
‘flagging, flaccid’ | (21) <i>chóáng váng</i> /cwa:ŋ ⁵ va:ŋ ⁵ /
‘dizzy, dazed, giddy’ |
| (22) <i>căn nhăn</i> /kan ² ɲan ² /
‘grumble, grunt repeatedly’ | (23) <i>chóí vớí</i> /cəj ⁵ vəj ⁵ /
‘reach up one’s hands repeatedly’ |

This type only occurs with some verbs.

iii. Change of initial consonant and tone

The initial consonant and the tone are different from that of the stem (transcribed in bold). In reduplication of open syllables the *ngang* tone occurs with *huyền* tone and the *huyền* tone occurs with the *sắc* tone as indicated in the following examples.

- (24) *chơi bời* /cəj¹ bəj²/ ‘be a playboy, lead a gay life’
 (25) *biu điu* /biu² ziw⁵/ ‘wrapped up in’
 (26) *bù khú* /bu² xu⁵/ ‘have heart-to-heart talk, talk together’

In reduplication of closed syllables the *nặng* tone occurs with the *sắc* tone and the *ngang* tone occurs with the *hỏi* tone as in the following examples:

- (27) *sạch bách* /sa:c⁶ ba:c/ ‘very clean’
 (28) *hưng hửng* /huŋ¹ huŋ³/ ‘be struck with disappointment’

The rhyme of either the first or second syllable may be reduplicated and the initial is different as follows:

- | FIRST | SECOND |
|---|--|
| (29) <i>biu điu</i> /biw ¹ ziw ⁵ /
‘wrapped up in’ | (30) <i>bù khú</i> /bu ² xu ⁵ /
‘have heart-to-heart talk, talk together’ |
| (31) <i>căn đăn</i> /kan ¹ zan ⁶ /
‘recommend’ | (32) <i>chơi bời</i> /cəj ¹ bəj ² /
‘be friends, indulge in play’ |

This type only occurs with some verbs.

b. Chiming

iv. (1) Reduplication of initial consonant and change of rhyme

The vowel in the reduplicated form is different from that of the stem (transcribed in bold) as follows:

- (33) *ngu* /ŋu¹/ ‘stupid, dull’
 => *ngu ngơ* /ŋu¹ ŋə¹/ ‘naively silly’ (back vowels)
 (34) *chúm* /cum⁵/ ‘purse, round (lips)’
 => *chúm chím* /cum⁵ cim⁵/ ‘slightly protruding’ (high vowels)

- (35) *đu* /đu¹/ 'swing'
=> **đu** *đu* /đu¹ đu¹/ 'swing' (high vowels)
- (36) *móp* /móp⁵/ 'sunken'
=> **móp** *mép* /móp⁵ mep⁵/ 'badly dented' (mid vowels)

The reduplicated syllable occurs in either the first or second position as follows:

- | FIRST | SECOND |
|--|---|
| (37) <i>bi</i> bô /bi ¹ bo ¹ /
'lisp (baby)' | (38) bet <i>bát</i> /bet ⁵ bat ⁵ /
'dog-tired, exhausted' |
| (39) <i>thuế</i> thỏa /thwe ³ thwa:3/
'to be satisfied' | (40) chúm <i>chím</i> /cum ⁵ cim ⁵ /
'open slightly' |

This type occurs with some verbs.

iv. (2) Change of rhyme and tone

The vowel and tone are different from that of the stem (transcribed in bold) as in the following examples.

- (41) **cãi** *cọ* /kaj⁴ kɔ⁶/ 'have an argument with'
- (42) **che** *chở* /ce¹ cə³/ 'give cover to protect'
- (43) **cò** *cử* /kɔ² ku³/ 'wheeze'

The reduplicated syllable occurs only in the first position. This type of reduplication seems to occur with relatively few verbs.

The tone and the final consonant are different from that of the stem (in bold), as in:

- | FIRST | SECOND |
|---|---|
| (44) <i>còm</i> cộ /kom ² kop ⁶ /
'unwieldly' | (45) bàn <i>bạc</i> /ba:n ² bak ⁶ /
'discuss' |
| (46) <i>hoang</i> hoác /hwa:ŋ ¹ hw:k ⁵ / 'open wide' | |

The reduplicated syllable can occur either in the first or second position but mainly occurs in the first position. This type of reduplication occurs with many verbs. The reduplicated words of this type are typically onomatopoeic expressions, while others are diminutive terms.

c. Rhyming and chiming

v. Change of tone

In the reduplicated form, which can either precede or follow the stem, the tone is different from that of the stem (transcribed in bold).

The reduplicated form occurs in the first position as in:

- (47) *liệu* /liəw⁶/ 'use the suitable thing'
=> *liệu* **liệu** /liəw² liəw⁶/ (repetitive meaning)

The prefix tone is *huyền* instead of *nặng*.

- (48) *khó* /xɔ⁵/ ‘hard, difficult’
=> *kho khó* /xɔ¹ xɔ⁵/ (repetitive meaning)
The prefix tone is *ngang* instead of *sắc*.
- (49) *deo* /zɛw³/ ‘flexible, malleable’
=> *deo dẻo* /zɛw¹ zɛw³/ ‘rather flexible, rather malleable’
The prefix tone is *ngang* instead of *hỏi*.
- The reduplicated form occurs in the second position as in:
- (50) *khít* /xít⁵/ ‘fit well’ => *khít khít* /xit⁵ xit⁶/ ‘close-fitting’
The tone of the suffix is *nặng* instead of *sắc*.
- (51) *ngờ* /ŋə²/ ‘doubt, suspect’ => *ngờ ngờ* /ŋə² ŋə⁶/ (repetitive meaning)
The tone of the suffix is *nặng* instead of *huyền*.
- (52) *bừ* /bu⁵/ ‘cloyed, satiated’ => *bừ bừ* /bu⁵ bu⁶/ (emphatic meaning)
The suffix tone is changed from *sắc* to *nặng*.

In my data the four tones *ngang*, *huyền*, *sắc* and *ngã* occur with stems. Table 5 below presents the summary of the changes of the tones in the reduplicated forms:

Table 5: Changes of tones in reduplicated forms

1. <i>ngang</i> tone	=>	<i>sắc, hỏi, ngã, nặng</i> tones / # _____
	=>	<i>hỏi, ngã</i> tones / _____ final nasal
	=>	<i>sắc</i> tone / _____ final stop and nasal
2. <i>huyền</i> tone	=>	<i>nặng</i> / # _____
	=>	<i>ngã, nặng</i> tones / final nasal
3. <i>sắc</i> tone	=>	<i>huyền</i> tone / # _____
	=>	<i>huyền</i> tone / _____ final nasal
4. <i>ngã</i> tone	=>	<i>huyền</i> tone / # _____

The reduplicated syllable occurs in either the first or second position but reduplication in the first position is more frequent. This is also the most frequent type of verbal reduplication. Many such reduplicated words convey a diminutive meaning.

4.2 Disyllabic partial reduplication

Where the verb stem has two syllables, there are three subtypes of partial reduplication. In two of these types the resulting reduplicated form has four syllables and generally follows the ABAB and ABCD patterns, with some differences parallel to those found in monosyllabic reduplication, as in rhyming and chiming, in one or both of the reduplicated syllables. In type (iii), the resulting form has three syllables in the ABB pattern; for this type the partially reduplicated middle syllable is different from the basic B syllable stem. The ABB reduplication type is relatively infrequent.

Of the various types below, the basic stem comes second in most types; the modified reduplicated form precedes in four-syllable types in the ABAB and ABCD patterns, and the modified reduplicated form also precedes in both three-syllable types in the ABB pattern.

a. ABAB modified reduplications

i. Change of rhyme and tone

All four syllables share the same initial consonant but the rhyme of the second syllable changes to vowel /a:/ and its tone changes to *ngang*.

- (53) *bấp bênh* /bap⁵ beŋ¹/ ‘unsettle, staggering’
=> *bấp **ba** bấp bênh* /bap⁵ ba:¹ bap⁵ beŋ¹/ (augmentative meaning)
- (54) *chấp chới* /cap⁵ cəj⁵/ ‘fly with a rolling motion’
=> *chấp **cha** chấp chới* /cap⁵ ca:¹ cap⁵ cəj⁵/ (consecutive meaning)
- (55) *hồn hên* /hon³ hen³/ ‘panting’
=> *hồn **ha** hồn hên* /hon³ ha:¹ hon³ hen³/ ‘to pant heavily’

The tone of the Ba syllable (B syllable with a vowel) is determined by the tone of the second syllable. When the B syllable has the *ngang*, *sắc* or *hỏi* tone, this syllable carries the *ngang* tone. This is the most frequent type of ABAB reduplication and the meaning is augmentative. The Bà (B syllable with à vowel) type in (ii) is less frequent.

The second syllable has the same initial as the second syllable plus the vowel /a/ with *huyền* tone (transcribed in bold).

- (56) *hì hục* /hi² huk⁶/ ‘be absorbed in’
=> *hì **hà** hì hục* /hi² ha:² hi² huk⁶/ ‘be completely engrossed in one’s work’
- (57) *khập khiễng* /xap⁶ xiəŋ⁴/ ‘limping’
=> *khập **khà** khập khiễng* /xap⁶ xa:² xap⁶ xiəŋ⁵/ (augmentative meaning)
- (58) *lảm cằm* /lam³ kəm³/ ‘forgetful, in one’s dotage’
=> *lảm **cà** lảm cằm* /lam³ ka:² lam³ kəm³/ (augmentative meaning)
- (59) *gập ghềnh* /ɣap⁶ ɣep²/ ‘rough, bumpy’
=> *gập **gà** gập ghềnh* /ɣap⁶ ɣa:² ɣap⁶ ɣep²/ ‘very rough’

The above examples include those where the B syllable has *ngã*, *huyền*, *hỏi*, or *nặng* tones. We may note that where the B syllable carries the *hỏi* tone, both type (i) and type (ii) can occur.

The vowel and the final consonant of the last syllable of the reduplicated form (transcribed in bold) are different from those of the second syllable of the stem as in:

- (60) *bộc tuệch* /bok⁶ twec⁶/ ‘ingenious’
=> *bộc tuệch bộc **toạc*** /bok⁵ twec⁵ bok⁵ twa:k⁵/ ‘very ingenious’
- (61) *tâm bậy* /tam² bəj⁶/ ‘perverse, wrong’
=> *tâm bậy tâm **bạ*** /tam² bəj⁶ tam⁶ ba:⁶/ ‘absolutely wrong’

The vowel and tone of the second syllable (transcribed in bold) are different from those of the stem as in:

- (62) *hốt hải* /hət⁵ haj³/ ‘panic’
=> *hốt **hơ** hốt hải* /hət⁵ hə¹ hət⁵ haj³/ ‘panicky’

The first two syllables (transcribed in bold) carry a different tone from that of the stem as in:

- (63) *bần thần* /bʌn² thʌn²/ ‘haggard, worried’
=> ***bần thần*** *bần thần* /bʌn³ thʌn³ bʌn² thʌn²/ (repetitive meaning)
- (64) *bồi hồi* /boj² hoj²/ ‘fret, be fretty’
=> ***bồi hồi*** *bồi hồi* /boj³ hoj³ boj² hoj²/ ‘to fret intensely’

The tone and final consonant of the second syllable (transcribed in bold) are different from those of the stem as in:

- (65) *bắt nét* /bat⁵ net⁵/ ‘break in by strictly finding fault with’
=> ***bắt ne*** *bắt nét* /bat³ nɛ³ bat² net²/ (augmentative meaning)
- (66) *tí toét* /ti⁵ twet⁵/ ‘grin and speak continually’
=> ***tí toe*** *tí toét* /ti⁵ twe¹ ti⁵ twet⁵/ (emphatic meaning)

b. ABCD reduplication

ii. Change of rhyme

The vowels of both syllables of the reduplicated form (transcribed in bold) are different from those of the stem as in:

- (67) *bông lông* /boŋ¹ loŋ¹/ ‘aimless’
=> *bông lông* ***bang lang*** /boŋ¹ loŋ¹ ba:ŋ¹ la:ŋ¹/ ‘absolutely aimless’

This type and the following type (iv) are the least frequent types of reduplication for disyllabic verbs.

c. ABB reduplication

iii. Change of rhyme and/or tone of the second syllable

In this type of reduplication the final consonant of the rhyme may change from a stop to the homorganic nasal as in example (68), or the tone may change from 6 to 3 as in example (69):

- (68) *im thít* /im¹ thit⁵/ ‘be completely silent’
=> *im* ***thin thít*** /im¹ thin¹ thit⁵/ ‘to keep as silent as a grave’
- (69) *khắm lặm* /xam⁵ lam⁶/ ‘intolerably offensive’
=> *khắm* ***lắm lặm*** /xam⁵ lam³ lam⁶/ (augmentative meaning)

5 Summary

There are two types of reduplication in Vietnamese: full reduplication and partial reduplication. Full reduplication can occur with both monosyllabic and disyllabic words for which there is exact reduplication of the syllable. The disyllabic full reduplication follows the AABB pattern. Partial reduplication can occur with both monosyllabic and disyllabic words. Monosyllabic partial reduplication has the following six sub-types:

- a. Rhyming (same vowel, same final and sometimes same tone)
 - i. Reduplication of rhyme with loss of consonant
 - ii. Change of initial consonant
 - iii. Change of initial consonant and tone
- b. Chiming (same initial, but different vowel, final, or tone)
 - iv. Reduplication of initial consonant and change of rhyme
- c. Rhyming and chiming (with different tone)
 - v. Change of tone

The disyllabic partial reduplication has the following three sub-types:

- a. ABAB reduplication
 - i. Change of rhyme and tone
 - Replace the vowel and tone of the second syllable with /a/.
 - Replace the vowel and tone of the second syllable with /à/.
 - Change the vowel in the third and fourth syllables.
 - Change the vowel and final of the fourth syllable.
 - Change the vowel and tone of the second syllable.
 - Change the tone of first and second syllables from *huyền* to *hỏi* tone.
- b. ABCD reduplication
 - ii. Change of rhyme
- c. ABB reduplication
 - iii. Change of rhyme and tone of the second syllable

As we have observed from the examples presented in this paper, Vietnamese reduplication conveys augmentative, repetitive, consecutive, and emphatic meanings. Chiming reduplication has an onomatopoeic function of imitating sounds in the natural world. Some patterns of Vietnamese reduplication resemble those that are found in other East and Southeast Asian languages, such as Mandarin, Cantonese, and Thai.

References

- Bauer, Robert S. and Paul K. Benedict, 1997, *Modern Cantonese phonology*. Trends in Linguistics Studies and Monographs 102. Hong Kong student edition. Berlin and New York: Mouton de Gruyter.
- Hoàng Văn Hành, 1994, *Từ điển từ láy tiếng Việt* (Dictionary of reduplicated Vietnamese words). Hà Nội: Nhà xuất bản Giáo dục.
- Laksaneeyanawin, S., 1982, 'kham sám ?ùtca:ranawilât: ?ì:k ηε: nùη khǒ:η rabòp siəη lε? rabòp khwa:m mǎ:j naj phaisǎ: thaj' (Reduplication: phonology and semantic systems in Thai]. *Journal of Fine Arts*. Vol.1:11-34. Bangkok: Chulalongkorn University.
- Li, Charles N. and Sandra A. Thompson, 1981, *Mandarin Chinese: a functional reference grammar*. Berkeley, Los Angeles, London: University of California Press.

Nhan Ngô Thanh, 1984, The syllabeme and patterns of word formation in Vietnamese. PhD dissertation, New York University. Ann Arbor, Michigan: University Microfilms International.

Thompson, Laurence C., 1967, *A Vietnamese grammar*. Seattle and London: University of Washington Press.

II

Language classification

4 *Nùng An: origin of a species*

JEROLD A. EDMONDSON

Although the sound and lexical changes among the members of the Tai language branch are fairly regular, the devil of classification lies in the details that separate one language from another. Li Fang Kuei's classic *A handbook of comparative Tai* (1977), proposed one taxonomic system with three terms — Northern Tai (NT) versus Central Tai (CT) versus South-western Tai (SWT). This tripartite theory has garnered the allegiance of most Tai specialists. In the tripartite theory one essential feature is a well-established linguistic division of Tai that lies in the geographic north and thus I will call this approach the *NT-Hypothesis*.

In the summer of 2000 I was able to study the Nung An of Jingxi County in Guangxi Province, China. There are about 26,000 speakers of this language located there. A preliminary comparison shows that the language spoken in Jingxi and that of Hà Quảng District, Cao Bằng are nearly identical.

A smaller group of scholars, including most prominently André Haudricourt, insisted on a two-term system, dividing the Tai parent language first into Dioi or ?Yai (corresponding approximately, but not exactly, to Northern Tai) versus Tai Proper. Resting on his work on the Tai languages of Vietnam, I suppose Haudricourt may have been reluctant to accept the term 'northern Tai', as he was aware of many 'northern Tai' features among the Nùng languages of Vietnam, far to the south of the usual line of division between NT and CT languages, (Haudricourt 1956, 1960). Further suspicion about geography as a criterion emerged from the discovery of another 'NT language' much further to the south in Laos called *Saek* (Haudricourt 1963). The 'non-northern focus' of ?Yai in Tai in Haudricourt's work is a position we can call the *?Yai-Hypothesis*. While Haudricourt, to my knowledge, never framed the question so starkly, one can take two views, the NT-Hypothesis versus the ?Yai-Hypothesis, as in reality a discussion about the location of the original homeland of the Tai people and their migration thereafter. Crucial data bearing on the NT versus ?Yai Hypotheses is to be found mostly in Vietnam and this sort of data have often been meagre, recorded in Quốc Ngữ transcription, and, save for the EFEO survey of the 1930s, usually in the form of dictionaries of the language at one location. This paper will provide new data on several Tai languages all gathered in Cao Bằng Province in 1998.

There are still other problems about classifying Tai languages. Deciding whether a given Tai language is or is not a member of this or that branch is not a trivial question, as we have argued recently in writing about the Cao Lan language of northern Vietnam (see Gregerson &

Edmondson 2000). We stated in that paper that Cao Lan might constitute a case of a *tertium quid*, a kind of a third thing neither NT/Dioi nor CT/Tai Proper, in regard to:¹

its *phonology* — Cao Lan demonstrates some phonological rules regarded as exclusively Northern Tai (NT) and others that have been regarded as Central Tai (CT) features,

its *vocabulary* — Cao Lan shares common Tai vocabulary, but it also has uniquely Central Tai as well as Northern Tai forms too.

In this paper I will investigate the classification of another language of northern Vietnam called *Nùng An*, because Nùng An was among those regarded by Haudricourt (1960) and Strecker (1985) as ‘...exhibit[ing] a mixture of ?Yai and Tay-Nung traits’, (Strecker 1985: 481) and thus weakening the NT-Hypothesis. I will conclude that Nùng An is not a geographic anomaly at all, because it is not native to this area. Furthermore despite apparent similarities, Nùng An and Cao Lao are, in reality, unlike, because Nùng An appears to have imported its mixed contrasting colours from its progenitors in China. Thus, it did not evolve into a mixed language as a consequence of the migration process as perhaps happened in Cao Lan. Most importantly though, both of these are different from surrounding CT languages of Vietnam, such as Nùng Giang and Tày languages of Cao Bằng Province and NT languages of Vietnam, such as Giáy ?Yai).² Indeed, the evidence I have found suggests strong resemblance between Nùng An and Long’an, 108 km to the east in Guangxi Province, China, which I believe to have been the home of the Nùng An. Thus, Nùng An (and Cao Lan) data should be taken as support for ‘migration’ in the history of Tai and not for moving the NT/Dioi focus southward toward the Sino-Vietnam border. But, before we examine the linguistics and sociolinguistics of Nùng An, let us assemble a few pieces of information about Nùng An, the culture that speaks it, and the setting among other groups of the area.

1 Introduction

The Nùng An constitute a subgroup of the Nùng nationality, one of the larger of Vietnam’s 54 officially recognised minority ethnicities. The Nùng An population is concentrated in Cao Bằng Province at a location 37 km from Cao Bằng City in Quảng Hòa District at Phúc Sen Village and neighbouring Communes such as Đoàn Khôn, Quốc Phong, and Quốc Dân, where the population is 99 per cent Nùng An. Beth Nicolson (pers. comm.) reports that Nùng An speakers also live in Quảng Hòa District at Chí Thảo and Tụ Do as well as in Hà Quảng District at Nội Thôn Commune and in Ba Bê District at Cao Ché Commune, see map. In this region the Nùng An have a population of about 10,000. It is unknown how many Nùng An speakers there are altogether.³ The Nùng An are a distinctive and geographically concentrated group with their own native dress, customs, practices, and

¹ This research was supported by a National Endowment for the Humanities and the National Science Foundation grant to the author and Kenneth J. Gregerson: ‘Languages of the Vietnam-China borderlands.’

The author also wishes to thank Mr Lương Văn Đo, 44 years of age, who served as informant, and also my collaborators in gathering the data, Professor Nguyễn Văn Lợi, Professor Hoàng Văn Ma and Dr Thặng Văn Tò, all staff members of the Linguistics Institute of Vietnam and to Robert Bauer and Trần Thun for comments on this paper.

² It is to be noted though that Nùng Giang does show some surprising NT features lacking in Tày Bảo Lạc and Tày Trùng Khánh of Cao Bằng Province.

³ The Nùng had a total population of 696,300 according to the 1989 census.

language, which is not mutually intelligible with other kinds of Nùng. According to our informant's wife, who is studying the ethnic literature of the Nùng An as a part of a doctoral degree at Hanoi University, the Nùng An tell that they immigrated from Long'an in Guangxi Province some 8–10 generations ago and settled this territory. Today they continue their traditions of wet paddy agriculture using water buffaloes, planting trees to produce charcoal, shamanist medicine men, and using modified Chinese characters, similar to the Chừ Nôm of the Vietnamese.

2 Features of the Nùng An language

The following is intended to provide a preliminary treatment of the initials and tones of Nùng An. Like all Tai languages Nùng An has syllables with initials, rhymes (open syllables probably have vowels that are long as well as CVC and CVVC syllable shapes, which terminate in codas /-p -t -k -m -n -ŋ -i -u/), and a system of six tones. We will not provide details of the rhymes of Nùng An until we have gathered more data.

Initials in Nùng An:

p	<i>pəi¹</i>	'year'	<i>pət⁷</i>	'duck'
ph	<i>phu³</i>	'cloud'	<i>phən¹</i>	'rain'
ʔb	<i>ʔboot⁹</i>	'blind'	<i>ʔbau¹</i>	'leaf'
m	<i>ma¹</i>	'dog'	<i>miŋ²</i>	'name'
f	<i>fəi²</i>	'fire'	<i>fəən²</i>	'firewood'
v	<i>vən²</i>	'day'	<i>vaai²</i>	'buffalo'
t	<i>ta⁶</i>	'river'	<i>tə¹</i>	'wasp'
th	<i>tha¹</i>	'eye'	<i>thəu¹</i>	'louse, head'
ʔd	<i>ʔda¹</i>	'cloth sling for baby'	<i>ʔduut⁹</i>	'hot'
n	<i>nɔk⁸</i>	'outside'	<i>nən²</i>	'worm'
ʔ	<i>ʔaŋ¹</i>	'high'	<i>ʔom³</i>	'sour'
l	<i>laau²</i>	'fat'	<i>lət⁸</i>	'blood'
r	<i>raan²</i>	'house'	<i>rat²</i>	'mushroom'
s	<i>suu²</i>	'ox'	<i>saam²</i>	'indigo'
tɕ	<i>tɕuŋ³ tɕaau¹</i>	'spider'	<i>tɕai¹</i>	'plow'
ŋ	<i>ŋu³</i>	'grass'	<i>ŋoŋ²</i>	'mosquito'
j	<i>jaai²</i>	'mother's mother'	<i>jiu⁶</i>	'eagle'
k	<i>kuut⁹</i>	'fern'	<i>phak⁷ kaat⁹</i>	'mustard green'
kh	<i>kheen¹</i>	'arm'	<i>khim¹</i>	'needle'
ŋ	<i>ŋu²</i>	'snake'	<i>ŋən²</i>	'silver'
h	<i>ha²</i>	'cogongrass'	<i>hən²</i>	'person'
ʔ	<i>ʔaap⁹</i>	'to bathe'	<i>ʔooi³</i>	'sugarcane'
kw	<i>kwa¹</i>	'pumpkin'	<i>kwaan¹</i>	'sambar deer'

Tones of Nùng An. Nùng An has six tones on CV syllables and two tones each on CVVC₁ and CVC₁ syllables (where C₁ = /-p -t -k/).⁴

⁴ It will be convenient to transcribe Nùng An *Tone Values* in terms of Y.R. Chao's 'Scale-of-Five system' in which two or more numbers from 1–5 are intended to portray points on a pitch scale with 5 the highest

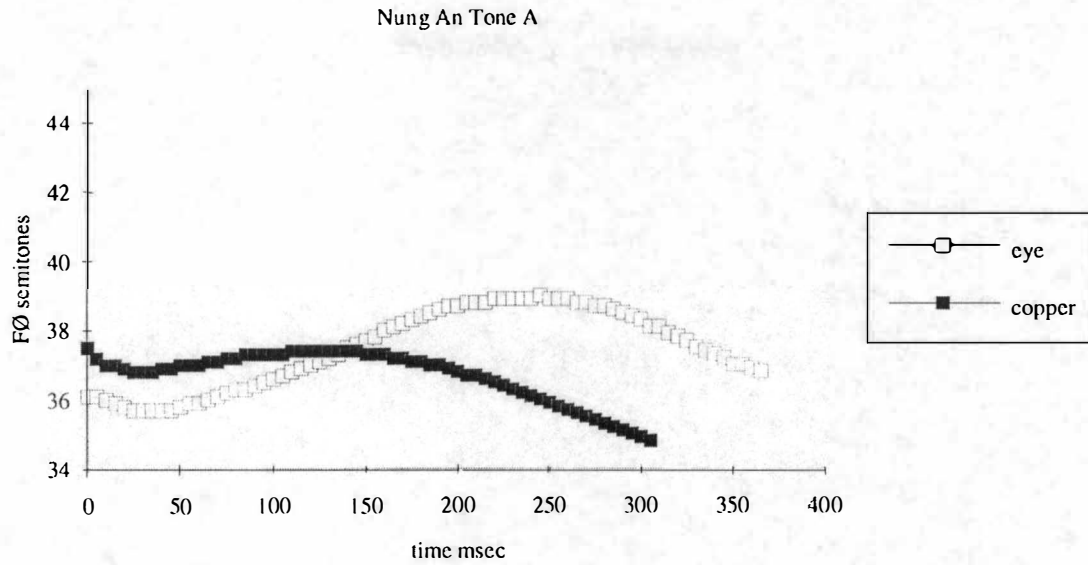


Figure 1: Nùng An Tone A illustrated with *tha*²¹³¹ 'eye' and *toor*²²¹ 'copper'

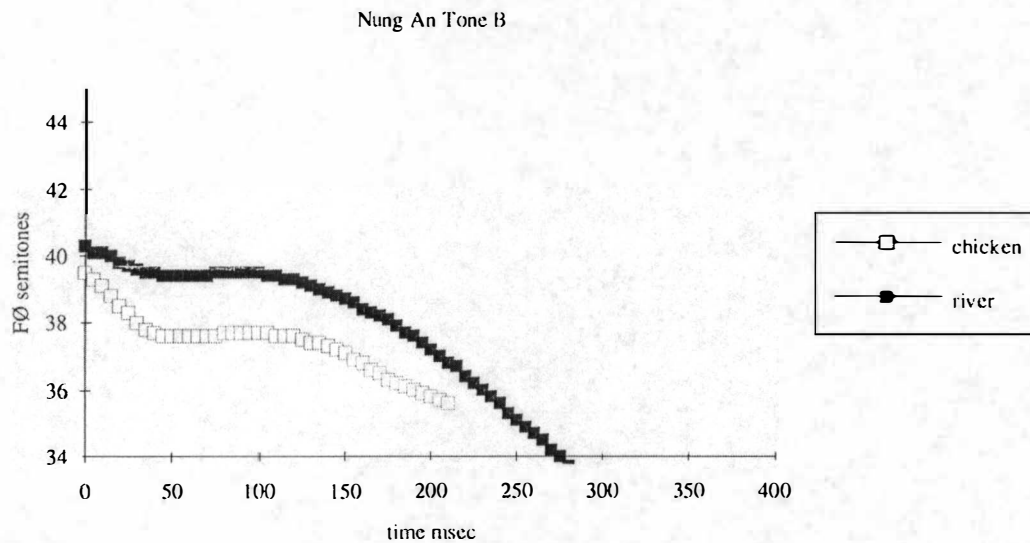


Figure 2: Nùng An Tone B illustrated with *kei*²¹ 'chicken' and *ta*³¹ 'river'

value and 1 the lowest. Thus, 35 will mean a mid-to-high rising tone. At other times I will use *Tone Categories*, which remain rather constant across members of the family even if the Tone Values change. There are thought to have been five original tones called *A*, *B*, *C*, *DL*, and *DS*. These four tones split, according to whether the initial consonant was voiced or voiceless into more tones. I will employ here the Chinese system, where Tone A became in daughter languages Tone 1 and Tone 2; Tone B became Tone 5 and Tone 6; Tone C became Tone 3 and Tone 4; Tone DL became Tone 9 and Tone 10; and Tone DS became Tone 7 and Tone 8. For more on this issue, see Edmondson and Solnit (1997), introduction.

The B tones in Nùng An are quite close in pitch trajectory. They are kept distinct not only by pitch but also by voice quality. The tone category 5 is accompanied by tight throat voice at the end of the syllable. Minimal pairs illustrating this difference are shown above, *ke²¹* ‘chicken’ versus *ta³¹* ‘river’, see Figure 2.

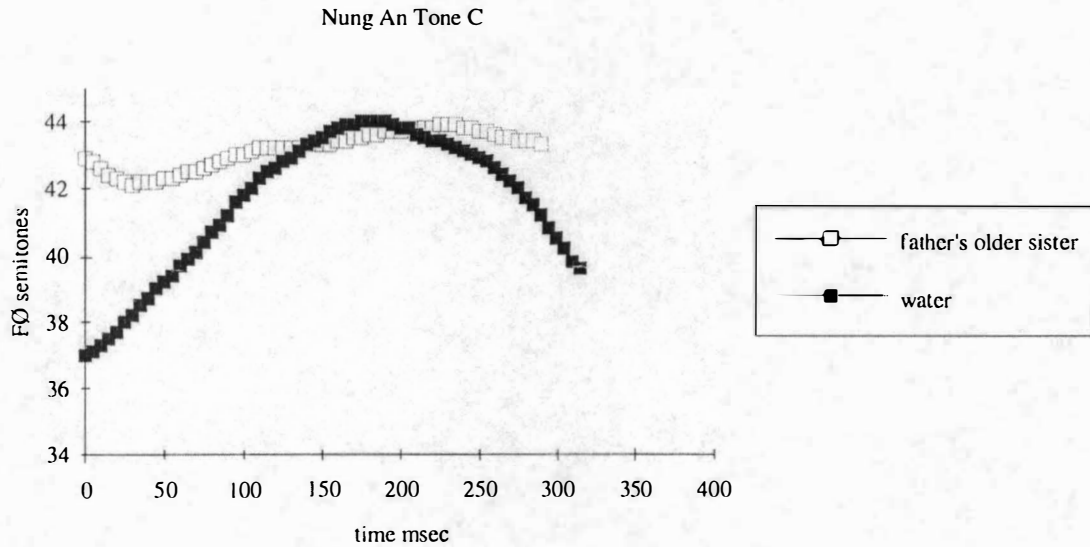


Figure 3: Nùng An Tone C illustrated with *pa⁴⁵* ‘aunt’ and *nem²⁵³* ‘water’

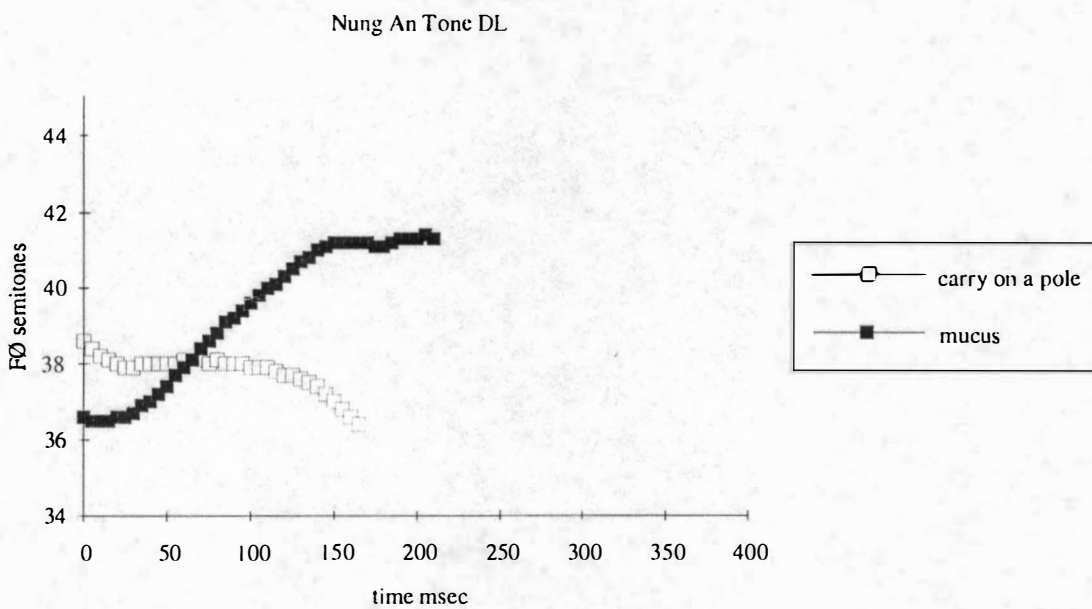


Figure 4: Nùng An Tone DL illustrated with *thaap³³* ‘to carry on the ends of a pole’ and *muuk²⁴* ‘mucus’

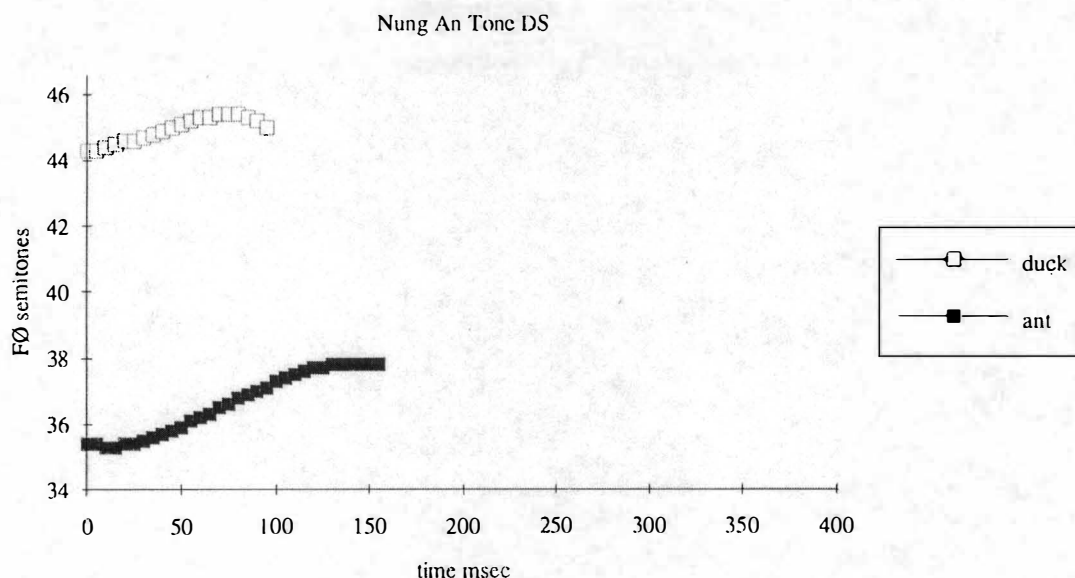


Figure 5: Nùng An Tone DS illustrated with $pət^{45}$ 'duck' and $mət^{23}$ 'ant'

There are also voice quality features in Nùng An. Tone 3 and Tone 4 end in glottal constriction that chokes off the voicing abruptly. This feature is not normally found in NT languages, but it is a prominent feature of CT languages. In addition to the constriction in Tone 3 and 4, Tone 5 shows a kind of slower change from modal to tight throat voice at the end of the syllable. This feature can be seen in Figure 6 below which is an inversely filtered airflow plot, for details on the recording and filtering of this sort of data (see Edmondson & Li 1994).

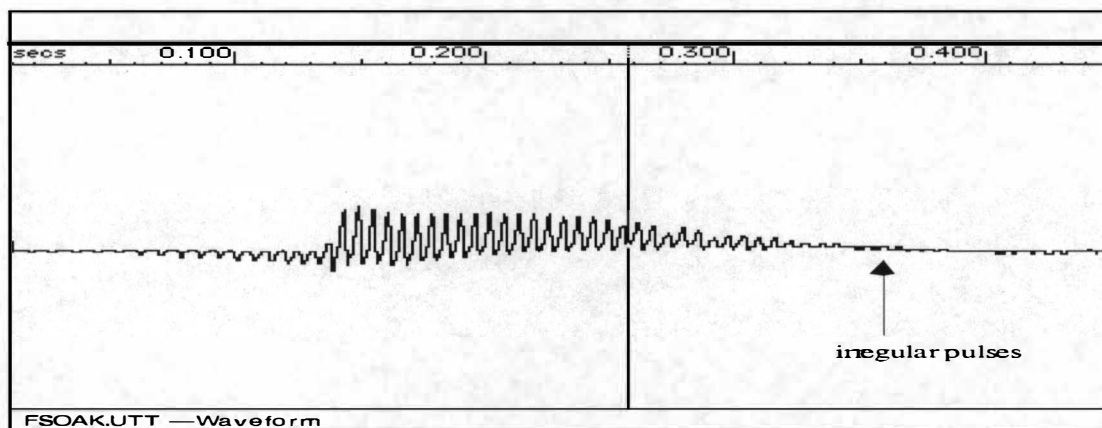


Figure 6: Inversely filtered airflow plot of the syllable ma^{21} 'to soak (rice)'

In Figure 6 particular attention should be paid to the irregular height (amplitude) and non-uniformity of distance (frequency) between glottal pulses at the end of the syllable.

3 Aspiration in Nùng An

With this preliminary look at some of the sound features of Nùng An, let us now turn to the comparative situation of Nùng An. Note first the sources of Nùng An initials. Some of the aspirated stops of the Tai parent language have merged in Nùng An with the plain stops, i.e. /*p- *ph-/ → p; /*t- *th-/ → t-; and /*k- *kh-/ → k-. Thus, Nùng An has:

Gloss	Nùng An	Giáy (NT)	Nùng Giang (CT)
'bamboo'	<i>mɛi⁴ pɛi⁵</i>	—	<i>pheu⁵</i>
'pond'	<i>təm²</i>	<i>tan²</i>	<i>təm²</i>
'male animal'	<i>tək⁸</i>	<i>tək⁸</i>	<i>tək⁸</i>
'chopsticks'	<i>tu⁶</i>	<i>tu⁶</i>	<i>thəu⁵</i>
'to kill'	<i>ka³</i>	<i>ka³</i>	<i>kha³</i>
'to ride'	<i>ke⁶</i>	<i>kwi⁶</i>	<i>khwe⁵</i>
'son-in-law'	<i>kwi¹</i>	<i>kwi¹</i>	<i>khi¹</i>
'maternal uncle'	<i>kɛu³</i>	—	<i>kjɛu³</i>
'body hair'	<i>phən¹</i>	<i>pun¹</i>	<i>khon¹</i>
'head hair'	<i>phom¹</i>	<i>pjim¹</i>	<i>phjam¹</i>
'taro'	<i>phuk⁹</i>	—	<i>phik⁹</i>
'vegetable'	<i>phak⁷</i>	<i>pjik⁷</i>	<i>phjak⁷</i>
'leg'	<i>kha¹</i>	<i>ka¹</i>	<i>kha¹</i>
'needle'	<i>khim¹</i>	—	<i>khjam¹</i>
'eggplant'	<i>khui¹</i>	<i>ku²</i>	<i>khi¹</i>

This rule is, of course, the well-known rule of the loss of aspiration in NT that Li (1977:12) noted. As the data above demonstrate, Giáy and Nùng Giang show typical features of NT and CT, respectively, whereas Nùng An is rather ambiguous, having about one-half NT and one-half CT features. As the data below show, the deaspiration processes bled NT of its aspirated stops. Later rules from CT sources appear to have 'fed' or repopulating the aspirated stop categories from original /*f-/ developing into *ph-*, whereas /*f-/ in Giáy often become *v-*.

'cloud'	<i>phui⁴</i>	<i>vui³</i>	<i>fa⁴</i>
'cotton'	<i>pha:i⁵</i>	<i>va:i⁵</i>	<i>pha:i³</i>
'rain'	<i>phən¹</i>	<i>pun¹</i>	<i>phən¹</i>
'hand palm'	<i>pha⁵</i>	<i>va⁵ fuŋ²</i>	<i>pha⁵ moŋ²</i>
'turtle, soft shell'	<i>phui¹</i>	—	<i>pha¹</i>
'dam'	<i>pha:i¹</i>	—	<i>pha:i¹</i>
'dream'	<i>pha¹ nən²</i>	<i>pan¹ huən²</i>	<i>pho¹ nən²</i>
'to sharpen'	<i>pən¹</i>		

Another source of aspiration came from original clusters involving the /*r-/ , such as **tr-* or **thr-*, which also developed into *th-*. See the list below.

4 Sonorants and sonorant clusters in Nùng An in comparative perspective

Strecker, supporting the claims of Haudricourt, points to the behaviour of old clustered elements **dl/r*, **l*, **r*, **hr*, **hw*, **thr* and finds that in NT languages there is a tendency for vocabulary to merge into *l-* or *n-* in NT languages, whereas CT tends to have separate reflexes for **r*, **l*, **w* and proto-clusters involving **-r-*, **-l-*, or **-w-*. In this case we will again use data from a NT language, Giáy of Lao Cai Province, Vietnam, and a CT language, Nùng Giang for comparison.

Gloss	Nùng An	Giáy	Nùng Giang
*l-			
'wind'	<i>lɛm²</i>	<i>ðum²</i>	<i>lɛm²</i>
'fingernail'	<i>lip⁸ mu²</i>	<i>zit⁸ fuŋ²</i>	<i>lap⁸ moŋ²</i>
'blood'	<i>lət⁸</i>	<i>luat⁸</i>	<i>luut⁸</i>
'pangolin'	<i>tɕit⁹ rən⁶</i>	<i>lin⁶</i>	<i>lin⁶</i>
*kl/r-			
'head'	<i>rau⁴</i>	<i>tɕau⁴</i>	
'far'	<i>tɕei¹</i>	<i>tɕai¹</i>	<i>kjwɛi¹</i>
'banana'	<i>ko³</i>	<i>tɕoi³</i>	<i>ku³</i>
*xr-			
'road'	<i>ran¹</i>	<i>ðan¹</i>	—
'rice pestle'	<i>rum¹</i>	—	—
'six'	<i>rok⁷</i>	<i>ðok⁷</i>	<i>khjok⁷</i>
'ear'	<i>ru¹</i>	<i>ðu¹</i>	<i>khjɛu¹</i>
*dl/r-			
'day after tomorrow'	<i>vən² ru²</i>	<i>ðu²</i>	<i>vən² ləi²</i>
'to steal'	<i>sək⁸</i>	<i>ðək⁸</i>	<i>lɛk⁸</i>
'to wash dishes'	<i>luŋ⁴</i>	<i>laŋ⁴</i>	
*r-			
'house'	<i>ra:n²</i>	<i>ða:n²</i>	<i>hru:n²</i>
'rice husk'	<i>ram²</i>	<i>ðam²</i>	<i>hrum²</i>
'dry field'	<i>reŋ⁶</i>	<i>ði⁶</i>	<i>hrei⁵</i>
'shout, to'	<i>ɾɔ:ŋ⁶</i>	—	<i>hro:ŋ⁶</i>
*hr-			
'bark'	<i>heu⁵</i>	<i>ðau⁵</i>	<i>heu⁵</i>
'mushroom'	<i>rat⁷</i>	<i>ðat⁷</i>	
*hw-			
'comb'	<i>roi¹</i>	<i>ðoi¹</i>	<i>vei¹</i>
*nl/r-			
'water'	<i>nɛm⁴</i>	<i>ðɛm⁴</i>	<i>nɛm⁴</i>
'bird'	<i>nok⁸</i>	<i>ðok⁸</i>	<i>nok⁸</i>
'outside'	<i>nɔ:k¹⁰</i>	<i>ðɔ:k¹⁰</i>	<i>nook¹⁰</i>

*thr-			
'carry on a pole'	<i>thaap⁹</i>	<i>ḍaap⁹</i>	<i>thaap⁹</i>
'tail'	<i>thuŋ¹</i>	<i>ḍuŋ¹</i>	<i>thuŋ¹</i>
'stone'	<i>thən¹</i>	<i>ḍin¹</i>	<i>thən¹</i>
'loom'	<i>thok⁷</i>		<i>thok⁷</i>
'louse'	<i>thau¹</i>	<i>ḍau¹</i>	<i>thau¹</i>
'cook, to'	<i>thoŋ¹</i>	<i>ḍuŋ¹</i>	
'carry hanging'	<i>thiu³</i>	<i>ḍiu³</i>	<i>thiu³</i>
'hailstone'	<i>thet⁷</i>	<i>thap⁷</i>	

Summarising the comparison above:

Giáy	<i>*l-, *kl/r-, *xr-, *dl/r-, *r-, *hr-, *hw-, *nl/r-, *thr-</i>	→ ḍ-
Nùng An	<i>*kl/r-, *xr-, *r-, *hr-, *hw-</i>	→ <i>h-, r-, tɕ-, k-</i>
	<i>*l-</i>	→ <i>l-</i>
	<i>*nl/r-</i>	→ <i>n-</i>
	<i>*thr-</i>	→ <i>th-</i>
Nùng Giang	<i>*l-, *dl/r-</i>	→ <i>l-,</i>
	<i>*kl/r-</i>	→ <i>k(l)-</i>
	<i>*xr-</i>	→ <i>khj-</i>
	<i>*r-</i>	→ <i>hr-</i>
	<i>*hr-</i>	→ <i>h-</i>
	<i>*hw-</i>	→ <i>v-</i>
	<i>*thr-</i>	→ <i>th-</i>

we can see that Nùng An occupies a position much closer to Giáy and is more apt to merge proto-initials than Nùng Giang, which preserves nearly all the contrasts in the parent language.

5 Tone differences, other phonological and lexical differences

Another pattern of NT vis-à-vis CT is that a number of common items are found in the lower tone set, whereas CT and SWT have these in the high set. On this litmus Nùng An sides with the North.

Gloss	Nùng An	Giáy	Nùng Giang
'excrement'	<i>khɿ⁴</i>	<i>ʔe⁴</i>	<i>khjɿ³</i>
'rice'	<i>hɿu⁴</i>	<i>hau⁴</i>	<i>khɿu³</i>
'bowl'	<i>tuɿ⁴</i>	<i>tuɿ⁴</i>	<i>thuɿ³</i>
'bean'	<i>tu⁶</i>	—	<i>thu⁵</i>
'bitter'	<i>ham²</i>	<i>ham²</i>	<i>kham¹</i>
'eggplant'	<i>khui¹</i>	<i>ku²</i>	<i>khi¹</i>
'son-in-law'	<i>kwɿ²</i>	<i>kwɿ²</i>	<i>khi¹</i>

There are also some differences of tone or segmental elements that seem to be restricted to individual words. First consider the items in which Nùng An resembles the NT languages.

Gloss	Nùng An	Giáy	Nùng Giang
'ladder'	<i>lai¹</i>	<i>lai¹</i>	<i>?dai¹</i>
'this'	<i>nəi³</i>	<i>nɪ⁴</i>	—
'a fly'	<i>neŋ²</i>	<i>neŋ²</i>	<i>miŋ²</i>
'sleep, to'	<i>nə:n²</i>	<i>nin²</i>	<i>nu:n²</i>
'meat'	<i>nɔ⁶</i>	<i>nɔ⁶</i>	<i>nɪ⁴</i>

Finally there are many items (see Zhang & Wei 1997) in which NT languages and CT languages have a lexical difference. Consider the items below, which show similarity between columns 1 and column 3, that is Nùng An and Nùng Giang, and differ from column 2, Giáy.

'cloth'	<i>pha:i⁵</i>	<i>paŋ²</i>	<i>pha:i³</i>
'moon'	<i>ha:i¹</i>	<i>?duan¹</i>	<i>ha:i¹</i>

This table shows cognate items between column 1 and column 2, that is Nùng An and Giáy, and are different from the items in column 3, Nùng Giang, a CT language.

'sun'	<i>thuŋ¹ ŋon²</i>	<i>tɕan³³ wan²</i>	<i>tha¹ wən²</i>
'clothing'	<i>pəu¹</i>	<i>pɯ⁶</i>	<i>θu³</i>
'sky'	<i>?bən¹</i>	<i>?buən¹</i>	<i>fa⁴</i>
'horn'	<i>kɔk⁷</i>	<i>kok⁷</i>	<i>kok⁷</i>
'yesterday'	<i>lu:n²</i>	<i>lu:n²</i>	<i>va²</i>
'tiger'	<i>kok⁹</i>	<i>kuk⁹</i>	<i>ɬwa¹</i>
'flower'	<i>va²</i>	<i>?dok⁹ va²</i>	<i>?bjok⁷</i>
'evil spirit'	<i>maŋ¹</i>		<i>phi¹</i>
'yellow'	<i>heŋ³</i>	<i>he:n³</i>	<i>lu:ŋ¹</i>
'wing'	<i>fəət¹⁰</i>	<i>fɯwt¹⁰</i>	<i>pek⁷</i>
'head'	<i>rau¹</i>	<i>tɕeu¹</i>	<i>thu¹ khjɛi¹</i>
'slippery'	<i>laau²</i>	<i>laau²</i>	<i>laau²</i>
'rice mortar'	<i>rum¹</i>		<i>kjok⁸</i>
'wide'	<i>kwaan⁵</i>		<i>kwaan³</i>
'right'	<i>kwa²</i>		<i>ɬa¹</i>
'bean'	<i>tu⁶</i>		<i>thu⁵</i>

6 Conclusion

The phonological and lexical features of Nùng examined above show that most lexical items found in Nùng An are typical of NT and that the phonological rules are NT as well, except for two, (1) the **f-* → *ph-* rule and (2) the **thr-* → *th-* rule. However, Nùng An seems not to have borrowed these two phonological rules from the neighbouring Tai languages of Vietnam, such as Nùng Giang, but instead to have brought them from Long'an in Guangxi, which has exactly these two phonological rules as well, as the data below show (Long'an data taken from Zhuangyu Yinxi 1959):

Gloss	Nùng An	Long'an ⁵
(Examples of *f- → ph-)		
'cloud'	<i>phur^t</i>	<i>phur^t</i>
'rain'	<i>phən^l</i>	<i>phun^l</i>
(Examples of *thr- → th-)		
'stone'	<i>thin^l</i>	<i>thin</i>
'carry on end of pole'	<i>thaap⁹</i>	<i>thaap⁹</i>
'to die'	<i>thai^l</i>	<i>thai^l</i>

Thus, Nùng An and Long'an appear to have the same mix of NT and CT lexical and phonological features. The simplest account of this similarity accords with the report of the people themselves; they brought these NT features with them when they immigrated from further north in an area located on the border between NT and CT areas. The explanation for the NT features is to be found in the immigrant status of a NT people to a CT-speaking area in Cao Bang Province and it should not be considered support for the ?Yai-Hypothesis.

References

- Edmondson, Jerold A. and Li Shaoni, 1994, Voice quality and voice quality change in the Bai language of Yunnan Province. *LTBA* 17/2:49–68.
- Edmondson, Jerold A. and David B. Solnit, eds, 1997, *Comparative Kadai: the Tai Branch*. Dallas: SIL/UTA Series in Linguistics.
- Gregerson, Kenneth J. and Jerold A. Edmondson, 2000, Puzzles in Cao Lao. In S. Burusphat, ed. *Proceedings of the International Conference on Tai Studies*, 151–163. Held in Bangkok, July 29–31, 1998.
- Haudricourt, André-Georges, 1956, De la restitution des initiales dans les langues monosyllabique: le problème du thai commun. *BSLP* 52/1:307–322.
- 1960, Note sur les dialecte de la région de Moncay. *BEFEO* 50:167–177.
- 1963, La langue Sek. *25th Congress of Orientalists*, Moscow, vol.4, 414–415.
- 1973, Mấy điều nhân xét về lý luận và thực tiễn nhân một chuyến đi thăm các dân tộc thiểu số ở Tây Bắc Việt Bắc. *Ngôn Ngữ* 3.5. [Some remarks about theory and practice on the occasion of a trip to visit the ethnic groups of northwest and northeast Vietnam.]
- Li Fang Kuei, 1977, *A handbook of comparative Tai*. Honolulu: The University Press of Hawaii.
- MVNP, Vietnam Social Sciences Committee, ed., 1978, *Các dân tộc ít người ở Việt Nam (các tỉnh phía bắc)*. [Minorities of Vietnam, the northern provinces.] Hanoi: Social Sciences Publishing House.
- Strecker, David, 1985, The classification of the Caolan languages. In S. Ratanakul, et al., eds *Southeast Asian linguistic studies presented to Andre-G. Haudricourt*. 479–492. Bangkok: Mahidol University.

⁵ The Long'an data available to me is rather limited, but it is already apparent that there are also some important differences between the Nùng An data I collected and the vernacular of Genju Siqu Village, Long'an County, as the velar series are rather different and the Genju Siqu Village vernacular appears to have more Chinese loan words (see Zhuangyu Yinxi 1959:317–323).

Zhang Yuansheng and Wei Xingyun, 1997, Regional variants and vernaculars in Zhuang. In Edmondson and Solnit, eds *Comparative Kadai: the Tai Branch*. 77–96. Dallas: Summer Institute of Linguistics.

Zhuangyu Yinxu, 1959, Beijing: Chinese Academy of Social Sciences.

5 *Morphosyntactic evidence for the position of Chamorro in the Austronesian language family*

LAWRENCE A. REID

1 Introduction¹

The Chamorro language is an Austronesian language spoken in the Marianas Islands. Its position within the Austronesian language family has been a continuing topic of discussion for more than a hundred years. The difficulty in determining its position has apparently resulted from linguistic contact with a variety of other languages, both Oceanic as well as Western Austronesian, ranging from chance settlement from drifting sailors, to established trade networks with other island groups, possibly exchange of wives, and probably also through invasion by other Austronesian-speaking groups and eventually in historic times, through colonialisation under Spanish, German, Japanese and American governments. Other factors, such as natural disasters (the Marianas is not infrequently devastated by typhoons) and introduced diseases, are known to have at various times severely reduced population levels in the Marianas, and this again would have probably hastened change in the language. Hertha Costenoble, a native speaker of Chamorro, who studied linguistics under Dempwolff, proposed three different strata discernible in the multiple Chamorro reflexes of reconstructed Austronesian sounds, resulting in many doublets, i.e. pairs of words with slightly different pronunciation and with the same or similar meanings, one of which may have been directly inherited, while the other was borrowed from another Austronesian language which had undergone different sound changes from the inherited vocabulary (Costenoble 1940).

¹ An early version of this paper was presented to the 16th Congress of the Indo-Pacific Prehistory Association in Melaka, Malaysia, 1–7 July 1998. A later version was presented to the International Symposium on Problems in Morphosyntactic Comparison and Reconstruction, ILCAA, Tokyo University of Foreign Studies, 13 December 1998. I would like to express my thanks to the participants in these meetings for their comments, and particularly to Stan Starosta for his detailed discussion of many of the points made herein. Although he is not convinced of the correctness of my interpretation of the evidence in some cases, that is to be expected, given the current uncertainty that exists regarding the nature of higher level subgrouping in Austronesian, and our limited understanding of the processes of morphosyntactic change.

In addition to the lexical doublets discussed by Costenoble, the language also exhibits a number of pairs of syntactic structures, one of which must have been inherited and the other probably borrowed, just as in English there are two possessive constructions, one of which (the 'apostrophe -s' form) is inherited, and the other (the 'of' construction) is borrowed. These will be discussed later in the paper.

This paper first gives an overview of various claims that have been made about the position of Chamorro vis-à-vis other Austronesian languages. It will then discuss, in the context of these claims, various problems which inevitably arise when attempting to do morphosyntactic comparison. These problems are theoretical as well as practical, although solutions to the latter often depend upon solutions to the former.²

The first more general theoretical problem is that of comparability between analyses. It is simply not possible to properly compare the grammars of two languages that have been written using different theories. One may well be able to recognise that a given form, or its cognate occurs across a set of languages, but if that form has a syntactic function, then a common theory of the syntax of those languages is necessary to determine whether they have an equivalent function or not. Even with a common theory, determining their equivalence can still be problematic. As will be seen below, there are as many analyses of Chamorro pronouns as there have been linguists who have described them. And I shall propose yet another.

A second theoretical problem is the problem that all comparativists face, that of the inherent circularity of the comparative method. Subgrouping hypotheses are based on the accumulation of shared innovations in the phonology, lexicon, morphology and syntax of sets of languages. Yet it is not possible to distinguish an innovation from a retention without a subgrouping hypothesis.

² Abbreviations used in this paper are as follows:

Austronesian subgroup names:

CEMP	Central-Eastern Malayo-Polynesian
NMP	Nuclear Malayo-Polynesian
PAn	Proto Austronesian
PEF	Proto Extra-Formosan
PMP	Proto Malayo-Polynesian
PNMP	Proto Nuclear Malayo-Polynesian
PPh	Proto Philippines
WMP	Western Malayo-Polynesian

Language names:

Chm	Chamorro
Ilk	Ilokano
Ivt	Ivatan
Pal	Palauan
Tag	Tagalog

Lexicase case forms:

Acc	Accusative
Erg	Ergative
Gen	Genitive
Lcv	Locative
Nom	Nominative
Obl	Oblique

Lexicase case relations:

PAT	Patient
AGT	Agent
COR	Correspondent
MNS	Means
LOC	Locus

Lexicase features:

actr	actor
dfnt	definite
plrl	plural
prdc	predicate
prnn	pronoun
sttv	stative
trns	transitive

Pronoun abbreviations:

1SG	first person singular
2SG	second person singular
3SG	third person singular
1PL.INC	first person inclusive plural
1PL.EXC	first person exclusive plural
2PL	second person plural
3PL	third person plural

Other abbreviations:

Det	Determiner
LIG	Ligature

Practical problems include establishing criteria for determining whether the absence of a grammatical form in a language is the result of loss (and therefore an innovation), or whether it had not yet been innovated at the time the ancestral speakers of that language separated from the group of languages in which the form is found. Another practical problem, somewhat related, is determining which of two comparable forms represents the innovation and which the retention, or whether one or the other is the result of language contact. Examples of each problem occur in the discussion below.

2 Earlier views on the relationship of Chamorro to other Austronesian languages

Costenoble's (1940) claim that Chamorro exhibits three different strata of lexical items in the language was based on his understanding of the phonology of Proto Austronesian as it had been reconstructed at the time (primarily by Otto Dempwolff), and he did not have the benefit of data from any of the Formosan languages. However, his statements of the reflexes are generally well supported, although views of the PAN sound system have undergone considerable change since that time (see Blust 1997 for an overview). Zobel summarises the subgrouping implications of the phonological innovations that have taken place in Chamorro, and notes that 'the sound changes are either found in many other WMP languages (...merger of **e* and **u*...), or are unique (...merger of **D* [Blust's **d*] and **k*, **j* and **q*...)' (Zobel 2002:406), implying that it is not possible to claim anything about the subgrouping of Chamorro based on phonological evidence alone.

Although Costenoble recognised that the presence of various linguistic strata implied language contact, it was probably Topping (1973:3) who first proposed a possible genetic relationship for Chamorro by appealing to morphosyntactic evidence, but noting also the possibility that the evidence could be the result of language contact. He claimed that

Chamorro is a Philippine type language, and its closest linguistic relatives are probably Ilokano and Tagalog. This opinion is based on the many similarities in the grammatical structures of the languages...It is quite possible that these similarities in the grammatical devices were borrowed from Filipinos with whom the Chamorro traded. However, this is very unlikely.

Starosta and Pagotto (1991) compared the syntactic features of Chamorro with those of a Philippine language (Tagalog), a Formosan language (Tsou) and a Micronesian language (Marshallese), in an attempt to discover whether or not there was any morphosyntactic evidence for subgrouping Chamorro with these languages, but concluded that each of the features examined was probably inherited from PAN and therefore provided no evidence for subgrouping. There were no exclusively shared innovations in the morphology or syntax to support a subgrouping argument. Subsequently, Starosta (1995) claimed that Chamorro shares a set of morphosyntactic innovations with a subgroup of languages (named F3) which includes all Austronesian languages except Rukai, Tsou, and Saaroa in Formosa, placing the Chamorro split from other Austronesian languages at a very early stage, considerably prior to that which resulted in the Austronesian settlement of the Philippines, Indonesia and Oceania.

Zobel (2002) in a wide-ranging paper attempts to provide evidence from verb morphology and morphosyntax to claim that Chamorro shares a number of innovations which he reconstructs for the parent language of a set of Malayo-Polynesian languages which exclude

the Philippines, North Sulawesi, and Northeast and Interior Borneo.³ This new subgroup he labels NUCLEAR MALAYO-POLYNESIAN. He considers Chamorro and Palauan to be ‘early offshoots from PNMP’ (Zobel 2002:431). He furthermore speculates that Pre-Chamorro and Pre-Palauan speakers probably sailed from Sulawesi, although he believes that NMP languages could have been spoken in the Southern Philippines and this area could also be considered to be a possible departure point for the first migrants to the Marianas and to Palau. Zobel makes a good case for his theory, but the question that must be asked is whether the evidence that he adduces is evidence for a genetic relationship, or whether it is evidence for contact. This problem will be addressed later in the paper (§4.5), but first it is necessary to examine his claims about the nature of Chamorro itself.

3 Is Chamorro an ergative language or not?

Zobel (2002) claims that Chamorro is not a language with a ‘focus’ system, as described by Topping (1973). Instead, following Cooreman (1987), he describes it as a split-ergative system with the split being conditioned by whether the verb is realis or irrealis. He states, ‘in realis there is ergative pronoun marking, while in irrealis there is nominative marking.’ (Zobel 2002:410). In order to determine the validity of this claim, it is necessary to take a close look at the Chamorro pronominal system. The system, as described by Topping (1973) is shown in Table 1.

Table 1: Chamorro pronouns (Topping 1973)

	A	B	C	D
1SG	<i>hu</i>	<i>yo'</i>	<i>-hu</i>	<i>guahu</i>
2SG	<i>un</i>	<i>hao</i>	<i>-mu</i>	<i>hagu</i>
3SG	<i>ha</i>	<i>gue'</i>	<i>-ña</i>	<i>guiya</i>
1PL.INC	<i>ta</i>	<i>hit</i>	<i>-ta</i>	<i>hita</i>
1PL.EXC	<i>in</i>	<i>ham</i>	<i>-(n)mami</i>	<i>hami</i>
2PL	<i>en</i>	<i>hamyo</i>	<i>-(n)miyu</i>	<i>hamyo</i>
3PL	<i>ma</i>	<i>siha</i>	<i>-(n)ñiha</i>	<i>siha</i>

According to Topping (1973:106–111, 262), the functions of these pronouns are as follows:

- A. Subject markers which always precede the verb. They are required in transitive sentences, even if there is a full noun phrase present. They are also required if the verb is intransitive and marked for future tense.

³ Specifically, Zobel (2002:430–431) claims that

the NMP [Nuclear Malayo-Polynesian] subgroup includes the languages of CEMP [Central-Eastern Malayo-Polynesian] group, Chamorro and Palauan, and most WMP [Western Malayo-Polynesian] of Malaysia and Indonesian. Not included in the NMP group are the following WMP languages: the languages of the Philippines, the three Northern Sulawesi groups (Gorontalo-Mongondic, Minahasan, Sangiric), the Sama-Bajau languages, Malagasy, and all languages of Borneo with the exception of the Malayic and Tamanic groups.

- B. Subject and Object pronouns which always follow the verb. They are subject pronouns if they occur in an intransitive sentence, or in a transitive sentence with a non-specific object. They are object pronouns if they occur as a specific object in a transitive sentence.
- C. Possessive pronouns. They are bound morphemes, or enclitics. The disyllabic forms also require an 'excrescent' consonant *n* before they are joined to a vowel-final stem. Some auxiliary verbs, such as *ilek* 'say', *ga'o* 'prefer', and *ya* 'want', always require this set of pronouns as subject. They also occur attached to verbs following certain question words.
- D. Emphatic pronouns. Usually these are emphasised subjects, but may occur in other contexts as well.

Zobel's (2002) analysis of the pronouns (see Table 2) is somewhat different from that of Topping. He labels set A 'agentive', set B 'absolutive', set C 'possessive, and A [agent] in certain nominalizations'. He also notes that there is another pronoun set, the forms of which are very similar to those of the agentive set A but which he labels 'nominative' since (as Topping had noted) they occur as the subject of irrealis (Topping's 'future') intransitive sentences. It is on the basis of this analysis that Zobel claims that Chamorro is a split-ergative language, with the split conditioned by mood distinction.

Table 2: Chamorro pronouns (Zobel 2002)

	A1	A2	B	C	D
	Agentive	Nominative ⁴	Absolutive	Possessive	Free
1SG	<i>hu</i>	<i>(bai) hu</i>	<i>yo'</i>	<i>-hu</i>	<i>guahu</i>
2SG	<i>un</i>	<i>un</i>	<i>hao</i>	<i>-mu</i>	<i>hagu</i>
3SG	<i>ha</i>	<i>u</i>	<i>gue'</i>	<i>-ña</i>	<i>guiya</i>
1PL.INC	<i>ta</i>	<i>(u) ta</i>	<i>hit</i>	<i>-ta</i>	<i>hita</i>
1PL.EXC	<i>in</i>	<i>(bai) in</i>	<i>ham</i>	<i>-mami</i>	<i>hami</i>
2PL	<i>en</i>	<i>en</i>	<i>hamyo</i>	<i>-miyu</i>	<i>hamyo</i>
3PL	<i>ma</i>	<i>uha/u/uma</i>	<i>siha</i>	<i>-ñiha</i>	<i>siha</i>

3.1 A re-analysis of the Chamorro pronoun system

Working without the benefit of any constraining theory allows one to multiply entities indefinitely, in this case pronouns sets and subsets, and to assign functions to them on the basis of translation equivalents (as in the case of Topping), or of apparent typological equivalents (as in the case of Zobel). The analysis that I provide in the following sections will be couched within Lexicase Dependency Grammar, a highly constrained theory originally proposed by Starosta and developed by him and his students at the University of Hawai'i and elsewhere.

I shall claim here that Chamorro is a pure ergative language, and that the pronominal system of Chamorro is much simpler than that shown by either Topping or Zobel. I shall claim that there is only one case-marked set of pronouns in Chamorro, that which in most

⁴ It should be noted that both *bai* and *u* are future or irrealis markers which occur either optionally or obligatorily in combination with the pronominal forms. They are not themselves pronouns.

typological studies of ergative languages is called Absolutive, but which I shall call, following Starosta, Blake and others, Nominative. It is probable that these forms are enclitics. The free pronouns (Topping's 'emphatic pronouns') are not themselves case-marked but receive case depending on the sentential context in which they occur.

3.1.1 *Possessor agreement markers*

I shall begin by discussing the so-called 'possessive set'. These forms are clearly reflexes of the Proto Extra-Formosan enclitic Genitive set. In many languages, such as the northern languages of the Philippines (Reid 2001), some of these pronouns have lost their enclitic status and have been incorporated into their former head word as agreement markers. Once this takes place, the forms lose their syntactic independence and no longer carry case-marking. It is very clear from Topping's description that all of the 'possessive pronouns' are now possessor agreement markers, phonologically and syntactically incorporated into their former head words. Topping considers them enclitics, although he also frequently refers to them as suffixes (Topping 1973:42), and usually writes them as hyphenated forms. But consider the following evidence, both phonological as well as syntactic, that strongly suggests they are no longer enclitics and therefore no longer pronouns.⁵ Starosta and Pagotto (1991) briefly noted some of the evidence outlined below and came to a similar conclusion that these are 'derivational affixes, rather than pronouns as such' (Starosta & Pagotto 1991:332).

3.1.1.1 *The phonological evidence*

There are three kinds of phonological evidence that may be considered. First, each of the forms is treated as an integral part of the word for the placement of primary stress. Most non-Spanish Chamorro words are stressed on the penultimate syllable. The addition of the possessive pronominal forms results in the movement of stress, one syllable to the right in the case of the single syllable forms, two syllables to the right in the case of the disyllabic forms, as in (1).

- (1) Chamorro (Topping 1973:42)
- a. /háso/ 'think'
 - b. /hináso/ 'thought'
 - c. /hinassómu/ 'your thought'
 - d. /hinassonmámi/ 'our thought'

Secondly, assimilative processes which are unique to the first person singular pronominal form suggest that this form, at least, is an integral part of the word. In most environments the first person singular form is *-hu*, a regular reflex of the reconstructed form **-ku*. There are several variant pronunciations, depending on the phonological shape of the form to which it is attached. According to Topping (1973:110), it is *-su* if the stem has a final *s*, and *-tu* if the stem has a final *t*. These assimilative processes are not found elsewhere in the language. If the stem has a medial consonant cluster, the form is *-ku*, a unique condition in the language blocking the operation of the sound change, **k* to *h*, as in (2).

⁵ See Zwicky (1977, 1985) and Zwicky and Pullum (1983) for criteria for distinguishing clitics from full words, and affixes from clitics.

- (2) Chamorro (Topping 1973:110)
- a. /lássas/ 'skin' /lassássu/ 'my skin'
 - b. /páchoť/ 'mouth' /pachóttu/ 'my mouth'
 - c. /lépblo/ 'book' /lepblóku/ 'my book'

Thirdly, the rule for the formation of continuative aspect applies not just to the basic form, but to the full word with the pronominal form attached. Continuative aspect is marked by reduplication of the syllable that carries primary stress. When the word happens to have a disyllabic pronominal form attached to it, it is the first syllable of the pronominal form that is reduplicated, since that is the part of the word that carries primary stress, as in (3).

- (3) Chamorro (Topping 1973:259)
- a. *Hafa bidanñiha?* 'What did they do?'
 - b. *Hafa bidanñíñiha?* 'What are/were they doing?'

3.1.1.2 The syntactic evidence

One of the characteristics of agreement markers is that they usually appear even when the full noun phrase with which they agree actually occurs in the sentence. In Chamorro there are two sets of possessive constructions (see §4.1 below), one of which requires what Topping (1973:223) refers to as the 'Full Possessive Form'. In this construction, a third person singular or plural possessive noun phrase must co-occur with an appropriate possessor agreement marker on the head noun, respectively either *ña* or *ñiha*. For example:

- (4) Chamorro (Topping 1973:223)
- a. *i gima'ña si Rosa* 'Rosa's house'
 - b. *i malago'ñiha i taotao* 'the people's wish'

Taken together, the phonological and syntactic evidence strongly suggests that Topping and Zobel's possessive pronouns are not separate words, but are possessive agreement markers, integral parts of their former head nouns.

3.1.2 Actor-agreement markers

In this section I shall discuss Topping's set A 'subject' pronouns. Although each of the forms in this set are written as separate words, their status as separate words is questionable. Nothing may intervene between them and the verb which follows them, implying that they are at least proclitics. That they may actually be part of the verb as agreement markers is suggested by two pieces of evidence.

3.1.2.1 Ambiguity between ma '3PL' and ma- 'passive marker'

Topping notes (1973:258) that the third person plural form in this set is homophonous with, and probably originated from what he refers to as 'the passive marker *ma-*'. He cites (5a,b) as cases of such ambiguity.

- (5) Chamorro (Topping 1973:258)
- a. *Malalalatde i patgon.* 'The child is being scolded,' or:
'They are scolding the child.'
- b. *Masangan na maolek iya Guam.* 'It is said that Guam is good,' or:
'They say that Guam is good.'

It is to disambiguate such sentences in the writing system, he says, that the 'pronoun *ma*' is written as a separate word, while the 'passive marker *ma-*' is written as a prefix, as in (6).

- (6) Chamorro
- a. *Malalalatde i patgon.* 'The child is being scolded.'
- b. *Ma lalalatde i patgon.* 'They are scolding the child.'

3.1.2.2 *Co-occurrence with a lexical noun*

It was noted above that one of the characteristics of agreement markers is that they usually appear even when the full noun phrase with which they agree actually occurs in the sentence. Topping notes (1973:203) that in those constructions which require his set A pronouns, the third person forms are obligatory, even though a lexical noun is present, as in (7a,b). He therefore chooses to refer to at least the third person forms as 'subject markers' rather than as subject pronouns. For example:

- (7) Chamorro (Topping 1973:203)
- a. *Ha li'e' i guaka.* 'He saw the cow.'
- b. *I patgon ha li'e' i guaka.* 'The child saw the cow.'

3.1.2.3 *Set A forms as Actor-agreement affixes*

When one examines the pronominal forms that appear in what are probably transitive constructions, it seems clear that Chamorro is an ergative language, the Patient (or 'O') is marked by the same pronominal set that also marks the Actor Patient (or 'S') of intransitive sentences (i.e. Zobel's Absolutive set, Topping's set B), while the Actor Agent ('A') is marked by either an oblique set of pronouns, or by set A forms. A problem arises for Zobel, however, because the set A forms also seem to mark the grammatical subject ('S') of intransitive sentences, when these carry one of the future tense auxiliaries, as in (8a,b).

- (8) Chamorro (Topping 1973:263)
- a. *Para un li'e' i lahi.*
future you see the man
'A' +trns 'O'
'You will see the man.'
- b. *Para un saga giya Yigo.*
future you stay in Yigo
'S' -trns
'You will stay in Yigo.'

It is on the basis of data such as these that Zobel concludes that the set A forms used in transitive sentences are Agentive, while those occurring in intransitive sentences are Nominative, and therefore the language must be split-ergative. Tchekhoff (1991:501) faced a similar problem in her analysis of Tongan, finding that that language had ergative morphology with nominal participants, but had accusative constructions with pronouns. Kikusawa's re-analysis of the Tongan data (1997) within the Lexicase dependency grammar framework, concluded that Tchekhoff's 'nominative pronouns' showed all the characteristics expected of agreement markers, although being written as separate words. They were analysed by her as Actor-agreement markers. She considers them to be part of the verb, not syntactically independent words, and therefore not eligible to carry case. The same analysis should also be considered for the Chamorro set A forms. The forms are not pronouns, they are Actor-agreement markers, leaving Chamorro as a pure ergative language. Starosta and Pagotto (1991:332) recognise that the forms being discussed here are Actor-agreement markers, however they also refer to them as 'clitic pronouns'. They correctly note that the forms are actor markers in transitive realis constructions and in transitive and intransitive irrealis clauses, but fail to recognise that treating them as pronouns in intransitive irrealis clauses would force them to treat Chamorro as a split-ergative language, as Zobel did. Gibson (1990:248) similarly notes that Chamorro has an agreement system, but describes it as an 'ergative agreement system in realis clauses', disregarding the fact that the same set of forms also occurs in intransitive irrealis clauses. Treating the forms as Actor-agreement markers captures a generalisation missed by her.

3.2 Source of Chamorro Actor-agreement markers

In many ergative Austronesian languages (including most Philippine and Formosan languages) the pronominal Agent of a transitive clause is identical in form to the possessor of a noun, and has therefore been described as having a Genitive case form. In Chamorro, however, the forms of the Actor-agreement markers are very different from those which mark possessor on a noun and which reflect PEF reconstructed Genitive pronouns (see Table 3). The two sets only share reflexes of the first person singular, *hu*, and the first person exclusive, *ta*. In addition, the Actor-agreement markers all occur at the beginning of a word (indicated by the opening square bracket '['), while the possessor-agreement markers all occur at the end of a word (indicated by the closing square bracket ']').⁶

⁶ The use of a right square bracket ']' after a form is a lexicase convention to indicate that the phonological sequence that precedes it occurs at the end of a word. A left square bracket '[' marks a form as occurring at the beginning of a word. Although in this paper these forms are given glosses as though they were affixes, strictly speaking in lexicase the forms are considered to be integral parts of the words which they end. Similarly, the agreement features that are given here as glosses are features of the full word, and should not be considered to be uniquely associated with the ending.

Table 3: Chamorro agreement markers

	Actor-agreement	Possessor-agreement
1SG	[<i>hu</i>	<i>hu</i>]
2SG	[<i>un</i>	<i>mu</i>]
3SG	[<i>ha</i>	<i>ña</i>]
1PL.INC	[<i>ta</i>	<i>ta</i>]
1PL.EXC	[<i>in</i>	(<i>n</i>) <i>mami</i>]
2PL	[<i>en</i>	(<i>n</i>) <i>miyu</i>]
3PL	[<i>ma</i>	(<i>n</i>) <i>ñiha</i>]

Two questions arise: First, how did Actor-agreement markers get to the beginning of the verb, when comparative evidence is clear that as far back as PAN, Genitive pronouns were enclitics, not proclitics? Second, what were the processes by which the forms became so different from the possessive-agreement markers?

3.2.1 Initial position agreement markers

Starosta and Pagotto (1991) point out that the most general account of the positioning of pronouns (not just in Austronesian languages) requires that so-called ‘auxiliaries’ be treated as main verbs. Under this hypothesis, it was not atypical for Genitive and Nominative pronouns to appear as enclitics to an initial, ‘auxiliary’ main verb and to be followed by a dependent ‘main verb’. Starosta, Pawley and Reid (1982) have shown that one of the processes by which Oceanic languages became SVO from an earlier VSO structure was through what was labelled in that paper as ‘Aux-axing’, i.e. the loss of an auxiliary verb and the resultant stranding of the pronominal clitics that were dependent on that verb. In many languages the pronouns then became phonologically dependent on what followed them, becoming proclitics to that verb. Thus:

$$V=\text{prnn } V \rightarrow \text{prnn}=V$$

This process has occurred independently in a number of languages, especially in contexts in which the function of the original auxiliary verb could be inferred either from context, or simply from the position of the pronoun itself in pre-verb position. In both Inibaloi and Ivatan, for example, loss of an initial verb meaning ‘go’ resulted in pronoun-initial imperative sentences, see (9) and (10 a, b). Notice the free translations from both sources indicate the meaning ‘go’ is implied in these sentences.

- (9) Inibaloi (Ballard et al. 1971:24)
Jo di olop jet idaw jo la'd ma Peshis.
 you here fetch and bring you just=to MA Peshis
 ‘Go fetch him and just bring him to Peshis.’
- (10) Ivatan (Hidalgo & Hidalgo 1971:214, 239)
 a. *Ka machinanaw, mu Marya.*
 you study you Mary
 ‘You (go and) study, Mary.’

- b. *Marya, mu rutungan u manuk.*
 Mary you cook Nom chicken
 'Mary, you (go and) cook the chicken.'

Similarly, the process has begun in Tagalog, with the imperative meaning 'Let's go, now!' consisting solely of what was earlier a Nominative pronoun, now re-analysed as a verb carrying a set of agreement features for 1PL.INC, followed by an adverb. Compare the following Ilokano and Tagalog constructions (11 and 12), one of which (Ilokano) retains the initial (auxiliary) verb, and the other in which the earlier verb has been lost.

- (11) Ilokano (Reid)

Intayon!

go.we.now

'Let's go now!'

- (12) Tagalog (Reid)

Tayo na!

go.we now

'Let's go now!'

4 Syntactic doublets in Chamorro

There are at least three sets of parallel constructions, or syntactic doublets, in Chamorro which suggest the possibility that the language has undergone massive change as the result of intensive language contact.

4.1 Possessive constructions

Topping (1973:223) discusses three types of possessive constructions, one of which he labels the 'full possessive form'; another he calls the 'construct form'. The third is a possessive classifier construction. The first requires the presence of possessive-agreement marking⁷ on the possessed noun, either *n̄a* '3SG' or (*n*)*n̄iha* '3PL', with the possessor noun being preceded by a determiner, either *si* for personal nouns, or *i* for common nouns, as in (13a,b). The 'construct form' has neither of the above agreement markers on the possessed noun, nor determiners before the possessor noun. It does however require that a possessed noun ending in a vowel be followed by *n*], which, in effect, is also a general third person agreement marker, as in (13c,d). Topping notes that there is very little difference in meaning between the two constructions, and that they somewhat parallel the difference expressed in English by the translations given. The possessive classifier construction is similar to the full possessive construction. It requires that the possessed head be one of the classifier nouns: *na* 'edible thing', *ga* 'non-human animal', *iyó* 'inanimate thing', or *gimen* 'drinkable thing'. The classifier noun must carry agreement marking, and the specific object to which it refers immediately follows, but without a determiner, as in (13e,f).

⁷ Topping's 'possessive pronoun'.

(13) Chamorro (Topping 1973:223)

Full possessive

- a. *i gima'ña si Rosa* 'Rosa's house'
 b. *i hagaña i rai* 'the king's daughter'

Construct

- c. *i gima' Rosa* 'the house of Rosa'
 d. *i hagan rai* 'the daughter of the king'

Possessive classifier

- e. *i na'hu mannok* 'my chicken (to eat)'
 f. *i ga'hu mannok* 'my chicken (pet)'

4.1.1 Possible sources of Chamorro possessive constructions

Costenoble (1940) discussed several lexical forms in Chamorro which he believed were evidence of borrowing from some Oceanic language. It is probable that at least the possessive classifier construction and possibly also the forms with the construct *n]* are also the result of contact with an Oceanic language. In discussing these forms, Starosta and Pagotto (1991) note that construct possessive constructions occur in Micronesian languages such as Marshallese, as do possessive classifier type constructions, however at least in Marshallese the word order is the opposite of that in Chamorro. Nevertheless other Oceanic languages have structures which exactly parallel the Chamorro structure, including Fijian, as in (14).

(14) Fijian (Kikusawa, pers. comm.)

- a. *ke-qu toa* 'my chicken (to eat)'
 b. *no-qu toa* 'my chicken (to sell, etc.)'

Although structures which appear to be similar to the construct forms of Oceanic languages are found in some Philippine languages, such as Bontok and other Central Cordilleran languages, they are probably only coincidentally similar. They are only parallel with respect to personal noun possessors (15a,b). Common noun possessors require a determiner between the possessed noun and the possessor, as in (15c,d). Note that *?ábuŋ* 'house' is a consonant final noun, while *?ásu* 'dog' is vowel final and therefore requires *n]* before Genitive noun phrases.

(15) Bontok (Reid)

- a. *nan ?ábuŋ Rosa* 'the house of Rosa'
 b. *nan ?ásun Rosa* 'the dog of Rosa'
 c. *nan ?ábuŋ nan laláki* 'the house of the man'
 d. *nan ?ásun nan laláki* 'the dog of the man'

It can be shown (Reid 1998) that the *n]* in the Central Cordilleran languages is the result of a local innovation. It is a reflex of PEF **ni* 'genitive determiner', a reflex of which can be seen in the Chamorro possessive-agreement markers *ña]* < **niya* '3SG', and *ñiha]* < **nida* '3PL'. The fact that Chamorro requires a construct *n* before *ñiha]*, as well as before other disyllabic possessive-agreement markers, argues for the introduction of the construct *n]* through contact or a local analogical development, rather than for its being an inherited form,

since the [*n̄i*-marked pronouns were already genitively marked. There would have been no syntactic motivation for the additional marking with the construct *n*].

I claim then that it is only the 'full possessive constructions' that are inherited. Such constructions are widespread throughout the family from Saaroa and Atayal in Formosa (Starosta & Pagotto [1991], citing Tsuchida 1976:279 and Egerod 1966:365), in Ivatan (16), Ilokano, and other northern languages of the Philippines, to some of the more conservative Western Fijian dialects in the Pacific (17).

(16) Ivatan (Reid 1966:125)

qo qama daaya no tatdo saaya ka makákakteh
 the father their.this Gen three these LIG sibling.PL
 'the father of these three sisters'

(17) Kadavu Fijian (Kikusawa, pers. comm.)

na taci.na ko Mere
 Det sibling.3SG Det Mere
 'Mere's sibling'

4.2 Existential possessive constructions

There are two clearly distinct sets of constructions headed by existential verbs (Topping 1973:80) which can be used to indicate possession. The first requires either the positive existential verb *guaha*, or the negative existential verb *taya'* as the predicate. It is followed by a possessed noun carrying a possessive-agreement marker, see (18a,b). The second requires a different set of existential verbs, either the positive *gai*, or the negative *tai*. The possessed noun without agreement marking follows the existential verb, while the possessor occurs last as a nominatively marked noun phrase, see (18c,d).

(18) Chamorro (Topping 1973:80)

- a. *Guaha salape'hu.* 'I have money.'
- b. *Taya' salape'hu.* 'I have no money.'
- c. *Gai salape'yo'.* 'I have money.'
- d. *Tai salape'yo'.* 'I have no money.'

Both of these constructions are found in Philippine (and Formosan) languages. The former is widespread both in the Cordilleran languages of the north, as well as in the south, and is probably the inherited form, see (19a,b), while the latter is an innovation, found also in the Central Philippine languages such as Tagalog (20a,b). It is probable that in Chamorro also, it is the former constructions that are inherited. The existential verb *guaha* is a regular reflex of PEF **wada* 'there is'. In Central Philippine languages, the reflexes of this verb, such as Tag *walà*, mean 'there is none'. The existential verbs of (18c,d) are possibly cognate with existential verbs found in Ivatan. Ivt *tayto* (< **ta-i-tu*) 'negative existential' is perhaps cognate with Chm *tai* [tay] and *taya'*. In both languages also, the *y*-glide is probably an incorporated enclitic determiner. Similarly Chm *gai* [gay] (< **a-y* with regular prothetic prevocalic *g*) is perhaps cognate with Ivt *ara* (< **a-da* 'exists-now') 'positive existential', as well as Ivt *ari* (< **a-di* 'exists-there') 'positive existential'. However Ivatan does not use a grammatical structure such as that illustrated in (18c,d).

- (19) Bontok (Reid)
- a. *wadʔay sipiŋku*
 exist money.my
 'I have money.'
- b. *maʔid sipiŋku*
 not.exist money.my
 'I don't have any money.'
- (20) Tagalog (Schachter & Otones 1972:274-275)
- a. *may Bangka ako*
 exist boat Nom.1SG
 'I have a boat.'
- b. *wala ako-ng relas*
 not.exist Nom.1SG-LIG watch
 'I don't have a watch.'

4.3 Transitive constructions

In his grammar, Topping (1973) describes two different types of construction, which although supposedly conveying almost identical information are very different in form. Noting the similarity with Philippine languages, he describes those verbs which carry 'focus' affixation and are followed by noun phrases which are case-marked by determiners, as constituting a 'focus' system. There are three 'non-actor focus' constructions.⁸ They are 'goal focus' (with verb initial [*Cin*]),⁹ 'referential focus' (with verb final *i*), and 'benefactive focus' (with verb final *iyi*). All have identical syntax. They are distinguished only by the forms of the verb. The 'goal focus' type is illustrated in (21a-d). Topping is careful to distinguish these constructions from those in which the verbs do not have 'focus' affixes, but which obligatorily carry initial agentive agreement markers, and whose nominal complements are case-marked only by word order (Agent precedes Patient), as in (21e-g). The former he calls 'focus' constructions. The latter he calls 'transitive' constructions. However in a theory such as Lexicase, which defines any construction which has both an Agent and a Patient complement as transitive, then the 'focus' constructions must also be transitive. For example:

- (21) Chamorro (Topping 1973:245, my analysis)

'Goal Focus' Constructions

- a. *Lini'e' i lahi ni palao'an.*
 see the man the woman
 +trns Nom PAT Erg AGT
 Nom Erg
 actr
 'The woman saw the man.'

⁸ Topping also analyses causative constructions as a type of focus.

⁹ Zobel (2002) analyses these constructions as passives, claiming that they are syntactically intransitive. But according to Topping the actor of such verbs is obligatory.

- b. *Lini'e' ni lahi i palao'an.*
 see the man the woman
 +trns Erg AGT Nom PAT
 Erg Nom
 actr
 'The man saw the woman.'

- c. *Lini'e' si Maria as Pedro.*
 see Det Maria Det Pedro
 +trns Nom PAT Erg AGT
 Nom Erg
 actr
 'Pedro saw Maria.'

- d. *Lini'e' as Maria si Pedro.*
 see Det Maria Det Pedro
 +trns Erg AGT Nom PAT
 Erg Nom
 actr
 'Maria saw Pedro.'

'Non-Focus' Constructions

- e. *Hali'e' i lahi i palao'an.*
 see Det man Det woman
 +trns +dfnt AGT +dfnt PAT
 ?3SG¹⁰ Erg Nom
 ?actr actr
 'The man saw the woman.'

- f. *Mali'e' i lalahi i palao'an.*
 see Det men Det woman
 +trns +dfnt AGT +dfnt PAT
 ?3PL Erg Nom
 ?actr actr
 'The men saw the woman.'

- g. *Hali'e' si Pedro si Maria.*
 see Det Pedro Det Maria
 +trns AGT PAT
 ?3SG Erg Nom
 ?actr actr
 'Pedro saw Maria.'

¹⁰ The question mark is a lexicase convention to mark a feature that the head looks for, expects, or implies. In this case it indicates that the head carries third person singular Actor-agreement marking. The use of traditional person and number marking on pronouns and agreement markers in this example and elsewhere in this paper is not a lexicase convention. Here, in order to reduce the number of interlinear lines, they are used as abbreviations for the lexicase formal device of referring to such forms as sets of speaker, addressee, and plural features, with appropriate value specifications.

4.4 Intransitive constructions

Parallelling the double sets of transitive constructions are double sets of intransitive constructions.

4.4.1 'Actor Focus' intransitive constructions

One set of verbs has either an initial [*Cum* (if the Actor is singular) as in (22a), or [*maN*¹¹ (if the Actor is plural) as in (22b), and requires that the actor pronoun come from the Nominative set, as in (22b,c). Antipassive constructions occur in which the notional 'object' is indefinite and is unmarked as in (22d,e). The verb of an antipassive construction usually has initial [*maN*, but with some verbs, such as 'eat', an initial [*Cum*.¹² There is also an intransitive, stative verb, analysed by Topping as an agentless 'passive', with initial [*ma*, as in (22f).

(22) Chamorro (Topping 1973:83, 84, 86, 121, 186, 258 my analysis)

- a. *Gumupu i paluma.*
 flew Det bird
 -trns Nom PAT
 -plrl Nom
 actr
 'The bird flew.'
- b. *Manggupu siha.*
 fly they
 -trns PAT
 +plrl Nom
 actr
 'They are flying.'
- c. *Sumasaga yo' giya Agaña.*
 living I in Agaña
 -trns PAT Lcv LOC
 -plrl Nom Lcv
 actr
 'I live in Agaña.'
- d. *Manli'e' yo palao'an.*
 see I woman
 -trns PAT MNS (?)
 -plrl Nom Obl
 actr
 'I saw a woman.'

¹¹ The N refers to an assimilating nasal, which in some environments results also in the deletion of the consonant to which it assimilated.

¹² Topping analysed antipassive constructions as transitive verbs with indirect objects.

- e. *Chumocho yo' mansanas.*
 eat I apple
 -trns PAT MNS (?)
 -plrl Nom Obl
 actr
 'I ate an apple.'
- f. *Malalalatde i patgon.*
 scold Det child
 -trns Nom PAT
 -plrl Nom
 +sttv actr
 'The child is being scolded.'

To this point, Chamorro looks very much like a Philippine language. The structures pretty much parallel what have been called 'Actor focus' constructions in descriptions of these languages. Upon closer examination, however, there are a considerable number of differences between the above structures and those typically found in Philippine languages. The first is that Philippine languages do not normally use [*maN* as a marker of plural subjects.¹³ Second, the base form of the antipassive [*maN* is [*man* in Chamorro, but [*maŋ* in Philippine languages. Third, there is no apparent evidence in Chamorro of a reflex of **maR-*, which has reflexes in most Philippine languages.¹⁴ Fourth, each of the sentences (22a-f) can only be realis constructions. Irrealis constructions have quite different structures in Chamorro, as we shall see in §4.4.2 below.¹⁵

4.4.2 'Non-focus' intransitive constructions

Unlike the transitive constructions, the 'non-focus' intransitive set is in complementary distribution with the 'focus' set. The former require the presence of one or more irrealis¹⁶

¹³ Tagalog and some other Central Philippine languages, such as Mamanwa (Miller & Miller 1976:53), however, do use a similar device for the optional marking of plural subjects. Compare: Tag *makakita* 'see' with *mangakakita* 'see (PL)', etc. (Schachter & Otnes 1972:335).

¹⁴ The possibility that PEF **maR-* may be irregularly reflected as Chm *man-*, as it is in Central Cordilleran languages of the Philippines occurred to me. However in the latter languages, PEF **R* is reflected as *l*, so the expected reflex of **maR-* is *mal-*. However in Chamorro **R* became *g*, so the expected reflex would be *mag-*.

¹⁵ Topping does not refer to these constructions as 'actor focus', but as examples of the 'verbalizing affix -*um-*' (Topping 1973:244). His 'actor focus' constructions require that the Nominative NP appear in initial position, using his 'emphatic' pronoun set if the reference is pronominal. He also notes that such constructions can have either a definite, or an indefinite 'object', as follows: *Guiya lumi'e' i palao'an*. 'He is the one who saw the woman.' and *Guiya manli'e' palao'an*. 'He is the one who saw a woman.' His translations suggest that the verb is nominalised. He does note, however, that 'there are good reasons to consider the verbalizing affix -*um-*...[to be] the same as the actor focus affix -*um-*' (Topping 1973:244). Zobel notes that constructions such as these cannot be 'actor focus'. He believes they are transitive constructions, and in sentences such as: *Hayi lumi'e' gue'?* 'Who saw him?', he says 'the object is marked by an Absolutive pronoun' (Zobel 2002:412). However, Zobel's analysis of them as 'active participles' throws little light on their syntactic status. The analysis of such constructions is problematic. I suspect that they may be recent innovations, possibly influenced by the grammar of European languages.

¹⁶ Topping analyses these forms as future tense markers.

adverbs, *para*, *bai*¹⁷ or *u*, while the latter are always realis and are interpreted as either present or past depending on context or the presence of time adverbs. As with 'non-focus' transitive constructions, the verbs do not carry any 'focus' marking, as in (23a,b).

(23) Chamorro (Topping 1973:263)

a. *Para husaga giya Yigo.*

will stay at Yigo

-trns Lcv LOC

?1SG Lcv

?actr

'I will stay in Yigo.'

b. *Para tafañocho gi ega'an.*

will eat in morning

-trns Lcv LOC

?1PL Lcv

?actr

'We will eat in the morning.'

4.5 Transitive and intransitive constructions: inherited, innovated or borrowed?

Zobel outlines a number of innovations which he claims have either taken place in Chamorro, or are (exclusively) shared by Chamorro with Palauan or one or more of the set of languages that constitute his NMP. Most of these apparent innovations are found in the forms of the constructions discussed in the previous section. This section will briefly examine each of Zobel's claims.

4.5.1 Development of Actor-agreement markers

Zobel (2002) claims that the differentiation of his set A pronouns from the possessive suffixes is an innovation shared by Palauan, and the presence of 'prefixed pronouns' is restricted to a set of languages that 'excludes the Philippines, Northern Sulawesi, and Borneo (except for Malayic and Tamanic)'. As I showed above, however, there are some languages in the Philippines which have independently developed pre-verbal pronoun clitics, and under the identical conditions that Zobel claims for the NMP languages that have them. So the possibility exists that they were independently innovated in Chamorro. However, the forms of the Chamorro Actor-agreement markers are similar in some respects to those that occur in Palauan, as seen from Table 4.

¹⁷ Probably from Spanish *voy* 'I go' (Topping 1973:262). It is often omitted in casual speech, but when it occurs it always follows *para*, and co-occurs only with verbs that carry first person actor agreement.

Table 4: Chamorro agreement markers

	Chamorro Actor-agreement	Palauan Actor-agreement
1SG	[<i>hu</i>	[<i>k</i>
2SG	[<i>un</i>	[<i>?om</i>
3SG	[<i>ha</i>	[<i>l</i>
1PL.INC	[<i>ta</i>	[<i>d</i>
1PL.EXC	[<i>in</i>	[<i>kim</i>
2PL	[<i>en</i>	[<i>?om</i>
3PL	[<i>ma</i>	[<i>l</i>

The nature of the similarity between the forms is perhaps revealing of their origins. Zobel claims that the Chamorro forms are 'related to' the Palauan forms. But this is only obvious of the 1SG forms, both from **ku*, and the 1PL.INC forms, from **ta*. Each of these forms is inherited from PEF. None of the other forms are clearly related. Palauan [*l* '3SG, 3PL'] is a reflex of **na* '3SG', while Chamorro [*ha* '3SG' (and also [*uha* '3PL, irrealis')] is probably from **da* (PEF **d* > Chm *h*), reconstructable as an adverbial clitic with the meaning 'now, already', and with reflexes in a number of Philippine (and Formosan?) languages. The fact that it does not co-occur with the irrealis marker of 3SG (and is optionally absent with 3PL) of intransitive verbs (see Table 5) suggests that it was probably not a pronominal form in Chamorro originally, and that there was no overt marking for either 3SG or 3PL on intransitive verbs, a feature of early Austronesian morphosyntax.

We noted above that the source of Chm [*ma* '3PL'] is the stative verb marker [*ma*. A closer look at the Palauan forms, however, is suggestive of other types of relationship. According to Zobel, the Palauan forms occur only when the verbs which they precede are imperative, so it is not surprising then that the Palauan 2SG and 2PL forms are identical, [*?om*. It is possible that this is a truncated form of an 'auxiliary' verb, such as PEF **?umay* 'go', which did not require the explicit mention of second person pronouns in imperatives. In the northern Philippine language Bontok reflexes of **?umay* 'go' are split between a regular intransitive verb, as in (24a), and a transitive 'auxiliary' verb carrying Actor-agreement marking, as in (24b), occurring in imperatives.¹⁸ A similar situation may have existed in Pre-Palauan.

(24) Bontok (Reid)

- a. *?umeyka!* 'You go!'
 b. *?umeymu ?alá?en!* 'You go get (it)!'

But what about the Chamorro 2SG and 2PL forms? At least the 2SG form [*un* may have the same source as Pal [*?om*. But are these independent innovations, or the result of contact between the two languages, or evidence for some subgrouping relationship? There is some evidence that contact (probably with Palauan) is the explanation for the Chamorro forms. That they are not inherited is clear from the fact that none of the vowel-initial forms underwent the sound change by which a prothetic *g* or *gu* developed in this environment, as on the free pronouns *guahu* from **aku* '1SG' and *guiya* from **iya* '3SG'. The contact must have taken place after that change was no longer operative. The final *n* in each of the

¹⁸ In Ilokano a similar split has occurred between intransitive *mapan* 'to go' and a transitive 'auxiliary' verb *mapan*, as in *Mapanmo alaen!* 'Go get it!' (Carl Rubino pers. comm., my analysis).

consonant final Chamorro forms is probably the result of levelling of the final consonant to an alveolar nasal after their introduction into the language, but the motivation for this change is not apparent. One further speculation related to Chm [*un* '2SG' is that this form is probably the source of the Chamorro irrealis marker [*u*, which has spread from [*un* to mark also verbs with 1PL.INC, 3SG and 3PL agreement markers as irrealis. See Table 5 below.

Table 5: Chamorro Actor-agreement markers on irrealis intransitive verbs

1SG	[<i>hu</i>
2SG	[<i>un</i>
3SG	[<i>u</i>
1PL.INC	[<i>uta</i>
1PL.EXC	[<i>in</i>
2PL	[<i>en</i>
3PL	[<i>uha</i>]/[<i>u</i>]/[<i>uma</i>

4.5.2 Development of [*maN* as marker of antipassive verbs

Zobel (2002) claims that the specialisation of [*maN* as a marker of antipassive verbs is restricted to Chamorro, Palauan and a set of his NMP languages, implying that this is not one of its functions in Philippine languages. However, as I have noted elsewhere (Reid 2000:34), its function as a marker of antipassive is common in Philippine languages, such as Ivatan (Reid 1966:34), compare (25a) and (25b), and Bontok, compare (26a) and (26b).

(25) Ivatan (Reid 1966:34)

- a. *Somali qo tao do vahay.*
 drop.by Nom man Loc house
 -trns Nom Loc
 'The man is dropping by the house.'
- b. *Manali qo tao so libro do vahay.*
 drop.by Nom man Obl book Loc house
 -trns Nom Obl Loc
 'The man is dropping by the house for a book.'

(26) Bontok (Reid)

- a. *Sak?en nan ?umála.*
 I Det getter
 +prdc Nom
 'I'll be the one to get (it).' (lit. 'I'll be the getter.')
- b. *Sak?en nan mañála=s nan ?ásu.*
 I Det getter =of Det dog
 +prdc Nom =Lcv MNS
 Lcv
 'I'll be the one to get the dog.' (lit. 'I'll be the getter of the dog.')

4.5.3 Cooccurrence of ‘actor focus’ and ‘non-actor focus’ verb affixation

Chamorro has verbs that are marked by more than one ‘focus’ affix. Topping (1973:253) refers to these as ‘multiple focus constructions’. Zobel (2002:414) notes that ‘the suffix *-i* is not a focus affix since it can combine with all the above-mentioned transitive affixes, giving *maN--i*, *<um₂> -i*, *<in> -i*, and *ma--i*.’ He describes it as a ‘symmetrisation of the focus system’. He claims it is found in all languages that have developed a verb-initial Actor-agreement system and that have not lost the affixes concerned, as well as in Balinese, Madurese, and Sundanese. Yet it is precisely this type of combination of affixation which is widespread in Western Austronesian, including Philippine languages and Paiwan¹⁹ in Taiwan, and which Zobel himself reconstructs for PMP, with sequences of affixes such as **maN--a*, **<um>-a*, **ka--i*, as well as **ma--an*, **mai-*, etc. The Chamorro forms may be retentions from PEF, they may be independently innovated, or they may be the result of contact.

4.5.4 Loss of tense-aspect involving [Cin

Chamorro does not use a reflex of **[Cin* to distinguish tense-aspect. This form still occurs in Chamorro, but according to Zobel (see footnote 8) is a passive verb marker. As noted above in §4.3, according to my analysis it marks a transitive verb with an Ergative Agent and a Nominative Patient; but as Zobel correctly notes, it does not mark either completive aspect, as it did in PEF, nor begun aspect, as it does in Central Philippine languages (Reid 1992). Zobel notes that this innovation, although not found in Palauan is common to ‘almost all’ of his NMP subgroup. However, it is not an exclusively shared innovation to this subgroup, as Zobel implies. Some Southern Philippine languages, such as Tboli, are like Chamorro in innovating the form as a marker of transitive verbs, regardless of tense or aspect. In (27a,b), the reflex of PEF **[Cin*, shown as verb initial [*nC* or [*Cn*, occurs in constructions that are clearly future.

- (27) Tboli (Porter 1977:134)
- a. *Tnaba-hu lemwo't kedeng.*
call-I leave later
‘I’ll call him when I leave later.’
 - b. *Nwit-en mulék kedeng.*
bring-he return later
‘He will bring it when he returns later.’

Some Manobo languages, such as Tasaday, have also lost, or are in the process of losing [*Cin* as a distinguisher of tense, as in (28a-c).

- (28) Tasaday (Reid 1993: sentences 17, 78, 115)
- a. *Kakay Dula, pinahuna ku.*
friend Dula first I
‘Friend Dula, I’ll be the first (to go).’

¹⁹ Starosta (pers. comm.).

- b. *Uman ka hinali? da? siya du.*
 again you move now it emph
 'Move it over there again!'
- c. *Pinehagtay de aken siya.*
 kill now I it
 'I'll be the one to kill it.'

4.6 Summary

The evidence adduced above suggests that each of the innovations proposed by Zobel for subgrouping Chamorro with his NMP languages, a group that excludes the languages of the Philippines, either has a parallel development in one or more of the Philippine languages, or could be reflexes of PEF forms, or could be the result of intimate contact between Chamorro and one or more of the languages with which it has interacted in the remote past.

In the following section I shall address some of the problems raised by Starosta's proposed subgrouping.

5 Starosta's morphological subgrouping hypothesis for Chamorro

Starosta (1995) proposes that Chamorro is one branch of a subgroup of Austronesian, which is labelled F3, while the other branch F4 comprises all Austronesian languages except Rukai, Tsou, and Saaroa. In other words, Chamorro reflects morphological innovations that are found in the latter three languages, but does not show evidence of morphological innovations that he claims developed in other Austronesian languages at later stages of development of the family. It is precisely the lack of these later innovations that forms the basis of his subgrouping argument, since, if they were present, Chamorro would be subgrouped at a later point in the tree. The developments that he claims are unique to Chamorro are irrelevant for subgrouping purposes, unless of course, they can be shown to also be shared with some other language or languages. And the presence in Chamorro of earlier innovations (at the F1 and F2 levels) are likewise irrelevant for subgrouping purposes, because for Chamorro they are retentions.

For Starosta, a morphological form can only be considered to be lost from a language if a reasonable explanation can be given of the processes by which the form was lost or traces of it can be found lying around as frozen forms in the lexicon. Even if such frozen forms can be found, an alternative explanation of borrowing is also often possible. Without either of these explanations, Starosta is required by his methodology to subgroup the language at a point in the tree prior to the innovation of the form.

Some of the forms that I shall claim had been innovated prior to the initial settlement of the Marianas by Pre-Chamorro speakers, but which do not appear in present-day Chamorro, include reflexes of the transitive verb endings **en*], **an*] and **a*], the transitive verb-initial **[?i*, **[paR* and its intransitive verbal counterpart **[maR*, and the causative verb-initials **[pa(ka)* and **[ka*, all of which were part of the inventory of affixes acknowledged by Starosta to already be present by the time PEF, his F9, began occupying the Philippines. The following sections will discuss each of these forms in turn.

5.1 Transitive verb endings **en]*, **an]* and **a]*

Although I claim that all of these forms have been lost as transitive verb endings, reflexes of the nominal counterparts of the first two both occur in the language, while the latter seems to have been lost without a trace. The reflex of **en]* in Chamorro is *on]* ‘capable of being V-d’ as in (29a-d). (Topping simply defines it as ‘capable of’). This, as Starosta notes, fits fairly well the expected semantics ‘thing to be V-ed’. A few examples apparently mean ‘capable of V-ing’ (the meaning cited by Starosta) as in (29d,e). However, I consider this to be a secondary development in Chamorro, with some forms, such as (29d), able to be interpreted either way. Starosta claims that these are verbs. However there is no proof of this. They occur in the examples as predicates, and could be either descriptive nouns or intransitive (adjectival) verbs. Topping, Ogo and Dungca (1975:159) provide one example which is clearly a noun: *i guasa'on* ‘something capable of being sharpened’.

(29) Chamorro (Topping 1973:181, Topping, Ogo & Dungca 1975:159)

- a. *punu'on* ‘can be killed’
- b. *kannu'on* ‘can be eaten, edible’
- c. *taitayon* ‘can be read, readable’
- d. *guasa'on* ‘can be sharpened, or sharpener’
- e. *falaguyon* ‘capable of running’

The nominal reflex of **an]* occurs in combination with the Chamorro reflex of the ‘distributive’ nominal initial **[paN*, as Chm *[faN...an]*. The result is a locative noun form, as in (30a-c).

(30) Chamorro (Topping 1973:180)

- a. *fanbinaduyan* ‘place abounding in deer’
- b. *fañochuyan* ‘eating place’
- c. *fano'makan* ‘shower, bathing place’

It also occurs in combination with *[Cin*, with a range of meanings that Topping characterises as ‘attributive’. They are nominalisations of what in Philippine languages would be completive, surface-affected ‘locative focus’ verbs, as in (31a,b).

(31) Chamorro (Topping 1973:180)

- a. *binesbusan* ‘having a skin rash’ (*busbus* ‘skin rash’)
- b. *minigu'an* ‘having secretion from the eyes’ (*mugu* ‘secretion from eyes’)

Of course, Chamorro is not unique in losing much of its verb-final transitive marking. Tboli, in the Southern Philippines subgroup, has completely lost all its verb suffixing morphology. The only traces in that language of **en]* are the transitive Actor-agreement markers, *em]* ‘2SG’ and *en]* ‘3SG’, which were transferred by an analogical ‘word formation strategy’ to also mark 2SG and 3SG possessor agreement on nouns.

5.2 Transitive verb-initial **[?i* and intransitive verb-initial **[maR*

The first of these two forms seems to have been lost without trace, unless the sequence reconstructed as PEF **[ma?i* is the source of Chm *[mí* ‘have lots of’, as it is in a number of Philippine languages, including Ifugao and Inibaloi. This is perhaps supported by the fact that the Chamorro form always carries primary stress, and has a long vowel. PEF **[maR* would be

reflected as Chm [mak before consonant-initial forms. There are at least two Chamorro words where this initial may be present as a frozen form. They are (32a,b). They cannot be reflexes of *[maka because *k regularly became Chm h between vowels, and either Chm ʔ or zero at the end of a syllable (Costenoble 1940:32-33).

- (32) Chamorro (Topping, Ogo & Dungca 1975:131)
- a. *makmata* 'to waken' (*mata* 'eye')
 - b. *maknganiti* 'devil, satan' (also: *manganiti, aniti* 'devil, satan')

A further possibility is that some of the Chamorro forms with initial [ma are actually reflexes of *[maR, but with irregular loss of the reflex of *g. One of the functions of PEF *[maR was to create an intransitive verb from a noun. In Chamorro, with many forms, [ma serves this purpose, as in (33a-b).

- (33) Chamorro (Topping 1973:226)
- a. *macho'cho* 'to work' (*cho'cho* 'work, job, employment')
 - b. *mata'chung* 'to sit' (*ta'chung* 'seat')

The reflex of the nominal counterpart of PEF *[maR in Chamorro, *[paR, should be [fak. This form does not occur, but I suspect that Chm [fa' 'pretend to, change' (with a final glottal stop) had its origins here, since according to Costenoble (1940) final voiceless stops became glottal stops in some dialects. This was a narrowing of the verbalising function that is illustrated above in (33), and may have also been influenced by a causative [fa which is now lost. See (34a,b).

- (34) Chamorro (Topping 1973:176)
- a. *Hu fa'bentana i petta.* 'I changed the door into a window.'
(lit. 'I windowed the door.')
 - b. *Ha fa'bunita gue' i palao'an.* 'The woman pretended to be pretty.'
(lit. 'The woman prettied herself.')

5.3 Causative verb-initials *[pa(ka) and *[ka

There are a number of words in Chamorro that have a frozen [fa' initial, the meanings of which, according to Topping, are unpredictable. The meanings of some, however, suggest that at some earlier point the affix may have been a causative, as in (35a-c). If this is correct, then it would appear that the reflex of PEF *[paka was Chm [fa', and that it fell together with the reflex of PEF *[paR.

- (35) Chamorro (Topping 1973:177)
- a. *fa'aila* 'tell on'
 - b. *fa'na'gue* 'teach'
 - c. *fa'nu'i* 'show'

Although there are two distinct initial [há markers in Chamorro which could be reflexes of *[ka, neither has a causative sense. Neither do the (borrowed) [ka initial forms have a causative meaning.

Chamorro causatives are now formed with initial [na', as in (36a). This form occurs on the left edge of verbs which, in some cases, carry the marking of earlier derivation, such as [ma, as in (36b), supporting the hypothesis that these verbs are the result of the fusion of the

Chamorro verb *na'i* 'to give', with a following verb, a commonly found grammaticisation producing causatives in languages around the world.

(36) Chamorro (Topping 1973:247)

- a. *na'gasgas* 'made clean'
- b. *na'malangu* 'made sick'

5.4 Summary

I have attempted to show in the above sections that Zobel's subgrouping hypothesis for Chamorro is probably flawed because the innovations that he cites are frequently not shared exclusively with the subgroup to which he claims Chamorro belongs. In addition there is evidence that some of the changes that have taken place in Chamorro could well be the result of contact, or of parallel innovation. There are problems with Starosta's subgrouping hypothesis because he considers the absence of some morphosyntactic features in Chamorro to be evidence for a very early split from the other Austronesian languages of Formosa, prior to the development of those features. I have attempted to show that there is evidence in Chamorro that these features were present in Pre-Chamorro but have been lost in the modern language. In the following section I shall try to justify another subgrouping hypothesis that brings into consideration some facts about the language that have not so far been considered.

6 Another hypothesis regarding position of Chamorro in Austronesian

I shall claim in this section that Chamorro is probably a first-order branch of Proto Extra-Formosan. In other words, the first Austronesians to settle the Marianas probably sailed from somewhere in the (northern) Philippines, prior to the differentiation of Philippine languages into the various subgroups that are attested today. Chamorro is clearly not a Philippine language, that is, it does not immediately subgroup with any of the branches of Austronesian in the country. This we know, not from morphosyntax, but from the phonology. All Philippine languages have collapsed the reflexes of **d* and **z* (Zorc 1987), whereas Chamorro has kept them distinct, **d* is reflected as Chm *h*, while **z* is reflected as Chm *ch* [ts].²⁰ There is also lexical evidence that Chamorro does not immediately subgroup with the Philippines. Chamorro, for example, reflects PEF **siwa* 'nine' as Chm *sigua* 'nine', while all Philippine languages reflect one or another of two innovations, either **siyam* or **siyaw*. Several pieces of morphosyntactic evidence suggest that Chamorro reflects innovations that took place in Proto Extra-Formosan, probably after the initial settlement of the Philippines, but prior to the dispersal of Philippine languages.

²⁰ However, Blust (pers. comm.) states:

[the distinction is maintained in] Pamaona of central Sulawesi..., where **d* > r, but **z* > j, Proto Bungku-Tolaki of southeast Sulawesi..., where **d* > r, but **z* > s, and Muna, of Muna and Buton islands in extreme SE Sulawesi, where **d* > r, but **z* > s. I'm sure there are other languages in Sulawesi which preserve the distinction... I might add that it is not altogether clear that Inati in Panay [Central Philippines] has merged these phonemes. Since Proto Philippines preserved the PMP distinction between alveolar and palatal nasals, there is a clear implication derived from general typological studies that the PPh phonological system would have included a palatal obstruent.

6.1 Nominative marking

I have claimed elsewhere (Reid 1978, 1979) that the parent language of the Philippine group had a set of Nominative determiners of the form **i* ‘common noun’, and **si* ‘personal noun’. The common noun determiner does not occur as a Nominative marker in Formosan languages, where it must be reconstructed as a non-Nominative determiner, as well as a formal marker of definite predicate nouns, and of topic phrases. It is its latter function of marking definite NPs that probably motivated its spread to also mark Nominative NPs since these are always necessarily also definite. In the Philippines there is no language that maintains a reflex of **i* as its sole marker of Nominative common noun phrases, but there is clear evidence from the forms of free Nominative pronouns and demonstratives which often have an initial [i, that this was its function at earlier stages of the family.

In Chamorro, Nominative definite common noun phrases are marked by *i*, as in (21a), repeated below as (37). This is not the only function that *i* has, it retains also the earlier general function of marking definite common noun phrases, as in the ‘non-focus’ construction (21e) repeated below as (37b).

(37) Chamorro (Topping 1973:245, my analysis)

- a. *Lini'e' ni lahi i palao'an.*
 see Det man Det woman
 +trns Erg AGT Nom PAT
 Erg Nom
 actr
 ‘The man saw the woman.’
- b. *Hali'e' i lahi i palao'an.*
 see Det man Det woman
 +trns +dfnt AGT +dfnt PAT
 ?3SG Erg Nom
 ?actr actr
 ‘The man saw the woman.’

Although the generalisation of the function of **i* to include common noun Nominative NPs is a feature that was apparently an innovation in PEF, and appears to be inherited in Chamorro, it is possible that it could be an independent innovation in this language. The fact that this form did not acquire a prothetic *g* in Chamorro (see below), implies that it was probably an enclitic, as it is in Inibaloi, Pangasinan and other Philippine languages that have retained the form.

Most Chamorro personal noun phrases are marked by *si*. Only ergative personal nouns are distinguished, being marked by *as*.

6.2 Genitive marking

Chamorro possessive constructions are outlined in §4.1. Only the construct form, which I claim is probably a borrowed form, maintains a relic of the initial alveolar nasal of PEF **(n)i* ‘genitive determiner’, that is the construct *n]* itself. Genitive noun phrases in Chamorro that occur in constructions that I claim are probably inherited are marked either by *i*, if a common noun, or by *si* if a personal noun, as in (13a,b) repeated below as (38a,b).

- (38) Chamorro (Topping 1973:223)
- a. *i gima'ña si Rosa* 'Rosa's house'
 - b. *i hagaña i rai* 'the king's daughter'

In PEF, actors of transitive constructions were also genitively marked. However in Chamorro as discussed above a distinct set of agreement markers developed from what were originally genitive clitic pronouns, while full noun phrases in such constructions are marked by *ni*. Topping states that this form has developed from the sequence Chm *nu i*, and is often pronounced as such. It is probable that the *ni* pronunciation is the result of influence from speakers of Philippine language in which genitive common nouns are so marked.

The Chamorro ergative noun phrase marker *nu i*, is of interest because it suggests that Pre-Chamorro acquired at some early stage (or inherited from PEF) a set of *u*-grade determiners, and that *nu* was the genitive common noun marker from this set, as it is in several Formosan languages, such as Tsou, Saisiyat, Paiwan, and Amis, and Philippine languages such as Yami, Ivatan, Itbayat, Casiguran Dumagat, Umiray Dumagat, and Ilongot, as well as in Murut in Borneo (Reid 1978:54).

6.3 Topic marking

The evidence that Chamorro possibly had a series of *u*-grade determiners becomes important when considering the source of the prothetic *g* and *gu* on otherwise vowel-initial forms in Chamorro. I shall argue in this section that Pre-Chamorro developed from a language that probably had **ʔu* as a topic marker, and that this form became a proclitic [w on certain nouns.

It has long been recognised (Costenoble 1940:39, 54) that the Chamorro labio-velar, represented in written Chamorro as *gu*, is the regular reflex in inherited words of **w* before vowels, both initially and medially, for example, Chm *gualu* < PEF **walu* 'eight', Chm *ʔasagua* < PEF **qasawa* 'spouse', Chm *sigua* < PEF **siwa* 'nine', Chm *pugua* < PEF **buwaq* 'fruit', etc. But [*gu* also appears as a prothetic element on a number of nouns that are not reconstructable with an initial **[w*, such as Chm *guihan* < PEF **ʔikan* 'fish', Chm *gui'eng* < PEF **ʔijung* 'nose', Chm *guafi* < PEF **ʔapuy* 'fire', and so on.

Words reconstructable with medial **[...aʔu* sequence are reflected as Chm [...*agu*, for example Chm *hagon* < PEF **daʔun* 'leaf', Chm *chaguʔ* < PEF **zaʔuq*. This change is probably independent of the change from PEF **w* to Chm *gu*. The latter change occurred also in Karaw in the Southern Cordilleran subgroup in the Philippines, while the former occurs also in Northern Kankanay²¹ in the Central Cordilleran subgroup. In Chamorro the change apparently spread to the beginning of words so that in many forms **ʔ* became *g* before any vowel, for example Chm *gugat* < PEF **ʔuRat* 'vein, tendon', Chm *gunom* < PEF **ʔenem* 'six', and Chm *gi* < PEF **ʔi* 'locative preposition'.

As I noted above, the [*gu* reflex occurs on a number of nouns that did not originally have an initial **[w*. We find the same reflex, for example, on otherwise vowel-initial independent pronouns and demonstratives. Note Chm *guahu* '1SG' < Pre-Chm **waku* < PEF **aku*, and Chm *guiya* < Pre-Chm **wiya* < PEF **(s)iya* '3SG'. Similarly, there are a set of 'static locative' forms which have an initial [*gu* where there was no initial **[w*. Compare the locative demonstratives with the demonstrative pronouns: Chm *guini* 'here in this place', but *ini* 'this'

²¹ In Northern Kankanay, **ʔ* became *ɣ*, not only between *a* and *u*. The reflex occurs between any sequence of vowels where a glottal stop would have occurred.

< PEF **ni* ‘this’; *guenao* ‘there towards you, in that place’, but *enao* ‘that’ < PEF **na* ‘that’; and *guihi* ‘there, that place’, but *ihi* ‘that’ < PEF **di* ‘that’. The locative demonstratives must have all been marked originally with the PEF **ʔi* ‘locative preposition’, but acquired an initial *[w: **wini* ‘here’, **wina* ‘there, close to hearer’, and **widi* ‘there, away from speaker and hearer’.

With evidence that early Chamorro probably had a set of *u*-grade determiners, the source of the initial [gu- in the above forms becomes apparent. The pronominal evidence suggests that they acquired an initial [w in the environment in which they occurred, that is, as fronted, topicalised forms (or perhaps as sentence initial predicate nominals), marked by a topic-marking determiner **ʔu*. This determiner then became attached as a proclitic to locative demonstratives and also to a limited set of commonly occurring nouns which frequently would have occurred in topicalised environments. The languages that today show a reflex of **ʔu* as a topic and/or predicate marker are limited. Amis is one such language, where the same process of [w attachment to some nouns, such as *waco* ‘dog’ also occurs, as in (39). There is some evidence that it also occurred in the parent of the Batanic languages. Yami *ʔu* marks both topic and predicate common nouns, while in Ivatan the same form marks predicate common nouns.

- (39) Amis (Chen 1987)
O nomako kina waco. ‘The dog is mine.’

The evidence suggests, then, that Proto Extra-Formosan used **ʔu* as a topic and/or predicate marker, and that Chamorro reflects it in the form of its [gu initial pronouns, demonstratives and other nouns which did not originally have an initial [w.

One small piece of additional evidence for the early split of Pre-Chamorro from Proto Extra-Formosan is the Chamorro pronoun *hagu* ‘2SG’. This form ultimately derives from PAn **kaSu* > PEF **kahu*. In Philippine languages, this pronoun developed as **kaw* (as in Tag *ikaw* ‘2SG’). However in Chamorro, the presence of a medial [...g clearly shows that PEF **kahu* developed as Pre-Chm **kaʔu*. A similar development of PAn **S* can be seen in Chm *hugua* < Pre-Chm **duwa* < PEF **duha* < PAn **dewSa* ‘two’.

6.4 ‘Ligature’ **na*

Like many other western Austronesian languages, Chamorro has a form commonly referred to in the literature as a ‘ligature’ that occurs between a head, whether noun or verb, and its modifier. Its actual form class is probably either noun or preposition, but I shall continue to refer to it here as a ligature. In Chamorro the form is *na* as in (40a-c).

- (40) Chamorro (Topping 1973:138, 149)
- a. *i díkike’ na patgon*
 Det small LIG child
 ‘the small child’
 - b. *ayu na lepblo*
 that LIG book
 ‘that book’
 - c. *Hutungo’ na machocho’cho’ i lahi.*
 1SG.know LIG working Det man
 ‘I know that the man is working.’

For the parent language of the Philippines I have reconstructed **(n)a* as the ligature, with the full form **na* occurring following vowels, and the shorter form **a* occurring following consonants. These are the forms that are commonly found in the languages of the Northern Philippines, as well as in a number of Formosan languages, although some languages have generalised one or the other to occur in all environments losing the alternate form, as Chamorro has apparently done. **(n)a* must also be reconstructed for Proto Extra-Formosan. The Central Philippine languages, including Pre-Tagalog, and languages to the south of the Philippines underwent an innovation which changed initial **[n* of **na* to a velar nasal, resulting in ligature **ŋa*. In many of these languages the form was again reduced to become final *ŋ*] on Determiners, such as Tag *ang* and *nang*, Malay *yang*, and so on. It is clear that Chamorro did not share in this particular innovation, placing its separation from the rest of the family at a point prior to the development of the innovation.

6.5 Antipassive **[maN* and plural **[maN*

The Chamorro forms of the antipassive **[maN* and plural **[maN* both end in a nasal which assimilates to the point of articulation of the following consonant, and, under certain conditions, results in the deletion of the following consonant. These processes are commonly found throughout languages of the Philippines and Indonesia, but are not generally found in Formosan languages, although there is evidence they occur in Amis, and there are sporadic occurrences of verbs in other Formosan languages that have been claimed (Blust 1999:68) to be frozen remnants of *[maN*, such as Puyuma *mangayaw* 'to hunt heads', but which could also be borrowings from Amis, or even from some Philippine language. What interests us about the Chamorro forms and what distinguishes them from similar forms in other western Austronesian languages is the fact that the base form of the Chamorro initial is not *[maŋ* with final velar nasal, as it is in other languages, but *[man* with final alveolar nasal. This seems like a trivial point, but when we recognise that there is probably a historical relationship between the plural and antipassive forms, and that the pluraliser in Philippine languages, as for example, Tag *mga* (/maŋa/), consists of a sequence of *ma* plus ligature *nga*, then the source of the Chamorro forms becomes clear. They are probably both ultimately from a form *ma* plus a ligature *na* with loss of the final vowel. This must have been the Proto Extra-Formosan source of both the antipassive, as well as the plural forms. In Philippine languages, PEF **ma na Noun* 'plural noun' became **maŋa Noun*, after the velar nasal initial was innovated, as noted in the previous section.

7 Conclusion

In this paper I have tried to show that much of the evidence that Zobel has produced to support his claim that Chamorro is most closely related to a subgroup that excludes the Philippine languages, is suspect because it is not exclusively shared by languages in his Nuclear Malayo-Polynesian subgroup. Furthermore, I have claimed that just as Chamorro shows the effects of extensive contact in at least two separate layers of lexicon, so it also shows the effects of contact in its morphosyntax. Although it is known that there was extensive contact between the Philippines and the Marianas, there is little that one can specifically point to in the morphosyntax that is clear evidence of that contact, and which could not, for example, be the result of contact with some language south of the Philippines, just as the Chamorro terms *babui* 'pig' and *mannok* 'chicken' are clearly borrowings because

they do not exhibit the regular sound changes that characterise inherited vocabulary, but are non-specific as to the language from which they must have come.

It is tempting to see in Chamorro *munga* 'prohibitive negative', as shown (41), a borrowing from a Tagalog structure such as (42) with loss of the original negative verb, *huwag*, and retention of the genitive pronoun *mu* '2SG' and an enclitic 'particle', *nga*, which changes a command to a polite request. However it may not be related at all, and if it were borrowed, then it might also have come from some other Central Philippine language.

- (41) Chamorro (Topping 1973:138,149)
Munga humanao! 'Don't go!'
- (42) Tagalog (Schachter & Otanes 1972:405,523)
Huwag mo nga! '(Please) don't (do it)!'

Similarly, it is tempting to see in the Chamorro term *palao'an* 'woman', a clue of a prehistoric time when women were brought from Palau. The term seems to be borrowed; there is no other known etymology for it. If so, then it would not be surprising to find extensive influence on the language from the fact that it would probably have been the language of Palau that children learned at their mother's knee.

Determining the actual subgrouping position of Chamorro is not possible from phonological evidence, and even the morphosyntactic evidence that I have outlined in this paper is not strong, and is perhaps open to other interpretations. However, it seems to point to two facts. The first is that Chamorro is not most closely related to the Formosan languages, because there are probably archaic remnants in Chamorro of innovations that Starosta claimed were innovated at points later than the time when he would have it separating from those languages. The other fact is that Chamorro seems not to have participated in certain innovations that are found universally in the Philippines. I conclude, therefore, that Chamorro is an Extra-Formosan language, but that it is a first-order branch of the family, separating from Proto Extra-Formosan, probably from the Northern Philippines, prior to the actual dispersal of the other branches of the family.

References

- Ballard, Lee D., Robert J. Conrad and Robert B. Longacre, 1971, *More on the deep and surface grammar of interclausal relations*. Language Data: Asian-Pacific Series 1. Santa Ana, California: Summer Institute of Linguistics.
- Blust, Robert, 1997, Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun and Paul Jen-Kuei Li, eds *Selected papers from the Eighth International Conference on Austronesian Linguistics*. Symposium series of the Institute of Linguistics (Preparatory Office). No.1:31–94. Taiwan, Taipei: Academia Sinica.
- Chen, Teresa, 1987, *Verbal construction and verbal classification in Nataoran Amis*. Canberra: Pacific Linguistics.
- Cooreman, Ann M., 1987, *Transitivity and discourse continuity in Chamorro narratives*. Berlin: Mouton de Gruyter.
- Costenoble, Hertha, 1940, *Die Chamoro Sprache*. The Hague: M. Nijhoff.
- Egerod, Søren, 1966, Word order and word classes in Atayal. *Language* 42:346–369.

- Gibson, Jeanne D., 1990, Categorical grammatical relations: the Chamorro evidence. In Paul M. Postal and Brian D. Joseph, eds *Studies in relational grammar*, 3:247–260. Chicago: University of Chicago Press.
- Hidalgo, Cesar A. and Araceli C. Hidalgo, 1971, *A tagmemic grammar of Ivatan*. Philippine Journal of Linguistics, Special Monograph No.2. Manila: Linguistic Society of the Philippines.
- Kikusawa, Ritsuko, 1997, A re-examination of the grammatical ergativity of Tongan. Paper presented to the *Third International Conference on Oceanic Linguistics (TRICOL)*, University of Waikato, New Zealand.
- Miller, Jeanne and Helen Miller, 1976, *Mamanwa grammar*. Language Data: Asian-Pacific Series No.8. Huntington Beach, California: SIL.
- Porter, Doris, 1977, *A Tboli grammar*. Philippine Journal of Linguistics, Special Monograph No.7. Manila: Linguistic Society of the Philippines.
- Reid, Lawrence A., 1966, *An Ivatan syntax*. Oceanic Linguistics Special Publication No.2. pp.160. Honolulu: University of Hawai'i Press.
- 1978, Problems in the reconstruction of Proto-Philippine construction markers. In Stephen A. Wurm and Lois Carrington, eds *Proceedings of the Second International Conference on Austronesian Linguistics. Fascicle I. Western Austronesian*, 33–64. Canberra: Pacific Linguistics.
- 1979, Evidence for Proto-Philippine nominative marking. *Philippine Journal of Linguistics* 10/1-2:1–20.
- 1992, On the development of the aspect system in some Philippine languages. *Oceanic Linguistics* 31/1:65–92.
- 1993, Translation of the 1972 Tasaday cave Tape 2. <http://aa2411s.aatufs.ac.jp/~reid/Tasaday/Early-Tapes.html>
- 2000, Sources of Proto-Oceanic initial prenasalization: the view from outside Oceanic. In Videia P. de Guzman and Byron W. Bender, eds *Grammatical analysis: morphology, syntax and semantics. Studies in honor of Stanley Starosta*, 30–45. Honolulu: University of Hawai'i Press.
- 2001, On the development of agreement markers in some northern Philippine languages, 235–258. In Joel Bradshaw and Kenneth L. Rehg, eds *Issues in Austronesian morphology: a focusschrift for Byron W. Bender*.
- Schachter, Paul and Fe Otnes, 1972, *Tagalog reference grammar*. Los Angeles: University of California Press.
- Starosta, Stanley, 1995, A grammatical subgrouping of Formosan languages. In Paul Jen-kuei Li, et al., eds *Austronesian studies relating to Taiwan*, 683-726. Symposium series of the Institute of History and Philology, No.3. Taipei: Academia Sinica.
- Starosta, Stanley and Louise Pagotto, 1991, The grammatical genealogy of Chamorro. In Ray Harlow and Robin Hooper, eds *VICAL 2: Papers from the Fifth International Conference on Austronesian Linguistics*, 319–348. Te Reo Monographs, Auckland: Linguistic Society of New Zealand.
- Starosta, Stanley, Andrew K. Pawley and Lawrence A. Reid, 1982, The evolution of focus in Austronesian. In Stephen A. Wurm and Lois Carrington, eds *Papers from the Third International Conference on Austronesian Linguistics*, Vol.2. *Tracking the travellers*, 145–170. Canberra: Pacific Linguistics.
- Tchekhoff, Claude, 1991, Two symmetrical verbal aspects in Oceania. In Robert Blust, ed. *Currents in Pacific linguistics: papers on Austronesian languages and ethno-linguistics in honour of George W. Grace*, 501–508. Canberra: Pacific Linguistics.

- Topping, Donald M., 1973, *Chamorro reference grammar*. Honolulu: The University Press of Hawai'i.
- Topping, Donald M., Pedro M. Ogo and Bernadita C. Dungca, 1975, *Chamorro-English dictionary*. PALI Language Texts: Micronesia. Honolulu: The University Press of Hawai'i.
- Tsuchida, Shigeru, 1976, *Reconstruction of Proto-Tsouic phonology*. Monograph Series No.5. Institute for the Study of the Languages and Cultures of Asia and Africa. Tokyo: Tokyo University of Foreign Studies.
- Zobel, Erik, 2002, The position of Chamorro and Palauan in the Austronesian family tree: evidence from verb morphology and morphosyntax. In Fay Wouk and Malcolm Ross, eds *The history and typology of western Austronesian voice systems*, 405–434. Canberra: Pacific Linguistics.
- Zorc, R. David, 1987, Austronesian apicals (*dDzZ) and the Philippine non-evidence. In Donald C. Laycock and Werner Winter, eds *A world of language: papers presented to Professor S.A. Wurm on his 65th birthday*, 751–761. Canberra: Pacific Linguistics.
- Zwicky, Arnold M., 1977, *On clitics*. Indiana University Linguistics Club.
- 1985, Clitics and particles. *Language* 61/2:283-305.
- Zwicky, Arnold M. and Geoffrey K. Pullum, 1983, Cliticization vs. inflection: English *n't*. *Language* 59/3:502–513.

6 *A brief look at thirteen Mon-Khmer languages of Xekong province, southern Laos*

THERAPHAN L-THONGKUM

1 Introduction

Thirteen Mon-Khmer languages are spoken in Xekong province in southern Laos, viz., Alak (also called Harak, Harlak), Chatong, Dak Kang, Kaseng, Katu, Laven, Lavi, Nge', Suai, Ta Oi', Tariat, Tariw, and Yaeh (L-Thongkum 1997a).¹ Based on linguistic criteria, the Suai ethnic group as known by the local authorities and the inhabitants of Xekong are in fact two separate groups, that is the Suai of Tha Taeng speak a Katuic language, while the Suai of Yok Thong, Chunla, and Ka Pue villages speak a West Bahnaric language (L-Thongkum 1997b). Therefore, it would be more correct to say that fourteen Mon-Khmer languages are spoken by the population of Xekong. Because the Katuic Suai (Kui, Kuay) spoken in Thailand and Lao PDR has been studied by many linguists and language specialists and reliable data are available, I have not included it in my study. I discovered the fact that the Suai at Yok Thong village had called themselves Juk /ju:k/ during my fifth field trip in April 1998, when I worked extensively with my Suai informant from Yok Thong village. I also learned that the village name 'Yok Thong' did not mean 'raising flags' as I had been told by some local authorities but instead 'the lowland or paddy-field Juk'. This implies that in the past this West Bahnaric-speaking group had moved from the highlands somewhere to lowland areas to cultivate wet-rice, perhaps not for their own sake but for their Khmer landlords when the Khmer Kingdom had been in power and had conquered the whole area of

¹ The research project on 'Languages and Cultures of the Ethnic Groups in Xekong province, Lao PDR: A Foundation for Research and Development Planning', funded by the Thailand Research Fund (TRF), was officially launched in January 1997 after a brief field survey conducted in December 1995. Without the kindness of Her Royal Highness Princess Mahachakri Sirindhorn, this project could not have been possible. The output and outcome of this project will be regarded as our tribute to Her Royal Highness. I would like to express my gratitude to the Embassy of Lao PDR in Bangkok, the Ministry of Foreign Affairs, and the Ministry of Information and Culture of Lao PDR for their kind co-operation. Last but not least, I feel thankful to the local authorities of Xekong province for their help and assistance in making many arrangements.

southern Laos. This assumption of Khmer sovereignty is based on the fact that many old Khmer loans can be detected in Suai (Juk), whereas there are fewer such loans in the other Mon-Khmer languages of Xekong on which I have been working. All the data cited as examples in this paper come from my fieldnotes. I have made seven field trips (each of two to three weeks) since December 1995. A tentative classification of the thirteen Mon-Khmer languages is presented in this paper. The language data given as examples are in broad phonetic transcription.

2 Sociolinguistic setting

The number of people inhabiting Xekong is about 64,000. Although the population of Xekong is small, the sociolinguistic setting of the province is very complex. Lao is the official language of the province and the lingua franca of the Tha Taeng district, whereas Alak, Tariang and Nge' are those of Lamam, Dak Chueng, and Kaluem districts, respectively. Each inhabitant of this province can speak or understand at least two or three languages. The 'Lao Lum' or lowland Lao or even half Lao Lum is considered to be the most prestigious group; ethnic groups ranked next in prestige are the Alak and Tariang. As the result of language contact, linguistic borrowings cannot be avoided, especially in Lamam where the municipality of Xekong province is located, and in Tha Taeng which is the most recently established district where almost all of the ethnic groups belonging to Xekong can be found. In short, Xekong may be regarded as a paradise for sociolinguistic studies. See maps in the Appendix.

3 Language classification

Broadly speaking, the thirteen languages of Xekong can be divided into two major groups based on lexical and phonological developments; namely, Group A: Ta Oi', Katu, Nge' (Kriang), Chatong, Tariw, Dak Kang; and Group B: Suai (Juk), Laven (Jru'), Lavi (Savoeng), Tariang, Kaseng, Yaeh, Alak.

Comparison of several cognate lexical items in Tables 1 and 2 confirms the dichotomy of the two groups of languages. Both lexical and phonological evidence supports the idea that the six languages of Group A belong to one branch of Mon-Khmer, while the seven languages of Group B belong to another branch of the family.

4 Katuic languages

Based on the lexical examples given in the following two tables, there is no doubt in my mind that Ta Oi', Katu, Nge' (Kriang), Chatong, Tariw, and Dak Kang belong to one branch of the Mon-Khmer language family, and Suai (Juk), Laven (Jru'), Lavi (Savoeng), Tariang, Kaseng, Yaeh, and Alak belong to another branch. The next step is to try to place the languages of Groups A and B within the right branches. After searching through the literature of comparative Mon-Khmer to find appropriate criteria for identifying the Mon-Khmer branches, I believe the criteria of lexical and phonological innovations as proposed by

Diffloth (1996a,b) have proved to be applicable to and useful in my work.² My hypothesis has been confirmed: the Group A languages are Katuic, whereas the Group B languages are Bahnaric. Diffloth (1996a:11) has suggested that 'lexically, all Katuic languages have innovated in replacing the Proto Mon-Khmer numerals 'six', 'seven', 'eight', and 'nine' by a new lexical set'. The reconstructed forms of these numerals in Proto Katuic are: **təbat* 'six', **təbool* 'seven', **təgɔl* 'eight', and **təgɛs* 'nine'. Somewhat different forms have also been reconstructed by Peiros (1996): **tapat* 'six', **dəbuul* 'seven', and **dagiejh* 'nine'. Based on the Brou, Pacoh, and Katu languages of Vietnam, D.M. Thomas (1976) reconstructed the forms of these numerals as **pout* 'six', **tupol* 'seven', **tikool* 'eight', and **tikeas* 'nine'. Her reconstruction implies that the Proto Katuic voiced plosives **b*, **d*, **j*, **g* have devoiced to become voiceless in Proto East Katuic.

In the West Katuic languages, such as Kui, Kuay, Bru, So, and so on, the breathy voiced vowels (or second register vowels) suggest the intermediate stage of the devoicing process; see the numerals in Kui (the Suai language of Surin, Thailand): /thph̥aət/ 'six', /thph̥ool/ 'seven', /thkh̥uəl/ 'eight', /thkh̥əh/ 'nine'; and in Bru (of Khongchiam district, Ubonratchathani, Thailand): /tap̥ət/ 'six', /tap̥uul/ 'seven', /tak̥uəl/ 'eight', /take̥h/ 'nine', and so on. Based on the lexical and phonological developments illustrated in Tables 1 and 2, the six languages in Group A should be classified as East Katuic. There is no geographical contradiction.

Among the six East Katuic languages of Xekong, three are 'newly discovered languages', since Tariw, Dak Kang, and Chatong have never been mentioned previously in any publications on Mon-Khmer linguistics. One interesting feature worth pointing out is that the Proto East Katuic (PEK) preglottalised plosives or implosives **ɓ* and **ɗ* became plain voiced plosives *b* and *d* in Tariw and Dak Kang; and in Chatong which is a register language both voiced plosives and implosives **b*, **d*, **j*, **g*, **ɓ*, **ɗ* have merged to become voiceless plosives *p*, *t*, *c*, *k* and are followed by the breathy voiced or second register vowels. See examples in Table 3.

Besides the numerals 'six', 'seven', 'eight', and 'nine' (see Table 1), the more reliable Proto Katuic lexical innovations form a particular set of lexemes according to Diffloth (1996a). These lexemes and his Proto Katuic reconstructions are as follows: 'wife' **kəndɛal*, 'unmarried woman' **kumoor*, 'to smile' **kəcaŋ*, 'year' **kəmɔɔ*, 'cobra' **duur*, 'mushroom' **triəə*, and 'to keep' **dɔɔk*. I do agree with his conclusion that 'any Mon-Khmer language which has cognates in all or most of these words, as well as the numerals 'six', 'seven', 'eight', and 'nine' can be presumed to belong to the Katuic branch'.

² From 1996–98, Gérard Diffloth, who now lives permanently in Thailand, has visited me at the Thai Studies Institute, Chulalongkorn University, many times. Whenever we met, I showed him some parts of my Mon-Khmer data collected during my field trips in Xekong province, especially the lists of cognates as illustrated in Table 1 and Table 2, or discussed what I had found. This is one of the reasons why almost all of his proposals presented in the manuscript of his book 'An Introduction to Mon-Khmer Languages' from which he kindly gave me three chapters concerning the criteria for identifying languages belonging to the Katuic and Bahnaric branches were very useful to my Xekong project. In February 1998 Gérard Diffloth spent a few days working with my Tariw, Dak Kang and Suai (Juk) informants during his field trips in southern Laos.

Table 1: Classification of Xekong M-K languages into two groups based on lexical evidence

	'six'	'seven'	'eight'	'nine'	'to smile'	'tongue'	'navel'	'snake'	'head'	'skin'
Group A										
Ta Oi'	<i>ta'pat</i>	<i>ta'pool</i>	<i>ta'kool</i>	<i>ta'kiəjh</i>	<i>ka'caŋ</i>	<i>n'taaʔ</i>	<i>pon</i>	<i>ka'seŋ</i>	<i>plɿɿ</i>	<i>ŋ'kar</i>
Katu	<i>tɛ'pat</i>	<i>tɛ'pɔol</i>	<i>ta'kool</i>	<i>tɛ'kees</i>	<i>ka'caŋ</i>	<i>n'taak</i>	<i>pon</i>	<i>ka'seŋ</i>	<i>plaaw</i>	<i>ŋ'kar</i>
Nge'(Kriang)	<i>ta'piiət</i>	<i>tɛ'puul</i>	<i>ta'kool</i>	<i>tɛ'kieh</i>	<i>ka'caŋ</i>	<i>n'taak</i>	<i>pun</i>	<i>ka'seŋ</i>	<i>plɿɿ</i>	<i>ŋ'kar</i>
Chatong	<i>ta'pat</i>	<i>ta'pool</i>	<i>ta'kool</i>	<i>ta'kiejh</i>	<i>ka'caŋ</i>	<i>n'taak</i>	<i>pun</i>	<i>ka'seŋ</i>	<i>plɿɿ</i>	<i>ŋ'kar</i>
Tariw	<i>tɕa'pat</i>	<i>ta'pool</i>	<i>ta'kool</i>	<i>ta'kiəh</i>	<i>ka'caŋ</i>	<i>ʔan'taak</i>	<i>pun</i>	<i>ka'seŋ</i>	<i>plɿɿ</i>	<i>ŋ'kar</i>
Dak Kang	<i>tɕa'pat</i>	<i>ta'pool</i>	<i>ta'kool</i>	<i>ta'kiəʔ</i>	<i>ka'caŋ</i>	<i>taak</i>	<i>pun</i>	<i>ka'seŋ</i>	<i>plɿɿ</i>	<i>ŋ'kar</i>
Group B										
Suai (Juk)	<i>traw</i>	<i>thoh</i>	<i>thaam</i>	<i>ciin</i>	<i>ca'ʔaal</i>	<i>la'piat</i>	<i>klok</i>	<i>bih</i>	<i>tuuh</i>	<i>sruat</i>
Laven (Jru')	<i>traw</i>	<i>pɔh</i>	<i>thɿɿm</i>	<i>ciin</i>	<i>sa'ʔaal</i>	<i>piat</i>	<i>klok</i>	<i>bih</i>	<i>tuəjh</i>	<i>sruat</i>
Lavi	<i>traw</i>	<i>pɿh</i>	<i>thaam</i>	<i>ciin</i>	<i>sa'ʔau</i>	<i>ha'piet</i>	<i>klɔk</i>	<i>pih</i>	<i>tuəjh</i>	<i>srɔt</i>
Tariang	<i>ta'raw</i>	<i>ti'pɿh</i>	<i>taŋ'haam</i>	<i>ka'ceen</i>	<i>jaal</i>	<i>ha'piat</i>	<i>klɔk</i>	<i>bis</i>	<i>tuus</i>	<i>ruat</i>
Kaseng	<i>ta'raw</i>	<i>pa'pɿh</i>	<i>ta'haam</i>	<i>ka'ceen</i>	<i>ca'ʔaal</i>	<i>piat</i>	<i>klok</i>	<i>biç</i>	<i>tuuç</i>	<i>ʔa'ruat</i>
Yaeh	<i>cə'raw</i>	<i>puɿh</i>	<i>thaam</i>	<i>ciin</i>	<i>ʔaal</i>	<i>piat</i>	<i>kɔk</i>	<i>bih</i>	<i>tuuh</i>	<i>sruat</i>
Alak	<i>ta'raw</i>	<i>ti'pɔh</i>	<i>thaam</i>	<i>ta'ciin</i>	<i>ca'ʔau</i>	<i>ha'peet</i>	<i>klɔk</i>	<i>bih</i>	<i>(gaw)</i>	<i>(ʔa'kar)</i>

Table 2: Classification of Xekong M-K languages into two groups A based on phonological evidence

	'bone'	'pain'	'I'	'thigh'	'moon'	'liver'	'earth'	'fire'
Group A								
Ta Oi'	<i>ŋ'haaŋ</i>	<i>ʔaj</i>	<i>kuuu</i>	<i>paɾ'law</i>	<i>ka'saj</i>	<i>lɔɔm</i>	<i>ka'tɛɛʔ</i>	<i>ʔuujh</i>
Katu	<i>ŋ'haaŋ</i>	<i>ka'ʔaj</i>	<i>ʔa'kɽɽ</i>	<i>pɛ'law</i>	<i>ka'saj</i>	<i>lɔɔm</i>	<i>ka'tiiək</i>	<i>ʔuus</i>
Nge' (Kriang)	<i>ŋ'haaŋ</i>	<i>ʔaj</i>	<i>kaw</i>	<i>pə'lɽw</i>	<i>ka'sɛɛ</i>	<i>lɔɔm</i>	<i>ka'tiiək</i>	<i>ʔuujh</i>
Chatong	<i>ŋ'haaŋ</i>	<i>ʔaj</i>	<i>kaw</i>	<i>pa'law</i>	<i>ka'sɛɛ</i>	<i>lɔɔm</i>	<i>ka'tiiək</i>	<i>ʔoojh</i>
Tariw	<i>ʔi'haaŋ</i>	<i>ka'ʔaj</i>	<i>kuuu</i>	<i>bi'lɽw</i>	<i>ka'sɛɛ</i>	<i>lɔɔm</i>	<i>ka'tiiək</i>	<i>ʔuujh</i>
Dak Kang	<i>ʔi'haaŋ</i>	<i>ka'ʔaj</i>	<i>kuuu</i>	<i>pa'law</i>	<i>ka'sɛɛ</i>	<i>lɔɔm</i>	<i>ka'tiiək</i>	<i>ʔuuʔ</i>
Group B								
Suai (Juk)	<i>ka'tiin</i>	<i>jiʔ</i>	<i>ʔaw</i>	<i>bluu</i>	<i>khaj</i>	<i>klɛɛm</i>	<i>pa'tɛh</i>	<i>(ʔuuh)</i>
Laven (Jru')	<i>ka'tuuwəŋ</i>	<i>jiʔ</i>	<i>ʔaj</i>	<i>blɛw</i>	<i>khɛj</i>	<i>klɔɔm</i>	<i>pa'tɛh</i>	<i>ʔuuj</i>
Lavi (Savoeng)	<i>ka'cuuwəŋ</i>	<i>chiʔ</i>	<i>ʔɽɽ</i>	<i>pluu</i>	<i>(kɛɛt)</i>	<i>klɽɽm</i>	<i>pa'tiɛh</i>	<i>ʔuuj</i>
Tariang	<i>tiiəŋ</i>	<i>ʔiiʔ</i>	<i>ʔau</i>	<i>blaw</i>	<i>ka'hee</i>	<i>kuuəɔm</i>	<i>ca'neh</i>	<i>ʔuuj</i>
Kaseng	<i>ʔa'tiiəŋ</i>	<i>ʔiiʔ</i>	<i>ʔau</i>	<i>m'blaw</i>	<i>ʔa'haj</i>	<i>klɔɔm</i>	<i>ʔa'mɛh</i>	<i>ʔuuj</i>
Yaeh	<i>kriiəŋ</i>	<i>ʔiiʔ</i>	<i>ʔau</i>	<i>blaw</i>	<i>khe</i>	<i>kɔɔm</i>	<i>ʔa'neh</i>	<i>ʔuuj</i>
Alak	<i>ka'tiin</i>	<i>ʔa'jiʔ</i>	<i>ʔau</i>	<i>bluu</i>	<i>ka'haj</i>	<i>klɛɛm</i>	<i>ta'neh</i>	<i>ʔuuj</i>

Table 3: Voiced plosives and implosives in Tariw, Dak Kang, Chatong

	PEK	Tariw	Dak Kang	Chatong
'to sleep'	*b-	<i>bek</i>	<i>bec</i>	<i>p̄ic</i>
'to burn'	*b-	<i>buh</i>	<i>buh</i>	<i>p̄uoh</i>
'two'	*ɓ-	<i>baar</i>	<i>baar</i>	<i>p̄īaar</i>
'bamboo shoot'	*ɓ-	<i>ka'baŋ</i>	<i>ka'baŋ</i>	<i>p̄īaŋ</i>
'house'	*d-	<i>duŋ</i>	<i>duŋ</i>	<i>t̄aŋ</i>
'long (of time)'	*d-	<i>duuŋ</i>	<i>duuŋ</i>	<i>t̄ūoŋ</i>
'water'	*ɗ-	<i>daak</i>	<i>daak</i>	<i>t̄īaak</i>
'wife'	*ɗ-	<i>ka'diiəl</i>	<i>ka'diiəl</i>	<i>ka't̄iiəl</i>
'to peck'	*ɟ-	<i>ɟoh</i>	<i>ɟoh</i>	<i>c̄ūoh</i>
'to dip up'	*ɟ-	<i>ɟet</i>	<i>ɟet</i>	<i>c̄it</i>
'dry'	*g-	<i>goh</i>	<i>?a'goh</i>	<i>ŋk̄ūoh</i>
'mountain'	*g-	<i>gɔɔr</i>	<i>gɔɔr</i>	<i>k̄ūɔɔr</i>

To confirm the above assertion, let us examine carefully the following set of words in Tariw, Dak Kang, and Chatong listed in Table 4 below.

Table 4: Test words for identifying Katuic languages

	Proto Katuic	Tariw	Dak Kang	Chatong
'wife'	* <i>kəndeal</i>	<i>ka'diiəl</i>	<i>ka'diəl</i>	<i>ka't̄iiəl</i>
'unmarried woman'	* <i>kumoor</i>	<i>ka'moor</i>	<i>kɔɔn moor</i>	<i>kam'moor</i>
'to smile'	* <i>kəcaŋ</i>	<i>ka'caŋ</i>	<i>ka'caŋ</i>	<i>ka'caŋ</i>
'year'	* <i>kəmɔɔ</i>	<i>kam'mɔɔ</i>	<i>ka'mɔɔ</i>	<i>kam'mɔɔ</i>
'cobra'	* <i>duur</i>	<i>tuur</i>	<i>tuur</i>	<i>tuur</i>
'mushroom'	* <i>triəə</i>	<i>truuu</i>	<i>truuu</i>	<i>trii</i>
'to keep'	* <i>dɔɔk</i>	<i>(lut)</i>	<i>dɔɔk</i>	<i>t̄ūɔɔk</i>

We can clearly see that the sets of forms from these languages are cognates.

Regarding phonological innovation, medial /-h-/ in forms for 'centipede' in Tariw, Dak Kang, and Chatong also supports the idea that these are Katuic languages, since /-ʔ-/ from Proto-MK *-ʔ- is expected in many other Mon-Khmer branches but becomes /-h-/ in Katuic languages; cf. 'centipede': **kalheep* (PEK), *ka'hip* (Tariw, Dak Kang), and *ka'hēep* (Chatong). This is another criterion set up by Diffloth (1996:2).

Not only has the register distinction, that is clear voice versus breathy voice, developed in Chatong, but so also have the post-glottalised consonants *mʔ*, *-nʔ*, *-ŋʔ*, *-wʔ*, *-lʔ*, *-rʔ*, *-jʔ* in the same way as in Ta Oi', although fewer words have post-glottalisation in Chatong. We may note the following examples:

Chatong: *beek tranʔ* 'flint rock', *?aʔaʔ* 'more, again', *puuʔ* 'dull (of the pointed end)', *trɔʔ* *prɔʔ* 'excessively', *kanʔrɔʔ* 'to tiptoe', *kanʔɔɔmʔ* 'to close (wings)', *mʔphɔʔ* 'light (of weight)', *ka'colʔ* 'to leap', *luʔkluuʔ* 'getting dim (of candle light)', *?a'duʔ* 'float on the water', *tarʔrɔʔmʔ* 'to belch', *ha'rɔɔbrɔʔ* 'kind of paddy', *kam'briʔ* 'to blink', *ceewʔ* 'to tickle', *ka'vɔʔjʔ* 'to snatch', *puʔjʔ* *puʔjʔ* 'high spirited'.

Ta Oi': *m'paan?* 'maggot', *ha'pan?* 'to pretend', *la'laaj?* 'flash (of lightning)', *priin?* 'banana', *ηiiən?* 'to dive', *huan?* 'to steam', *lxj?* 'no, not', *man?* 'eyes', *maaj?* 'face', *suaj?* 'elongated', *la'ηəən?* 'soft', *ka'vaan?* 'to scratch with claws', *juaj?* 'gibbon', *kəəm?* 'to snatch', *ka'siam?* 'wither', *leem?* 'bird, animal', *lɛəm?* 'older sister-in-law', *raaj?* 'grasshopper', *ka'taam?* 'to slap, to clap', *puuul?* 'dull (of the pointed end)'.

On the basis of the examples given above, we may notice that the PEK finals **-p*, **-t*, and **-c* in some words have become post-glottalised *-m?*, *-n?*, *-j?* in Ta Oi'; cf. *kəəm?* 'to snatch', *ka'taam?* 'to slap', *priin?* 'banana', *man?* 'eyes', *raaj?* 'grasshopper', *suaj?* 'elongated'. I have not heard any pre-glottalised final nasals in the speech of my Chatong and Ta Oi' informants. In fact, a high-rising-falling pitch always accompanies the post-glottalisation in Ta Oi' and Chatong. I suspect that both pre- and post-glottalisation are perhaps part of the process of 'a non-tonal language becoming tonal'. Moreover, I think that Chatong and Ta Oi', which are spoken by hundreds of people, belong to the Yir-Ong-Ta Oih or Katang-Ta Oih group of the East Katuic sub-branch.

5 Bahnaric languages

Lexical innovations in numerals 'five', 'six', 'seven', 'eight', and 'nine' and phonological innovations in forms for 'tongue', 'bone', and 'fire' support the idea that Suai (Juk), Laven (Jru'), Lavi (Savoeng), Tariang, Kaseng, Yaeh, and Alak spoken in Xekong province are Bahnaric languages (see Table 5). No contradiction exists between my findings and the above criteria for identifying Bahnaric languages proposed by Diffloth. The only problems are that forms for 'five' in Suai, Laven, and Lavi and the form for 'fire' in Suai are Katuic borrowings; compare the following forms:

'five' **sâang* (PEK) > *saurŋ* (Brou), *səŋ* (Pacoh), *sâang* (Katu) (Thomas 1976:79), *ɤɤŋ* (Ta Oi', Katu, Nge', Chatong, Tariw), *pa'sɤɤŋ* (Dak Kang), *ɤɤŋ* (Suai, Lavi) *ɤəŋ* (Laven).

'fire' **ʔ(uu)s* (P-MK, Diffloth 1998), **ʔ-s* (PEK, Thomas 1976), *ʔuus* (Katu), *ʔuujh* (Ta Oi', Nge' Chatong, Tariw), *ʔoojh* (Chatong), *ʔuuʔ* (Dak Kang), *ʔuuh* (Suai or Juk).

Diffloth points out that final **-j* in 'fire' is a phonological innovation in Proto Bahnaric, that is Proto MK **-s* became **-j*, whereas the other Mon-Khmer branches still retain the Proto voiceless fricative. The Bahnaric languages of Xekong, except Suai (Juk), do retain the Proto Bahnaric **-j* (see Table 5). The Proto Bahnaric etymon **ləmpɨət* 'tongue' has been derived from the verb **liət* 'to lick', not the noun **ləntaak* 'tongue' in Proto MK (Diffloth). Forms for 'tongue' in Bahnaric languages are as follows:

**ləmpɨət* 'tongue' > Suai *la'piiət*, Laven *piiət*, Lavi *ha'pieet*, Tariang *ha'piat*, Kaseng, Yaeh *piat*, Alak *ha'peet*.

In most of the other Mon-Khmer branches, forms for 'bone' have a medial *-ʔ-*, but in the Bahnaric languages they have medial *-t-*; for example, Suai, Alak *ka'tiiŋ*, Laven *ka'tuuuəŋ*, Kaseng *ʔa'tiiəŋ*, and so forth.

Based on the data collected solely by myself, I believe that the Bahnaric languages spoken in Xekong province comprise three subgroups as illustrated below. First, Suai, Laven, and Lavi form one subgroup; the second subgroup includes Tariang Kaseng, and Yaeh spoken in

Dak Chueng district; Alak is a third subgroup by itself. Both lexical and phonological developments support this classification.

Table 5: The three subgroups of Bahnaric languages spoken in Xekong

	'house'	'long'	'shoulder'	'star'	'egg'	'neck'
Suai (Juk)	<i>srΛΛm</i>	<i>goh</i>	<i>paal</i>	<i>pa'toor</i>	<i>klɛɛ</i>	<i>lɔŋ</i>
Laven	<i>nΛΛm</i>	<i>goh</i>	<i>paal</i>	<i>pa'tuər</i>	<i>klɛɛ</i>	<i>lɔŋ</i>
Lavi	<i>sa'nΛΛm</i>	<i>goh</i>	<i>paa</i>	<i>pa'tɔər</i>	<i>klɛɛ</i>	<i>ha'lɔŋ</i>
Tariang	<i>suu</i>	<i>ruus</i>	<i>ca'laaŋ</i>	<i>φɛn</i>	<i>ka'leh</i>	<i>ndɔŋ</i>
Kaseng	<i>suu</i>	<i>?a'ruuɿ</i>	<i>laaŋ</i>	<i>ta'mɛŋ</i>	<i>leh</i>	<i>ndɔŋ</i>
Yaeh	<i>suu</i>	<i>ruuh</i>	<i>ku'laaŋ</i>	<i>?a'maŋ</i>	<i>keh</i>	<i>dɔŋ</i>
Alak	<i>jii</i>	<i>ja'rooŋ</i>	<i>ca'maa</i>	<i>bluuŋ</i>	<i>ka'tap</i>	<i>ŋɔk</i>

After searching through the available literature, such as Thomas and Headley (1970), Thomas (1979), Thomas and Srichampa (1995), and having discussions with Diffloth and Ferlus, I am certain that the first subgroup is West Bahnaric. My findings agree very well with the classification proposed by other Mon-Khmer specialists, no matter what methods or criteria they have used, that is Laven, Lavi, and Suai (Juk) must be West Bahnaric languages.³ For more details see L-Thongkum (1997a,b).

As for the second and third subgroups, the data from my fieldnotes contradict the classifications proposed by some Mon-Khmer specialists, that is that Alak belongs to Central Bahnaric, which is a very loose subgroup comprising Bahnar, Tampuan, and Alak (Thomas 1979); however, Diffloth places Alak, Tariang, and Kaseng within a new branch which is named 'Northwest Bahnaric'. If the Yaeh language of Xekong and the Jeh language of Vietnam are in fact the same language or dialects of the same language, then Yaeh should be a North Bahnaric language. Diffloth claims that 'bone' which has undergone very specific changes in different sub-branches of Bahnaric can be used as a test word for subgrouping within Bahnaric; for example, medial *-s-* is found only in the North Bahnaric sub-branch. To support his idea, he has cited forms for 'bone' from many languages of the North Bahnaric sub-branch; for example, Kacho' *kəɿ^hɛŋ*, Hre *kəseŋ*, Sedang *kəseŋ*, Jeh (*kə*) *siaŋ*, Halang *kəsiaŋ*, Cua *khiak*. Smith (1972) has reconstructed Proto North-Bahnaric **katseŋ* for 'bone'. Xekong Yaeh *kriiəŋ* 'bone' does not seem to fit any of the above North Bahnaric characteristics. On the basis of this criterion, I do not think that Yaeh belongs to the North-Bahnaric sub-branch. If Yaeh and Jeh turn out to be dialects of the same language, the only

³ Relying almost totally on vocabulary compiled from a variety of sources, Thomas and Srichampa (1995) tried to identify the affinities of the Mon-Khmer languages spoken in southern Laos. Their comparison of the distinctive vocabulary or distinctive forms of words yielded the following result: almost all of the languages belong to the West-Bahnaric sub-branch which comprises five subgroups, i.e. Northwestern, West Central, Southern, Central and Northeastern. Loven (Laven) is placed within the Northwestern subgroup, whereas Kraseng (Kaseng) and Trieng (Tariang) are placed within the Northeastern subgroup. As for Alak, it is a non-West Bahnaric language (Thomas & Srichampa 1995:306–307). On the basis of 'a unique Bahnaric phonological development in the word "bone"' (Diffloth 1996b:17), the languages within Bahnaric are sub-classified as West Bahnaric, Northwest Bahnaric, North Bahnaric, Central Bahnaric, and South Bahnaric sub-branches. In Diffloth's opinion, Alak, Tariang and Kaseng belong to the same sub-branch, i.e. Northwest Bahnaric. The classifications proposed by Thomas and Diffloth still need to be confirmed when more solid data are available.

possible explanation of the close relationship between Tariang and Yaeh is ‘language contact’, due to the fact that the Tariang and the Yaeh of Dak Chueng district have lived in the same area for a very long time. Moreover, it is more prestigious to be a Tariang since many members of this ethnic group have become the VIPS of Xekong province, and this can also encourage the borrowing of a lot of Tariang loanwords into Yaeh.

At this point, let us presume that Yaeh, Tariang, and Kaseng belong to the same sub-branch, namely Northwest Bahnaric. However, I do not think that Alak should be placed within this sub-branch as proposed by Diffloth; different types of lexical and phonological developments in Alak make me reject his classification of Alak. At the same time, I also do not support the idea that Alak is a member of the Central Bahnaric sub-branch as proposed by Thomas; this is due to the distribution of the Alak-speaking group which is in the northwest of the Bahnaric area and far away from the Tampuan and Bahnar areas, unless it is a displaced ethnic group as the result of wars or migrations in the past. Table 6 below lists some cognate items in Tariang, Kaseng, and Yaeh, and the non-cognate forms in Alak.

Table 6: Lexical evidence for separating Alak from Tariang, Kaseng, and Yaeh

	‘new’	‘head’	‘sky’	‘tooth’	‘tail’
Tariang	<i>naaw</i>	<i>tuus</i>	<i>pliij</i>	<i>puuɿŋ</i>	<i>suuəj</i>
Kaseng	<i>naaw</i>	<i>tuuɕ</i>	<i>pliij</i>	<i>puuɿŋ</i>	<i>sɔɔj</i>
Yaeh	<i>naaw</i>	<i>tuuh</i>	<i>pliij</i>	<i>puuɿŋ</i>	<i>sɔɔj</i>
Alak	<i>ta'maj</i>	<i>gauu</i>	<i>brah</i>	<i>ca'neŋ</i>	<i>teenj</i>

	‘hundred’	‘intestines’	‘skin’	‘to drink’	‘to steam’
Tariang	<i>riiəŋ</i>	<i>klaak</i>	<i>ruat</i>	<i>seet</i>	<i>coŋ</i>
Kaseng	<i>?a'riiəŋ</i>	<i>klaak</i>	<i>ruat</i>	<i>seet</i>	<i>coŋ</i>
Yaeh	<i>ka'riiəŋ</i>	<i>kaak</i>	<i>sruat</i>	<i>seet</i>	<i>coŋ</i>
Alak	<i>klam</i>	<i>pa'tuu</i>	<i>?a'kar</i>	<i>jɔɔk</i>	<i>sruu</i>

The phonological developments in the Alak cognates differ from those in the Tariang, Kaseng, and Yaeh cognates: diphthongs *uuə*, *ua*, *iiə*, *ia*, *au* (*aw*) in Tariang, Kaseng, and Yaeh correspond to Alak monophthongs *uu*, *oo*, *ii*, *e/ee*, *uu*; and Alak diphthongs *ai* (*aj*) and *ua* correspond to monophthongs *εε* and *oo* in Tariang, Kaseng, and Yaeh, respectively. See examples in Table 7.

Table 7: Vowel correspondences in Tariang, Kaseng, Yaeh and Alak

	‘road’	‘four’	‘bone’	‘banana’	‘rain’	‘tongue’
Tariang	<i>truuəŋ</i>	<i>puan</i>	<i>tiəŋ</i>	<i>priiət</i>	<i>miiə</i>	<i>piat</i>
Kaseng	<i>truuəŋ</i>	<i>puan</i>	<i>?a'tiəŋ</i>	<i>priiət</i>	<i>miiə</i>	<i>piat</i>
Yaeh	<i>truuəŋ</i>	<i>puan</i>	<i>kriiəŋ</i>	<i>priiət</i>	<i>miiə</i>	<i>piat</i>
Alak	<i>n'tuuŋ</i>	<i>poon</i>	<i>ka'tiiŋ</i>	<i>priit</i>	<i>mii</i>	<i>ha'peet</i>

	'root'	'nail'	'thigh'	'bathe'	'husked rice'	'three'
Tariang	<i>ʔa'rias</i>	<i>ka'rias</i>	<i>blaw</i>	<i>hoom</i>	<i>ba'hεε</i>	<i>pεε</i>
Kaseng	<i>rias</i>	<i>ʔa'riah</i>	<i>blaw</i>	<i>hoom</i>	<i>ʔa'hεε</i>	<i>pεε</i>
Yaeh	<i>ʔa'riaç</i>	<i>kriah</i>	<i>blaw</i>	<i>hoom</i>	<i>phεε</i>	<i>pεε</i>
Alak	<i>reh</i>	<i>ki'neh</i>	<i>bluu</i>	<i>huam</i>	<i>pa'haj</i>	<i>paj</i>

Even though the Tariang and Alak peoples are officially classified as 'Lao Thoeng' and live in nearby districts, several aspects of their cultures are different. Geographically, southern Laos, which comprises the four provinces of Champasak, Saravan, Xekong and Attapue, can be divided into three major areas: 'Thong Phiang (low land)', 'Phu Phiang (plateau)', and 'Phu Doi (mountain)'. Before liberation, the three most prestigious groups of 'Lao Thoeng', that is the Laven (Jru'), Alak, and Tariang, separately inhabited the three major areas mentioned above. At present, the majority of the Alak villages are still scattered in Lamam district which is a valley surrounded by high mountains; as a result, the Alak house style suits very well the warm climate of 'Thong Phiang' area. Dak Chueng district which is located in high mountains or 'Phu Doi' near the Lao-Vietnamese border is the center of the Tariang. The cold climate has caused the differences between the house styles of the Tariang and Alak and also their ways of living. The Laven who are quite rich in comparison with the other 'Lao Thoeng' ethnic groups live comfortably in the Boloven Plateau area which enjoys a temperate climate and they plant coffee trees. Although insignificant to language classification, these supporting reasons lead me to believe that languages spoken by the Alak and Tariang should not be placed within the same sub-branch. If Suai (Juk), Laven (Jru'), and Lavi (Savoeng) belong to the West-Bahnaric sub-branch, and Tariang, Kaseng and Yaeh belong to the Northwest sub-branch, then the place of Alak is still a problem which should be solved soon.

I heard an informal source state that the Kaseng and Tariang were in fact the same Mon-Khmer ethnic group. Most Kaseng people live in San Xai district of Attapue province which shares a common border with the Dak Chueng district of Xekong province in the south. I doubt the claim they are the same group of people is true. Among Mon-Khmer languages spoken in Xekong, Tariang and Kaseng are the most closely related; however, they are not the same language in my opinion. A comparison of basic vocabulary items (based on Matisoff's 200-wordlist) in these languages as in Table 8 below seems to support this view.

Table 8: Basic vocabulary differences between Tariang and Kaseng

	'mouth'	'finger'	'palm'	'fish'	'otter'	'flower'
Tariang	<i>nuuŋ</i>	<i>caaç</i>	<i>ta'paan</i>	<i>kaa</i>	<i>ba'hee</i>	<i>ʔa'piər</i>
Kaseng	<i>buəl</i>	<i>ʔool</i>	<i>klɔəm</i>	<i>ʔa'dum</i>	<i>biər</i>	<i>ʔa'raan</i>
	'brook'	'branch'	'shade'	'shadow'	'crossbow'	'to sleep'
Tariang	<i>nɔɔŋ</i>	<i>baak</i>	<i>ŋuut</i>	<i>mook</i>	<i>sroʔ</i>	<i>leek</i>
Kaseng	<i>huuŋ</i>	<i>n'daaj</i>	<i>kuuwp</i>	<i>ʔa'raʔ</i>	<i>ʔa'nan</i>	<i>ŋiər</i>
	'close (eyes)'	'to weep'	'to run'	'red'	'to eat'	'to kick'
Tariang	<i>ka'soot</i>	<i>rɯ'vɯh</i>	<i>n'duuəç</i>	<i>bruuh</i>	<i>caa</i>	<i>ŋ'jraat</i>
Kaseng	<i>kluw</i>	<i>kliiəw</i>	<i>ʔa'har</i>	<i>broon</i>	<i>cɔŋ</i>	<i>daç</i>

As for cognate words, we detect the two different patterns of phonological changes:

1. Proto Northwest-Bahnaric *CV(C) ' CV(C) is retained in Tariang but has become CV(C) in Kaseng, and vice versa. See examples in Table 9.
2. Proto Northwest-Bahnaric *ai and *ua are retained in Tariang but have become ee and oo in Kaseng, and vice versa. See examples in Table 9.

Table 9: Different patterns of phonological changes in Tariang and Kaseng

	'egg'	'nail'	'knee'	'foot'	'person'	'bitter'	'to tie'
Tariang	<i>ka'leh</i>	<i>ka'rias</i>	<i>ta'kool</i>	<i>ca'kooŋ</i>	<i>ma'ŋaaj</i>	<i>ŋ'jaŋ</i>	<i>m'blaŋ</i>
Kaseng	<i>leh</i>	<i>riah</i>	<i>kool</i>	<i>kooŋ</i>	<i>ŋaaj</i>	<i>jaŋ</i>	<i>blaŋ</i>
	'bone'	'rice'	'ten'	'to stand'	'long'	'sweat'	'itchy'
Tariang	<i>tiiəŋ</i>	<i>saaj</i>	<i>jit</i>	<i>juuəŋ</i>	<i>ruus</i>	<i>hil</i>	<i>kiiət</i>
Kaseng	<i>?a'tiiəŋ</i>	<i>n'saaj</i>	<i>ka'jit</i>	<i>?a'juuəŋ</i>	<i>?a'ruusɔ</i>	<i>?a'hul</i>	<i>?a'kiiət</i>
	'horn'	'tail'	'child'	'you'	'fruit'	'moon'	'otter'
Tariang	<i>?a'kuaj</i>	<i>suaj</i>	<i>?a'kuan</i>	<i>maj</i>	<i>plee</i>	<i>ka'hee</i>	<i>ba'hee</i>
Kaseng	<i>?a'kooj</i>	<i>soj</i>	<i>?a'koon</i>	<i>mee</i>	<i>plaj</i>	<i>?a'haj</i>	<i>biar</i>

6 Conclusion

On the basis of both lexical and phonological developments, we can identify and group within the two Mon-Khmer branches the thirteen Mon-Khmer languages spoken in Xekong province of southern Laos as follows: the Katuic branch includes the languages of Ta Oi', Katu, Nge'/Kriang, Chatong, Tariw, and Dak Kang; and the Bahnaric branch includes the languages of Suai/Juk, Laven/Jru', Lavi/Savoeng, Tariang, Kaseng, Yaeh, and Alak. The six Katuic languages belong to the East Katuic sub-branch; and the seven Bahnaric languages belong to three sub-branches: (1) West Bahnaric with Juk, Jru', and Savoeng; (2) Northeast Bahnaric with Tariang, Kaseng, and Yaeh; and (3) Northwest Bahnaric with Alak (?). The classification of Alak is still problematic: whether Alak is Central Bahnaric (Thomas), Northwest Bahnaric (Diffloth), or something else is still to be decided.

Once I have finished collecting data for the Xekong project, I might be able to give a better solution. While I have not yet worked extensively on Tariang, Kaseng, Yaeh, and Alak, the present stage of my studies leads me to modify previous classifications proposed by other Mon-Khmer specialists into the following subgroupings:

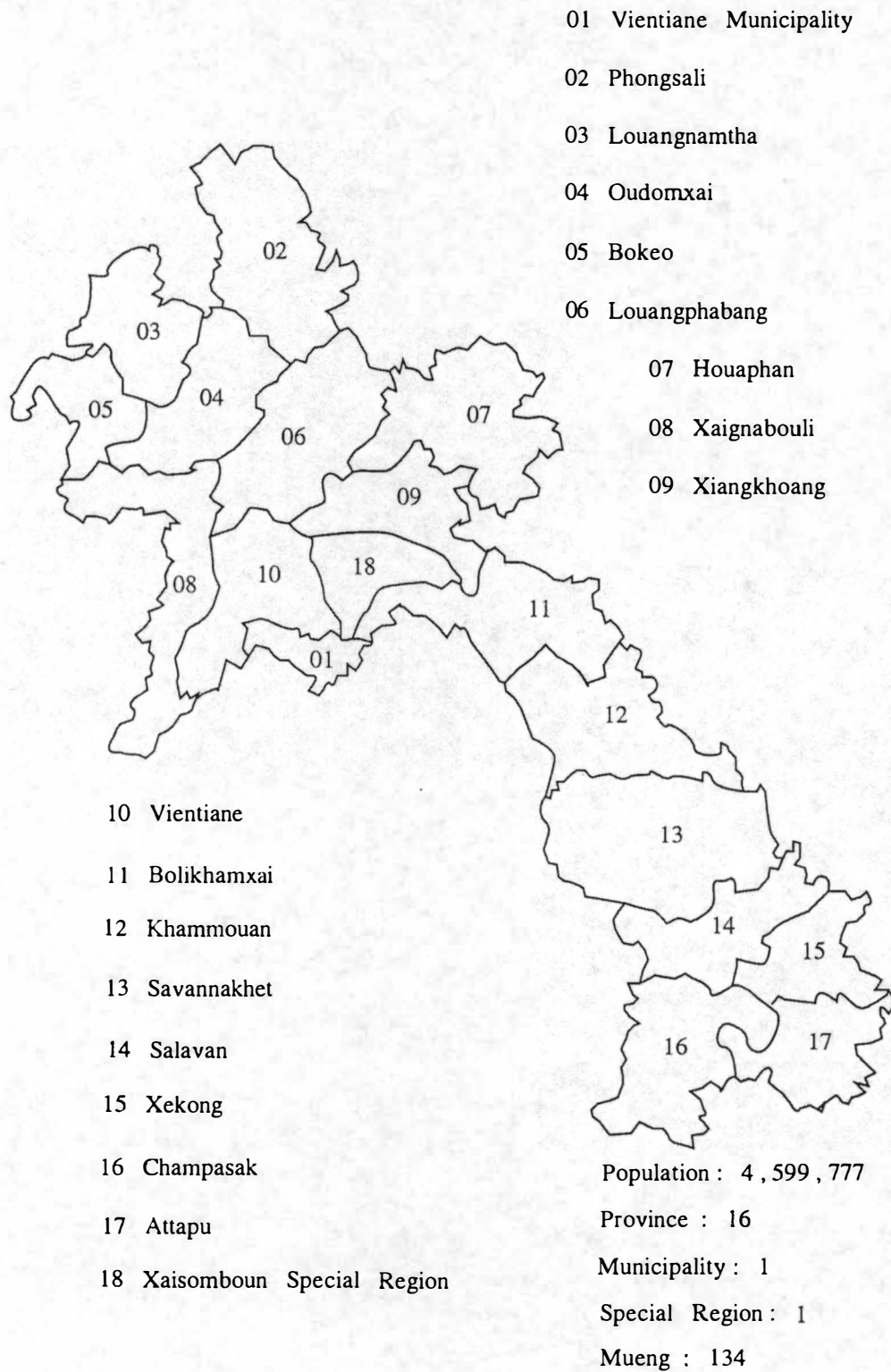
- | | | |
|----------|-----------------|---|
| Katuic | <i>West:</i> | Kui, Kuay, Bru, So, etc. |
| | <i>Central:</i> | Nge', Ta Oi', Chatong, Tariw, Dak Kang, etc. |
| | <i>East:</i> | Katu, Pacoh, Phuong, etc. |
| Bahnaric | <i>North:</i> | Northeast: Tariang, Kaseng, Yaeh, etc.
Northwest: Alak, etc. |
| | <i>West:</i> | Laven, Lavi, Juk, Brao, Su', Nyah Hueny, etc. |
| | <i>East:</i> | Takau, Cua, Sedang, Hre, etc. |
| | <i>Central:</i> | Tampuan, Bahnar, etc. |
| | <i>South:</i> | Mnong, Stieng, Sre, Chrau, etc. |

My commitment to the Thailand Research Fund (TRF) is to produce a book in Thai and a few papers in English as products of this project. When well-organised sets of data are available and more solid evidence is produced, then the members of our small Mon-Khmer community will be quite well-equipped to search for the truth. At present we can only hypothesise. I would like to stress that my present opinion is based on my preliminary research and is subject to future change.

References

- Diffloth, Gérard, 1992, East Mon-Khmer languages. *International Encyclopedia of Linguistics*, vol.1:390–393.
- 1996a, Katuic languages. MS, 15pp.
- 1996b, Bahnaiic languages. MS, 21pp.
- 1998, An introduction to Mon-Khmer languages. MS.
- L-Thongkum, Theraphan, 1997a, The place of Lawi, Harak and Tariang within Bahnaiic. *Mon-Khmer Studies* 27:109–117.
- 1997b, Suai, a newly discovered West Bahnaiic language of southern Laos. *Proceedings of Chulalongkorn University 80th Anniversary Research Conference*, 347–354. Bangkok: Chulalongkorn University.
- Peiros, Ilia, 1996, *Proto Katuic comparative dictionary*. Melbourne, Australia. Draft edition.
- Smith, Kenneth D., 1972, *A reconstruction of Proto-North-Bahnaiic*. Language Data, Asia-Pacific Series, No.2. Sant Ana, California: SIL.
- Thomas, Dorothy M., 1976, *A phonological reconstruction of Proto-East-Katuic*. Grand Forks: SIL.
- Thomas, David, 1979, The place of Alak, Tampuan, and West Bahnaiic. *Mon-Khmer Studies* 8:171–186.
- Thomas, David and Robert K. Headley, 1970, More on Mon-Khmer subgroupings. *Lingua* 25/4:398–418.
- Thomas, David and Sophana Srichampa, 1995, *Some languages of Siam by Phraya Prachakij-karacak*. Translated and annotated. Bangkok: Institute of Language and Culture for Rural Development.

Appendix



Map 1: Lao PDR

III

Discourse analysis

7 *The tiger mother's child and the cow mother's child: a preliminary look at a Bru epic*

JOHN MILLER AND CAROLYN MILLER

1 Background and origin of the text

When we lived in a Bru village in the mountains of central Vietnam in the 1960s, we listened to many traditional Bru stories. Villagers liked us to tape-record these stories on the small reel-to-reel tape recorder we had with us so that they could come to hear them again and again. This was a favourite activity on moonlit nights when people from other villages came to visit. We heard stories about the 'orphan boy' and others who were 'underdogs' with whom the Bru identified and whose triumph they celebrated against 'Anha', the rich and powerful overlord. Some stories were explanatory in telling how things got to be the way they are.

In the 1970s and 1980s we worked in Sabah, Malaysia and came to appreciate the wealth of stories of the Kadazan people. We did some study of narrative texts both oral (Miller 1987) and written by Kadazan authors (Miller & Miller 1990). We found Bru and Kadazan to be very different linguistically: while Kadazan was phonologically simple but morphologically complex, Bru was the reverse with complex phonology and simple word structure.

Another difference between these two groups of people became apparent when in the 1990s we found ourselves in Northeast Thailand doing research on languages closely related to Bru. While the Kadazan had developed an appreciation for their language and its oral literature and were involved in preserving this in writing for their children, the Bru-related groups in Northeast Thailand had for the most part not retained their oral literature. We suggested to Bru friends in Vietnam and Laos that they begin to write down these stories for the benefit of their children and grandchildren and for the enjoyment of those groups whose cultural background is similar to theirs. The text analysed in this paper is the result of our suggestion.

The author is a Bru man, Mpoaq Saniq, who has lived in many areas of Vietnam and Laos. He is fluent in both Vietnamese and Lao, but he has not lost an appreciation for his own language. And he has been able to preserve the flavour of the oral text while bringing the

Robert S. Bauer, ed. *Collected papers on southeast Asian and Pacific languages*, 111-139. Canberra: Pacific Linguistics, 2002.

© John Miller and Carolyn Miller

111

refinement and precision of the written page. He wrote the text in a modified Roman-script orthography developed in the 1960s and still used by the Bru in Vietnam. For purposes of this paper, the text has been reproduced in a regularised phonetic representation of the original.

As the text was written without paragraph notation and minimal punctuation, the numbered paragraphs included in the appended text have been introduced by us. It was not an easy task to determine paragraphs, since as Longacre suggests, internal structure of paragraph relations are more 'loose and diffuse' than those within sentences (Longacre 1985:235). Some are quite obvious having a 'unity of theme or action in a specified time setting' (Howard 1978:273), but others are more arbitrary.

We have generally broken paragraphs to indicate change of speaker as well as change of focus. Pickering (1980) speaks of a 'span' in which 'one participant has the initiative with respect to other participants and events, or is in focus', and says that 'a switch of initiative of focus to another participant would signal a boundary at some level' (1980:24). We have generally taken this to be a grammatical paragraph.

Numbers in square brackets used throughout the paper indicate paragraph numbers. Since the text is too long to include in its entirety, only the part from the peak episode to the end of the narrative has been appended to this article, but the examples have been taken from the entire text.

2 Scope of this study

It is impossible within one article to attempt in-depth investigation of all the various features of this extensive work. In the sections below, we will look briefly at the structure of the narrative, the plot development, and some of the highlights and outstanding features of the text. More detailed analysis will have to await future study.

The narrative provides a rich source of information for the study of Bru language and culture. We gain insights into Bru cosmology, metaphysics, marriage, family interaction, ideas of politeness and appropriateness, and each of these areas would provide profitable study. Additionally, the interplay between discourse features and grammatical features which are used by the author to highlight participants and introduce events in such a way as to maintain a coherent whole are well worth more serious study. Stylistic features of word compounding and the selection of vocabulary in order to produce phonological symmetry and elegance are features found not only within this text, but have also been described for related languages (Watson 1977, 1980). In addition to these areas of study, the devices used to indicate emotion and attitudes and the terms of address used by the various speakers in different situations to indicate social distance and levels of politeness all constitute invaluable material for future research.

3 Structure of the narrative

The story as written by Mpoaq Saniq has all the features of a good narrative text — title [0], aperture [1–19], stage setting [20–38], several pre-peak episodes [39–224], a peak episode [225–292] in which suspense, setback, hope and climax are evidenced, a post-peak [293–296], a closure [297–300], and a finis [301] (Longacre 1996:36). However, although all of these parts are tied together into a coherent whole, each episode is in itself a story in microcosm with most of the elements — setting, problem, suspense, solution, and closure — of a good narrative text.

Even the aperture and stage-setting sections display this characteristic. For example, while we believe (and the title supports this) that the main focus of the story is on the adventures of the two sons of the half-tiger and half-cow mothers, the story of the unfaithful wife and the husband's clever dispatching of the tiger by trickery in the opening section forms a story within a story. Likewise, the account of the ill-will between the two half-sisters leading to the murder of the younger by the elder, and the subsequent murder of the elder by her own son in the stage-setting section, is told in detail, including conversational interchange.

The four pre-peak episodes each centre around some central figure or figures met by the two heroes (Tasuo, the elder, and Tangu, the younger) in their journey together through life. Central to the whole is their commitment to each other even though in the end they are separated, each to rule in a different sphere, one in the overworld and one in the underworld. Also emphasised throughout the whole by various rhetorical devices is their respectful treatment of others and their concern to help those being mistreated by others.

4 Plot development

Episode One [39–113] involves interaction with Grandfather and Grandmother Soc Catuc. This couple is discovered to have cannibalistic intentions when the test is made of showing them an ostensibly bleeding hand. The two 'brothers' are helped in their escape by a termite and a rat, and the rest of the episode details their flight through various obstacles until at the end Grandfather and Grandmother Soc Catuc are dispatched by sympathetic villagers.

Episode Two [114–164] brings them to Grandfather and Grandmother Seng Toc Seng Lua, where the test of the bleeding hand reveals these two to be good persons. They take the boys in and raise them as their grandchildren, and when the boys want to leave home the Grandfather prepares them with special supernatural abilities and arranges for special weapons to be designed for them.

Episode Three [165–199] brings them in contact with two sisters who have been left as human sacrifices for the fearsome Phep Nhac. The girls, who have accepted their fate to die for the sake of the village, try to get the two visitors to leave. But the guests stay to eventually defeat the Phep Nhac by their extraordinary powers. Though the girls cling to them and beg them to stay, the two heroes fly away, leaving behind the girls, who have managed to hang on to only a fragment of their loincloths.

Episode Four [200–224] finds the girls getting a chilly reception from their parents upon their return to the village. The villagers, initially angry and fearful that the deliverance of the girls will expose them to retribution from the Phep Nhac, have to be persuaded by the girls that the Phep Nhac have been defeated. The boys, meanwhile, fearful of incurring the wrath of the villagers who have for years been under the control of the Phep Nhac, are hiding in the home of a widow. The villagers now want to honour the two young men, and the parents of the girls have promised to give the two girls to them in marriage, but they are not able to be found. Many hopefuls try to pass the test of matching the pieces of loincloth, but only when the two half-brothers are eventually found at the widow's house is the match made. Marriage and material rewards ensue.

Episode Five [225–295], which seems to be the peak episode, involves only the younger brother at the start. He develops the desire to wander, and drawn by a very sweet smell, he manages to find his way to a lower level of earth. (According to Bru cosmology there are multiple levels of earth and the same number of levels of heaven.) Here he finds Miss Creq

Phuom weaving. Enamoured of Miss Creq Phuom he asks her to marry him. She is already engaged to the terrifying Tiau Calang Cloc, but she agrees to the union if he will leave the world above and come to live with her. He goes back and talks to his older half-brother who disapproves but leaves the choice to the younger. A fairly strong Bru value as reflected in the exchange between the two brothers is that of individual personal choice. The brothers exchange flowers (reminiscent of an incident in Episode One where the Cow Mother did this with her son) so that each will know if the other is in danger or need. Younger brother leaves his family, goes to the lower world, and marries Miss Creq Phuom.

The Tiau Calang Cloc, when they come to visit, are not happy to find this man with the woman they claim as their fiancée. They become increasingly abusive (especially after Tangu says he has married the girl) and a fight ensues. Tangu defeats them, but they have a source of healing and eventually wear him out and kill him.

Older brother Tasuo, who knows by the wilting of the flower that Tangu is in trouble, comes to help. He also encounters the boastful and disrespectful Tiau Calang Cloc and learns that they have killed Tangu. Finding Tangu's bones scattered far and wide, Tasuo enlists the help of the animals to bring them back together where he uses special powers to restore Tangu to life.

In the final section of the peak episode, the two protagonists strategise to defeat the dreaded Tiau Calang Cloc. Things very nearly turn out disastrously in a repeat of the earlier battle, but Tasuo has observed the source of renewed strength and healing for the Tiau Calang Cloc and removes the teeth from the Pangolin and the tongue from the Flying Fox. The direction of the battle reverses; the Tiau Calang Cloc are destroyed with the exception of the leader who flees the scene.

In the post-peak section, the king of the Tiau Calang Cloc, when confronted with the misdeeds of his men, apologises for all the harm they have done. In the final resolution, the two heroes separate, Tangu to stay with Miss Creq Phuom and rule the underworld, Tasuo to return and rule the world above.

5 Major markers used in the story

In our study of Kadazan narratives, we found that some of the most commonly used functors and affixes had both syntactic and discourse functions and that these two uses were not unrelated. Bru does not have the same interplay of tense and aspect allowed by the highly inflected Kadazan language, but much the same phenomenon is seen in the use of the time words, grammatical particles, and deictics.

5.1 Time

5.1.1 Chronological progression

Time throughout the story is frequently indicated on the discourse level by time margins. Examples of this are:

- (1) *ho:j ce:ʔ e: di: dɔw luʔ cɣ:.*
 when near want midnight truly already
 '...when it was already midnight...' [4]

- (2) *maho:j ba:r...*
moment two
'In a moment or two...' [15]
- (3) *khʌn ho:j lɛ? ma: mɔj hu:m pjar nʌj,...*
if when which but 2SG see flower this
'Whenever you see this flower,...' [24]
- (4) *te: ho:j raj ki:.*
from before that
'...from before.' [225]
- (5) *du:n te: ki:...*
long.time from that
'A long time later...' [2]

5.1.2 Main action time

Time is also indicated by tense markers within the main action clauses of the text. The following chart notes some of these and examples are given below. These do not indicate salience (Longacre 1996), but merely summarise the major tense markers which move the story along. Very frequently the marker for completed action and the marker for perfective action occur together as in example (8) below.

Completed action	NP <i>khɔ:j?</i> VP	136 occurrences
Perfective action	NP VP <i>cɔ:</i>	662 occurrences
Immediate action	NP VP <i>tʌ:p</i>	40 occurrences
Unrealised action	NP NEG VP <i>juah</i> NP NEG <i>juah</i> VP	41 occurrences
Imminent action	<i>ŋke?</i> NP VP	5 occurrences
Future action	<i>cɔ: pʌ: mat</i> NP VP	1 occurrences
Preliminary action	NP VP <i>wʌ:j</i>	27 occurrences

- (6) *tɔaw kala:ŋ klɔ:k nʌj khɔ:j? kaci:t an.*
PN this PST kill 3SG
'The Tiau Calang Cloc killed him.' [276]
- (7) *an saŋa:t cɔ:,*
3SG dry.up already
'...it has already withered.' [24]
- (8) *mpi:ʔ aŋja khɔ:j? ca: cɔ: mpi:ʔ he? pʌ: saraj ki:.*
mother 2PL PST eat already mother 1PL(excl) direction field that
'Your mother has already eaten my mother out in the field.' [29]
- (9) *khʌn an tʌ:ʔ la: he? cil an tʌ:p.*
if 3SG arrive be 1PL(excl) fight 3SG right.away
'If he comes, we'll fight him immediately.' [68]

- (10) *khlan kuaj ale? ma: tɿ: juah dan, ki: ci:m wə:j.*
if person whoever but **not yet** know that test beforehand
'If anyone doesn't yet know, let him first try.' [243]
- (11) *kaja:k an tɿ: juah cu: te: ce? ca:ŋ.*
husband 3SG **not yet** return from sell buy
'...her husband had not yet returned from buying and selling.' [2]
- (12) *ŋke? acujh kɿp aja:? sɔk katu:k ki: sata:m ʔj?*
soon grandfather and grandmother PN that catch.up.with EMPH
'Soon Grandfather and Grandmother Soc Catuc will catch up with us.' [74]
- (13) *cɿ: pɿ: mat, la: ku? tɿ: bu:n tah mɿj ma:,*
return direction future be 1SG not leave 2SG NEG-EMPH
a?e:m ʔj.
younger.sib. VOC
'In the future, I will not abandon you, younger brother.' [233]

5.1.3 Comparative time

Very frequently throughout the text time is indicated by the interplay of dependent and independent clauses. Consecutive time is indicated by /bɔ:/ in the dependent clause, if the event is viewed as a whole, and /tu:/ in the dependent clause, if the event is viewed as a process leading to a point in time. The following independent clause is usually introduced by either /cɿ:/ or /ŋki?/.

- (14) *bɔ: an se:ŋ te: saru:ŋ ki:, cɿ: an pa:j nɛ:?,*
when 3SG go.down from loft that **then** 3SG say like.this
'When he came down from the loft, he said...' [10]
- (15) *tu: an arɔ: sa?u:j tro:, ŋki? kula: nɿj, an co:n tɿŋ saru:ŋ*
when 3SG call many times **thus** tiger this 3SG go.up in loft
doŋ.
house
'When he had called many times, then the tiger went up into the loft of the house.' [4]
- (16) *lu? tu: mɿ:m kunɿj nɿj ca: cɔk pasɿj lu? cɿ:,*
very **when** finish rat this eat one's fill very already
ŋki? kunɿj nɿj pa:j nɛ:? e:n cɔ? alɿj ba:r na?
thus rat this say like.this in.addition for 3PL two person
se:m a:j nɿj.
older.bro. this
'When the rat had eaten till he was very full, he said to the two brothers...' [56]

Example (17) shows concurrent time. This example has two dependent clauses: one introduced by /bɔ: nɿ:ŋ/ and the second by /ntɔ:m/. The independent clause is introduced by /la:/. (See §5.2.3)

- (17) *bɔː nɔːŋ won waj, ntɔːm nɑːw proːm prəj baːw kiː,*
when still troubled **while** someone gather common.people that
laː alɔj koŋ ɕaː? nik ŋɛː tɛː.
be 3PL also look.for always idea also
 ‘...while they were still troubled, and while someone was gathering the people,
 they were also looking for an idea...’ [212]

In examples (18) through (20), simultaneous action is indicated by the repeated use of the word /ntɔːn/ before each verb which is a part of the action being carried out at the same time.

- (18) *kɔːn mpiː? ntrɔː? ntɔːm paːj maː ntɔːm ɲiam.*
 child mother cow **simultaneously** say but **simultaneously** weep
 ‘The cow child said while weeping...’ [31]
- (19) *ntɔːm awɔːj maː ntɔːm blɔh, “ntrɔw aɲia eː*
simultaneously hold.out but **simultaneously** ask what 2PL want
ta? bɔː ɕɔw?”
 do huh grandchild
 ‘As they gave it to them, they asked, “What do you want to do, grandchildren?”’
 [57]
- (20) *ŋki? alɔj baːr na? seːm aːj nɔj huːm baːr na? ɲiaŋ*
 thus 3PL two person brothers this see two person young.woman
nɔj ntɔːm ɕaː maː ntɔːm ŋkah.
 this **simultaneously** eat but **simultaneously** afraid
 ‘The two brothers saw the girls eating even while they were filled with fear.’ [181]

Other illustrations of comparative time are shown in examples (21) through (24). In example (21) both clauses contain a time word. In the other examples, only one clause contains a time word, but the second clause is marked by a deictic or copula.

- (21) *hoj raj laː kɔːn kumɔːj? a? mpaː?,*
before be child corpse lack father
maː sanɔa haj lu? kaːt kɔːn tuːj eːn.
but now 1PL(incl) truly become child lacking.both.parents in.addition
 ‘Before, we were orphans without a father, but now we have become double
 orphans.’ [38]
- (22) *ŋke? alɔj tɔː?, kiː he? tamprɔːh alɔj ɕaːj joːn aɲia.*
soon 3PL arrive **that** 1PL(excl) fight 3PL help for 2PL
 ‘When they come, we will fight them on your behalf.’ [68]
- (23) *tɛː? aɲia pɔː?. maː ɲuaŋ aɲia pɔː?, laː kɔ? buːn*
 can 2PL go but **before** 2PL go **be** must have
 ‘You can go. But before you go, you must have...’ [142]
- (24) *mɔm kajaːk an ci caː ɕɔː, ŋki? an thruan be?*
finish husband 3SG eat already **thus** 3SG prepare sleep.
 ‘After her husband had eaten, he prepared to sleep.’ [7]

Some of the words which commonly occur as time markers have related spatial uses also. For example, /tu:/ not only indicates comparative time leading up to the main action as in examples (15) and (16), it may also indicate ‘up to’ in space as in example (25) below. Similarly, the words for ‘before’ and ‘after’, for ‘arrive’, and for ‘return’ have spatial meanings as illustrated in example (26)

- (25) *alɣj mbΛ:j? tu: ntruəj tɔ:ŋ*
 3PL just **up to** chicken male
 ‘...they had just gotten as far as the male chickens...’ [76]
- (26) *to:p kuaj thΛw kɣp kuaj karne:n lɣ: lɣ: la: puaj ntun*
 group person elderly and person children small be follow **after**
kuaj pɔŋ kj:
 person young that
 ‘The elderly people and children followed after the young people.’ [203]

5.1.4 Discourse linkage time

By far the most commonly used time marker in Bru narrative is the perfective word /cɣ:/. It occurs 662 times in this text. It functions on various levels as a boundary marker. A similar feature was noted by Gregerson and Martens (1986) in Uma for the perfective /mi/. It is used in verbal, adverbial or adjectival constructions with the meaning of ‘already’, as illustrated in examples (7) and (8) above and in the two below.

- (27) *Λ:t mumiaq cɣ:*
 be.at pregnant **already**
 ‘...was already pregnant.’ [3]
- (28) *yua? rana? mɣj ŋki? cɣ:*
 because work 2SG thus **already**
 ‘...because you have already done this...’ [19]

Also illustrated in example (14) above as well as in examples (29) through (31) below is its very common use following a clause introduced by another time word such as /tu:/ or /bɔ:/ ‘when’. In this usage, it marks an action which follows that previously mentioned one. In this position it carries the meaning of ‘then’.

- (29) *tu: kaja:k an cu: te: ce? cΛ:ŋ, cɣ: kaja:k an*
 when husband 3SG return from sell buy, **then** husband 3SG
nΛj hu:m...
 this see
 ‘when the husband returned from buying and selling, then he saw...’ [3]
- (30) *bɔ: an se:ŋ te: saru:ŋ kj:, cɣ: an pa:j nɛ:?,...*
 when 3SG go.down from loft that **then** 3SG say like.this
 ‘When he came down from the loft, then he said,...’ [10]
- (31) *ŋki? tu: an daŋ kula :nΛj khɔ:j? ɲɛ:ʔ rɛ:ŋ cɣ:, cɣ:*
 thus when 3SG know tiger this PST all strength already **then**

an hɔ:r su:m...

3SG drag basket

'So when he knew that the tiger had exhausted his strength already, then he dragged the basket...' [17]

Very commonly the word /cɣ:-/ introduces a sentence within a paragraph. In this usage it indicates continuity of action or sequence of action. In these cases it might be translated 'so', 'so then' or 'next'.

(32) **cɣ:** *ɬ:t tɒŋ doŋ e:n, la: an abu:n kɣp ntrɔ:ʔ hɬ:*
then be.at in house again be 3SG desire with cow also
 'Then back at home again, she wanted to sleep with a cow also.' [2]

(33) **cɣ:** *an ka: lakuəj pɬ:ʔ pɣh tɔ:ŋ yo:n kaja:k an mɔ:t.*
then 3SG who wife go open door allow husband 3SG enter
 'Then the wife opened the door for her husband to come in.' [4]

As an exclamation, it occurs on its own, where it means 'That's enough already!'

(34) *kula: nɬj, an pa:j nne:ʔ, "cɣ: kalɬ: ɣ:j,..."*
 tiger this 3SG say thus **already** friend VOC
 'The tiger said, "That's enough, friend..."' [14]

On the discourse level /cɣ:-/ plays an important part in moving the story along. Occurring at the start of a paragraph it indicates closure of one incident and the setting of the new action as following that which precedes in the previous paragraph.

(35) **cɣ:** *mah pɔ:ng taru:p pa:nɬ:,...*
then like light morning tomorrow
 'Then the next morning,...' [9]

(36) **cɣ:** *kɔ:n mpi:ʔ ntrɔ:ʔ nɬj, an taʔɬ:j lɔ:h ne:ʔ,...*
then child mother cow this 3SG reply again like.this
 'Then the child of the cow mother answered,...' [27]

5.2 Linkage

A number of devices are used throughout the text to indicate juncture. Some of these are the subordinating markers for purpose, reason, condition, concession, substitution, addition as described by Thompson and Longacre (1985). Others are more coordinating or boundary marking.

5.2.1 Inclusion

Several words are used for joining or indicating inclusion. These are /deh/ 'including', /ɛ:n/ 'in.addition', /hɬ:/ or /tɛ:/ 'also', /koŋ/ (preceding a verb phrase) 'also', or /nuŋ/ 'with'. Sometimes more than one of these is used together. This does not seem to indicate emphasis, but reflects the tendency to give weight to an utterance by doubling or compounding expressions as discussed in §6. The word /nuŋ/ is also commonly used with a verb of request to indicate that the one requesting places himself in the position of a supplicant as illustrated in examples (44) and (45).

- (37) *he?* *ŋkah mpi:?* *aŋja ca: deh* *to: he?* *en.*
1PL(excl) afraid mother 2PL eat **including** body 1PL(excl) **in.addition**
'I'm afraid your mother is going to eat me too.' [29]
- (38) **deh** *aŋja la: ka:t rawe:h alxj tɛ:*
including 2PL be become soup 3PL **also**
'...you too are going to become soup for them also...' [52]
- (39) *ŋki? alxj ce? ntɔŋ lu?, deh pra? je:ŋ, deh*
thus 3PL sell expensive very **including** silver gold **including**
tarjak en hɔ:
water.buffalo **in.addition also**
'So they asked a very high price including silver, gold and water buffalo also.' [112]
- (40) **deh** *alxj ba:r na? se:m aj ki: la: satɔ:h ɲe:? tɔŋ dɔ:?*
including 3PL two person brothers that be fall all in water
kro:ŋ ki:
river that
'The two brothers also fell into the river.' [113]
- (41) *to:p alxj nɔ:ŋ thre:k pa:j deh ku? aj en hɔ:*
group 3PL still insult say **including** 1SG older.bro. **in.addition also**
'This group insulted even me your older brother also.' [273]
- (42) *cɣ: kɔ:n ntrɔ: kɣp kɔ:n kula: nɔj, alxj koŋ bu:n kaja:k tɛ:*
then child cow and child tiger this 3PL **also** have husband **also**
'Then the cow child and the tiger child also had husbands.' [21]
- (43) *he? se:? lɣ:j? nuŋ hɔ:*
1PL(excl) request play **with also**
'I want to come up and play with you too.' [28]
- (44) *se:? aja:?* *ama:jh cɔaj he? cɔw nuŋ jo:n*
request grandmother awaken help 1PL(excl) grandchild **REQ** so.that
alxj tamɔ:
3PL wake.up
'We ask you, grandmother, to please help us wake them.' [216]
- (45) *he? se:m aj nɔj se:? kanɔ:m ba:r lam aŋi: an*
1PL(excl) brothers this request ask.favour two CL uncle 3SG
cɔaj amɔ:ŋ he? nuŋ.
help live 1PL(excl) **REQ**
'...we respectfully beg you two uncles to help save our lives.' [84]

The word most commonly used to indicate inclusion and show linkage on various levels is the word /kɣp/. This word is used 301 times throughout the text. On the phrase level, /kɣp/ is used most frequently to join nouns, but it is also used to join verbs or descriptives in a coordinate way. In this sense it is best translated 'and'. It is used frequently in the expressions /lakuəy kɣp kaja:k/ 'husband and wife' and /acɔajh kɣp aya:ʔ/ 'grandfather and grandmother', but its use in these expressions is optional, since /lakuəy kaja:k/ and /acɔajh aya:ʔ/ may be used without the joining word /kɣp/.

- (46) *bu:n ba:r na? lakuəj kɔp kaja:k la: kadi:t lu?*
 have two person wife **and** husband be poor very
 'There were two persons, a husband and wife, who were very poor.' [1]
- (47) *ta? na? la: saŋ tɔ: rɔj kunja? su:m ki: jo:n ruŋ*
 do how be only not able break.open basket that give spoil
kɔp raha?
and tear
 'Whatever he did he could not break open the basket to make it ruined and torn.' [14]

When used in a verbal phrase the meaning of inclusion or accompaniment is best conveyed by a translation of 'with' or 'toward'.

- (48) *ŋki? kaja:k nɔj pa:j kɔp lakuəj an nne:ʔ,*
 thus husband this say **with** wife 3SG thus
 'So the husband said to his wife...' [1]
- (49) *mɔj kɔ? bu:j nik kɔp ba:r lam kɔ:n mɔj nɔj,*
 2SG must happy always **with** two CL child 2SG this
 '...you must be satisfied with these two children...' [19]

On the sentence level, /kɔp/ is used to connect clauses which relate to each other in a coordinate way.

- (50) *an ka: ɔj kɔ:t kɔ:n kula:*
 3SG who older.sister become child tiger
kɔp an ka: aʔe:m kɔ:t kɔ:n ntrɔ:ʔ.
and 3SG who younger.sib. become child cow
 'The elder sister was a tiger's child, and the younger was a cow's child.' [19]

At the discourse level, /kɔp/ is used to bring in additional information which is related to that previously mentioned, but where time sequence is not in focus. Often it is used to introduce speech where this is seen as accompanying action, as in example (50). It is also used to introduce serial questions as shown in example (51).

- (51) *cɔ: an hɔ:r su:m kula: nɔj du:ŋ cɔ: dɔ:ʔ.*
 then 3SG drag basket tiger this bring return water
kɔp an pa:j nne:ʔ, "taŋaj anɔj ɔ:ʔj?"
and 3SG say thus day this EMPH
la: taŋaj aŋia cɔ: kruaŋ iən khe:, kruaŋ aklɔ:k...ɔ
 be day 2PL return country peace country drowning
 'Then he dragged the basket to the river. And he said, "This very day is the day you return to the country of peace, the country of drowning..."' [17]
- (52) *ŋki? ba:r na? se:m a:j nɔj blɔh nne:ʔ, "sakɔ:r nɔj nɔ:w*
 thus two person brothers this ask thus drum this someone
da:ʔ ta? ntrɔw? kɔp mah prɔa acɔ: ta?nɔj la: nɔ:w
 put do what **and** PLUR bushhook knife this be someone

e: ta? akan ntrɔw, nu: nɪaŋ? kɔp to:p phe:p na:k
 want do thing what huh young.woman and group PN
 naɟ, alɔj kluw mahle?? kɔp rɔ:p rɪaŋ alɔj la: na?
 this 3PL many how.many and form likeness 3PL be how
 tɔ:p mahle?? tɔ:r to: alɔj?"
 exactly how.much big body 3PL
 'Then the two brothers asked, "What is this drum for? And what is this bushhook
 for, huh girls? And this Phep Nhap group, how many are there? And what do they
 look like? How big are they?"' [182]

5.2.2 *Adversative*

Another important boundary marker used within the text is /ma:/. The word /ma:/ carries an adversative meaning. It is used on the sentence level in the usual way to contrast two ideas or conditions. In this use, it is best translated 'or' or 'but'.

- (53) mɔj saŋ te:, ma: tɔ: bu:n?
 2SG hear also but not
 'Do you hear it too or not?' [7]
- (54) aŋia khɔj? pupa:ʔ te: doŋ, ma: he? co:n taŋ saru:ŋ
 2PL PST walk from house but 1PL(excl) go.up in loft
 doŋ aŋia ŋki?
 house 2PL thus
 'You were gone from home, but I came up into the loft of your house.' [10]
- (55) ŋki? alɔj ba:r na? lu? ka:t se:m a:j te: mɔj aja:ʔ
 thus 3PL two person truly become brothers from one grandmother
 ma: mɪar acɔajh.
 but different grandfather
 'So those two became brothers from one grandmother but different grandfathers.'
 [20]

At the discourse level, /ma:/ is often used to introduce an action or situation which is contrary to expectation. In the opening section of the text, for example, a number of paragraphs begin with /ma:/. Where the husband is expected to go away as he said he would, he instead hides and watches his house (56). When the husband asks his wife why she's so slow opening the door (57), she does not tell him the truth, but instead lies about it (58). She lies again about hearing the noise of the tiger hiding in the loft (59). When the husband confronts the tiger, instead of pouring abuse on him, he speaks to him in a friendly manner (60).

- (56) ma: kaja:k an tɔ: bu:n lu? pa:ʔ nɔ:ŋ ce? ca:ŋ, ma:
 but husband 3SG not truly go still sell buy but
 an ɔ:t tɔ:ʔ mpe:r ki: tɔ:p
 3SG be.at hide around that right.away
 'But the husband didn't really go to buy and sell, but he hid close by...' [4]

- (57) **ma:** *kaja:k an na:j lu? khə:j? daŋ cɣ: mah ŋe: alɣj.*
but husband 3SG this truly PST know already PLUR way 3PL
cɣ: an ta? na:n blɣh lakuəj an nne:?, "na? mɣj tɔ: bu:n
 then 3SG pretend ask wife 3SG thus why 2SG not
pɣh caj tɔ:ŋ jo:n ku??"
 open fast door for 1SG
 'But her husband already knew what they intended. So he pretended to ask his wife, "Why didn't you open the door quickly for me?"' [5]
- (58) **ma:** *lakuəj an kake? nne:?, "ku? be? laŋe:t lu?*
but wife 3SG lie thus 1SG sleep soundly very
ku? tɔ: bu:n saŋ mɣj aɔ: ki: an kɔ: ɣ:j?."
 1SG not hear 2SG call that 3SG long.time EMPH
 'But his wife lied and said, "I was sleeping very soundly. I didn't hear you calling. That's why I was so slow."' [6]
- (59) **ma:** *lakuəj an ta?ɔ:j la:h ne:?, "ku? tɔ: bu:n saŋ nɔ:r*
but wife 3SG reply again like.this 1SG not hear noise
akan ntrɣw."
 thing what
 'But his wife replied, "I don't hear any noise at all."' [7]
- (60) **ma:** *an tɔ: bu:n pa:j ntrɣw ak saw?*
but 3SG not say what thing bad
 'But he didn't say anything bad at all...' [11]

In a construction in which a conditional clause is followed by a counter-conditional clause, it conveys the idea of 'if, on the other hand'.

- (61) *khɔn aŋja raja:jh jɛ:ʔ raŋuɣh rahɔ: ma: su:m na:j nɔ:ŋ patə:t*
 if 2PL move.around all strength but basket this still similar
tja?, ki: he? e: cɔ? ali:k he? tɔŋ na:j ɣ:j?. khɔn ma:
 old that 1PL(excl) want put pig 1PL(excl) in this EMPH if **but**
an ruŋ, ki: haj ta:n la:h mɣəj lam tamaj ta? khɔm
 3SG ruined that 1PL(incl) weave again one CL new make sturdy
hɔ:n te: na:j e:n.
 more from this in.addition
 'If you move around with all your strength but the basket remains as usual, I'll put my pig in it. But if it is ruined, then we'll weave a new one stronger than that one.'
 [13]

5.2.3 Sentence level linkage

Other important markers of juncture within the text are /ka:/ and /la:/. These are lower level linkage markers, but they aid cohesion throughout the text. The marker /ka:/ is used as a relative clause marker. It is frequently used in participant reference, not only upon initial introduction, but often throughout the narrative. It is sometimes used to contrast one versus

another, as in example (64). And it is frequently used in WH-questions. According to Comrie (1981), a relative clause construction must minimally contain a head and a relative clause. In the Bru text, the head is always present, but the restrictive clause is sometimes very minimal as in examples (64) and (65).

- (62) *he? jɔ:k e: pɬ:ʔ i:t ali:k ka: he? khɔ:j? cɬ:ŋ cɣ:,*
 1PL(excl) want want go take pig **which** 1PL(excl) PST buy already
 'I want to go get a pig which I have already bought.' [12]
- (63) "ɣ:ʔ! nɬ:w **ka:** ɬ:t tɬŋ pɬ:ŋ ki: ɣ:j, se:ŋ tɔ:h!
 Oh someone **who** be.at in above that VOC go.down implore
 'Hey! Whoever you are up there, please come down!'
- (64) *cɣ: an ka: lakuəj pɬ:ʔ pɣh tɔ:ŋ jo:n kaja:k an mɔ:t.*
 then 3SG **who** wife go open door allow husband 3SG enter
 'So the wife went and opened the door for her husband to come in.' [4]
- (65) *an ka: ɬj... kɣp an ka: aʔe:m...*
 3SG **who** older.sister and 3SG who younger.sib
 'The older sister...and the younger sister...' [19]
- (66) *cɣ: kaja:k an blɣh lakuəj nne:ʔ, "ntrɣw ka: nnɬ:r tɬŋ*
 then husband 3SG ask wife thus what **which** noise in
sarɔ:ŋ doŋ haj?"
 loft house 1PL(incl)
 'Then the husband asked the wife, "What's that making a noise in the loft of our house?"' [7]

The form /la:/ functions as a copula in equational sentences. In other cases it marks clause juncture of other types such as topic-comment constructions. A very common use of /la:/ is with the negative, as in example (72), in which the meaning is 'or'. The form /la:/ is found 253 times in the text.

- (67) *cɣ: kula: nɬj, an taʔɬ:j ne:ʔ, "ɣ:ʔ! he? la:, kalɬ:."*
 then tiger this 3SG reply like.this all.right 1PL(excl) **be** friend
 'Then the tiger replied, "Okay! It's I, friend." [10]
- (68) *ma: ali:k he? cɬ:ŋ nɬj la: tɬ:r luw? tanajh tɔ:l.*
 but pig 1PL(excl) buy this **be** big very huge male
 'But the pig I bought is a very large male pig.' [12]
- (69) *ŋki? tɬŋ mɔ:j sadɬw ki:, luw? pa:j an sɬŋ nɬ:r saʔw:j*
 thus in one night that truly say 3SG hear noise many
tro:, la: an tɬ: bu:n taʔ bo:ʔ nɬ:ŋ.
 times **be** 3SG not speak any.more
 'So that whole night, even though he heard the noise many times, he didn't say anything more about it.' [8]
- (70) *haj bu:n doŋ su: la: da:ʔ jo:n jɣw baj klah klɣ: ɣ:j?*
 1PL(incl) have house shelter **be** put for friends pay.a.visit EMPH
 'The reason we have a house is so that friends can visit us.' [11]

- (71) *luʔ pa:j heʔ tʌ: bu:n ʌ:t ndoŋ la: tʌ: bu:n ntrɔw ka:*
 truly say 1PL(excl) not be.at house be not what which
taʔ ma:, kalʌ: ɣ:j.
 do NEG-EMPH friend VOC
 ‘Even though I’m not at home, there is no problem, friend.’ [11]
- (72) *lɔ:jh tʌ: la tʌ: bu:n lɔ:jh la: aɣj ca: dajh, kalʌ: ʌ:j.*
 wrong or not wrong be 3PL eat EXCL friend VOC
 ‘Whether you are wrong or not, they will eat you, friend.’ [52]

5.3 Deixis

Bru has a well-developed series of deictics, and these play an important role in a Bru narrative. In their prototypical role, they show three levels of distance – ‘here’, ‘there’ and ‘farther over there’. They also distinguish ‘this’ and ‘that’. In a more abstract way, they may be used to refer to a certain situation or even to a speech act itself. These more abstract deictics may be either anaphoric with ‘backward-looking identification’ (Grimes 1978:316), or cataphoric, with ‘forward-looking reference’ (Grimes 1978:317). The table below gives the more common deictics and the number of occurrences within the text under study.

Table 1: Bru deictic distinctions and number of occurrences in Bru text

	Near		Far		Farther	
Deictic	<i>nʌj</i>	915	<i>ki:</i>	346	<i>tih</i>	6
Demonstrative	<i>anʌj</i>	25	<i>aki:</i>	10	<i>atih</i>	4
Locative	<i>nnʌj</i>	6	<i>ŋki:</i>	3		
Abstract	<i>nnɛ:ʔ/nɛ:ʔ</i>	224	<i>ŋkiʔ</i>	452		

As can be seen from the above table, the term /tih/ ‘over there’ occurs less frequently and seems to occur only in the simple deictic use of ‘over there’ or the demonstrative ‘that one over there’.

- (73) **tih**, *an ʌ:t kə:h aŋja tih nɣ:.*
there 3SG be.at side direction **there** huh
 ‘...over there, he’s over in that direction.’ [265]
- (74) **tih** *ce:ʔ nʌ:m aɣaj ka: tʌ:r tʌŋ saruŋ atih nɣ:ʔ*
there near trunk tree which big in young.forest **over.there** huh
 ‘Over there near the trunk of the big tree in that forest over there.’ [275]
- (75) *mɔʔj sadʌw kɔh nʌj, mɔʔj taŋaj saruŋ atih.*
 one night mountain **this** one day young.forest **over.there**
 ‘...one night at this mountain, one day in that forest.’ [39]
- (76) *pʌ:ʔ de:r da:w wə:j nɣ:, pʌ: ntoʔ aɣajh atih.*
 go forge sword beforehand huh direction place grandfather **over.there**
 ‘Go forge swords first, okay, over there at that grandfather’s place.’ [142]

Forms /nʌj/ and /ki/ following a noun carry the usual meaning of ‘this’ or ‘that’. Additionally, however, /nʌj/ is used throughout the text to identify participants who are ‘on stage’ at a given time. Looking at the start of the text, for example, the husband and wife are brought into the story without a deictic. But very quickly, in the first paragraph, the husband is singled out to be /kaja:k nʌj/ ‘this husband’. The wife, who is not on stage, is referred to as /lakʊəj an/ ‘his wife’.

Other props such as the loft, the pepper, etc. are backgrounded with the use of /ki:/. Then in paragraph [10], the tiger also comes on stage and is identified as /kula: nʌj/ ‘this tiger’. Since the main characters in the story are the two brothers, they are identified by the word /nʌj/ throughout much of the text, sharing the stage from time to time with others.

Used in a number of constructions, /ki:/ very commonly introduces a main clause which follows a conditional clause, as in example (77), or a clause which expresses the meaning of ‘whatever’, ‘whenever’, ‘however much’, etc. as in examples (78) through (80).

- (77) *cɣ:*, *an ka: aj pa:j e:n nne:?*, “*khʌn bu:n an ka:*
then 3SG who older.bro. say in.addition thus if have 3SG who
hu:j cɣm haj nʌj, ki: haj pʌ:ʔ tʌ:p te: nʌj,
singe bird 1PL(incl) this **that** 1PL(incl) go right.away from this
nʌŋ e:m.”
right younger.sib.
‘So the older brother said, “If there were someone to singe this bird of ours, we’d go right away, wouldn’t we younger brother?”’ [172]
- (78) *ntrɔw ʌŋja jɔ:k e: pa:j cɔʔ acʊajh ʌŋja, ki: ʌŋja pa:j nɣ:*
what 2PL want want say for grandfather 2PL **that** 2PL say huh
‘Whatever you want to say to your grandfather, you say, okay?’ [123]
- (79) *acʊajh ɣ:j! mahleʔ acʊajh jɔ:k e: pa:j, ki: ta:m*
grandfather VOC how.much grandfather want want say **that** it’s.up.to
acʊajh.
grandfather
‘Grandfather, however much you want to say, it’s up to you.’ [156]
- (80) *ma: ho:j leʔ ʌŋja saŋ tʌ: te:ʔ pʌ:ʔ nʌ:ŋ, ki: ʌŋja*
but when which 2PL feel not can go any.more **that** 2PL
pih cʊ: lə:h pʌ: doŋ haj, nɣ: cʌw.
go.back return again direction house 1PL(incl) huh grandchild
‘...but whenever you feel you can’t go any more, then you come back to our house, won’t you grandchildren?’ [163]

In the phrase /ki: nʌʔ/, it conveys the meaning ‘then only’.

- (81) *ʌŋi: tɔ:ŋ ɣ:j. heʔ se:ʔ ʌŋi: pa:ŋ rana: jo:n*
uncle male VOC 1PL(excl) request uncle make.way road allow
heʔ ba:r naʔ se:m a:j luh wə:j hʌ: ki: nʌʔ
1PL(excl) two person brothers run beforehand also **that how**

aŋja karcit la:h sia.
 2SG fight.together again as.before
 ‘Uncle Roosters, we ask you to make way for us to run first. Then you can go back to fighting.’ [65]

The forms listed in the table above as being demonstrative seem to carry a feeling of more specificity. They occur only following nouns. One might hypothesise that these are a contraction of the relative pronoun /ka:/ plus the deictic /ki:/ since this combination is found to occur with a similar meaning as shown in example (82) below. Further study is needed to confirm this.

- (82) *ŋki? kula: nΛj, an ta:n su:m jo:n kaja:k mpi:ʔ ka:*
 thus tiger this 3SG weave basket for husband mother **who**
ki:, la: tΛ:r khΛm luʔ.
that be big sturdy very
 ‘So the tiger wove a basket for the husband of that woman, large and very sturdy.’ [13]
- (83) *juaʔ alxj canxɯm pa:j la: kuaj nΛ:ŋ be? tΛŋ nto? aki:.*
 because 3PL think say be person still sleep in place **there**
 ‘...because they thought someone was still sleeping in that place.’ [63]
- (84) *cɣ: an sΛŋ hΛ:j phuəm luʔ tΛŋ saruŋ aki:.*
 then 3SG smell smell fragrant very in young.forest **there**
 ‘Then he smelled a very sweet smell in that forest.’ [226]
- (85) *taŋaj anΛj x:j? la: taŋaj aŋja cɯ: kruaŋ iən khe:.*
 day **this** EMPH be day 2PL return country peace
 ‘This very day is the day you will return to the peaceful country (i.e. die).’ [17]
- (86) *ŋki? se:ʔ aŋja pΛ:ʔ e:n te: nto? anΛj.*
 thus request 2PL go in.addition from place **this**
 ‘Please leave this place.’ [168]

The forms which begin with a syllabic nasal convey the meaning of place – /nΛj/ ‘here’ and /ŋki:/ ‘there’. With the use of this form, no noun is required, since place is implied in the form itself. A case might be made for considering the form a contraction of /tΛŋ/ ‘in/on’ plus the deictic, particularly since the syllabic nasal has this function in a few other contexts such as /ndon/ ‘at home’ or /nnx:j/ ‘at that place’.

- (87) *khə:j? du:n cɣ: aŋja ralɯ: nΛj.*
 PST long.time already 2PL rest **this**
 ‘You have rested here for a long time already.’ [168]
- (88) *cɣ: pΛ:ʔ la:t plaŋ baŋ kəh. sadΛw nleʔ be? ŋki:.*
 then go wander all.over mountain night where sleep **there**
 ‘So they wandered all over the mountains, sleeping wherever they ended up at night.’ [39]

The two abstract deictics are those with the widest range of uses. The word /nne:ʔ/ or /ne:ʔ/ (as it sometimes appears) is cataphoric and is used to introduce new information to which attention is being drawn. Sometimes the new information is in the form of a list. At

other times it relates to thoughts or decisions arrived at. Sometimes it is in the form of an announcement on the part of the narrator.

- (89) *kɔ:n kula: bu:n muəj lam kɔ:n samjan, kɔp an amuh ramuh*
 child tiger have one CL child male and 3SG name name
kɔ:n an nne:?, a:j tasuə.
 child 3SG **thus** older.bro. PN
 ‘The tiger (mother) had a male child, and she named him this: Tasuo.’ [20]
- (90) *lu? ku? a:j hu:m nne:?, to:p tɔaw kala:ŋ klɔ:k nɔj,*
 truly 1SG older.bro. see **thus** group PN this
lu? alɔj uan nik kɔp mɔj.
 truly 3PL unhappy always with 2SG
 ‘But I have seen clearly this: this Tiau Calang Cloc group is really angry with you.’ [273]
- (91) *cɔ: alɔj cuə? ŋe: ne:?, alɔj i:t aha:m cɔm. cɔ: atja tɔŋ*
 then 3PL look.for idea **like.this** 3PL take blood bird then rub on
ati: an ka: a?e:m.
 hand 3SG *who* younger.sib.
 ‘Then they got this idea: they took the blood of a bird and rubbed it on the hands of the younger brother.’ [41]

But the most common use of the word /*nne:?*/ in the text is to indicate a quote margin. Of the 224 occurrences in the text, all but twenty mark direct quotations. And of the twenty which do not, several are used to mark indirect quotations.

- (92) *ŋki? kaja:k nɔj pa:j kɔp lakuəj an nne:?,*
 thus husband this say with wife 3SG **thus**
 ‘Then the husband said to his wife...’ [1]
- (93) *cɔ: an ta? na:n blɔh lakuəj an nne:?,*
 then 3SG pretend ask wife 3SG **thus**
 ‘Then he pretended to ask his wife...’ [5]
- (94) *ma: lakuəj an kake? nne:?, “ku? be? laŋe:t lu?”*
 but wife 3SG lie **thus** 1SG sleep soundly very
 ‘But the wife lied, saying, “I was sleeping very soundly.”’ [6]
- (95) *cɔ: kumɔ:t pa:j nne:?, “kɔn ŋki?, aŋja co:j krq:? ntrɔw.*
 then termite say **thus** if thus 2PL don’t fear what
he? tɛ: cuəj aŋja.”
 1PL(excl) can help 2PL
 ‘Then the termite said, “If it’s like that, don’t be afraid. We can help you.”’ [48]

The use of /*nne:?*/ as a quote margin is not required, and a few instances are found where it does not occur. But the majority of direct quotes are preceded by this word, and the places where it is omitted, as in the latter part of example (96), seem to be in a context of rejoinder where speakers have already been identified.

- (96) *cɣ: tɔ:p pʰe:p ja:k nɔj lu? tɔ:?. cɣ: an saba:w nɛ:?,*
 then group PN this truly arrive then 3SG call.out **like.this**
“bat baj asa:j sakjəw kham cɣ:, ba: njaŋ?”
 vegetables ginger onion sufficient already huh young.woman
ŋki? aɔj ba:r na? se:m a:j nɔj taʔɔ:j lə:h nɛ:?, “ɣ:?”
 thus 3PL two person brothers this reply again like.this yes
kham cɣ:.” “pɾja acu: kham cɣ:, ba: njaŋ?”
 sufficient already bushhook knife sufficient already huh young.woman
“kham cɣ:.”
 sufficient already
 ‘Then the Phep Nhap group arrived. They called out, “Are all the condiments ready, girls?” The two brothers replied, “Yes. Everything is ready.” “Are the bushhooks and knives ready?” “Yes, they are ready.” [190]

A similar range of uses is found for the anaphoric deictic word /ŋki?/. As object of a verb or comment in a topic-comment construction, the word is used fairly commonly. It often occurs as the object of verbs of sight or speech.

- (97) *tu: kaja:k an hu:m rɔ:p kɔ:n an ka:t ŋki? cɣ:,...*
 when husband 3SG see form child 3SG become **thus** already
 ‘When the husband saw that his children had an appearance like that,...’ [19]
 (98) *ma: he? co:n tɔŋ saru:ŋ doŋ aŋja ŋki?*
 but 1PL(excl) go.up in loft house 2PL **thus**
 ‘...but I went up into the loft of your house like that.’ [10]

The word /ŋki?/ is also frequently found after a time clause to introduce the main action of the sentence.

- (99) *tu: an arɔ: sa?u:ɔj tro:, ŋki? kula: nɔj, an co:n*
 when 3SG call many times **thus** tiger this 3SG go.up
tɔŋ saru:ŋ doŋ.
 in loft house
 ‘When he had called many times, the tiger went up into the loft of the house.’ [4]
 (100) *mɔ:m kaja:k an ci ca: cɣ:, ŋki? an thruan be?*
 finish husband 3SG eat already **thus** 3SG prepare sleep.
 When the husband had eaten, he prepared to sleep.’ [7]
 (101) *khɔn ho:j lɛ? ma: mɔj hu:m pjar nɔj*
 if when which but 2SG see flower this
an saŋa:t cɣ:, ŋki? lu? ku? mpi:ʔ kuci:t
 3SG dry.up already **thus** truly 1SG mother die
 ‘Whenever you see this flower wilt, then surely your mother has died...’ [24]

With the preposition /jua?/, /ŋki?/ is used to denote cause.

- (102) *he? sɔŋ ŋkah lu? jua? ŋki? ma: he? luh puaj*
 1PL(excl) feel afraid very **because thus** but 1PL(excl) run follow

rana: nʌj.

road this

'We are really afraid. That's why we are running along this road.' [67]

On a discourse level, /ŋkiʔ/ connects ideas or actions which are related more in a cause/result or thought progression sequence than a time sequence. In this usage /ŋkiʔ/ occurs sentence or paragraph initially.

- (103) *ŋkiʔ, du:n kasʌj te: taŋaj ki: lakuəj an beʔ tapəh.*
thus long.time month from day that wife 3SG sleep fireplace
 'So, a long time after that, the wife went into labour.' [19]

- (104) *ŋkiʔ, mpi:ʔ kɔ:n ntrɔ:ʔ kɔp kɔ:n kula: nʌj, luʔ an cuum sɿəm*
thus mother child cow and child tiger this truly 3SG raise
mʉəj nɔ:ʔ ansʌŋ tɔ:w kɔ:n an tʌ:r kʌ:t kumu:r.
 one.person 3SG only until child 3SG big become young.woman
 'So the mother of the cow child and the tiger child raised them by herself until the children grew up to become young women.' [20]

6 Compounding

A common feature of Mon-Khmer languages is the use of what has been variously termed expressives, ideophones, binomials, or double words (Watson 1980:76). In some cases these expressions have parts which are either rhyming or phonologically similar in some way. They may be very common expressions or those which are created by the speaker.

Some of these expressions take a word which has meaning by itself and add a second word which has no meaning of its own simply to add fullness or weight to the expression. In many cases these show some form of reduplication or phonological similarity. For example, in the following list, the first word of the pair may be used on its own with the meaning indicated. The second word only occurs as an extension of the first.

/lɣ:ɿʔ lɣ:/	'play, visit'
/raŋwʰ rahʌ:/	'strength'
/krɣ:ŋ krɣw/	'possessions'
/hu:n hi:t/	'kiss' (lit. 'sniff')

Other expressions take two words with similar meanings and put them together to strengthen or give emphasis to the whole. For example, both /kɣt/ and /ŋkʉjʔ/ mean 'small'. Used together, they mean 'very small' or 'tiny'. The same is true of the other combinations given below. The use of two words which have the same general meaning either intensifies or makes the expression more complete.

/kɣt ŋkʉjʔ/	'tiny'
/iəŋ khe:/	'easy, peaceful'
/rʌ:p riəŋ/	'appearance'
/kɔŋ rɛ:ŋ/	'strength, labour'
/pheʔ thu:/	'deceive'
/tah tɔ:ɿʔ/	'discard'
/bɣk cu:l/	'injury, wound'
/ŋi:n khɔ:/	'confess'

In some cases two items of a class are used to indicate the entire category or class (Anderson 1985:40).

/doj ravə:h/	rice + soup = 'food'
/prja acu:/	bushhook + knife = 'utensils'
/alɿ:k tarijak/	pig + buffalo = 'domestic animals'
/doj su:/	house + shelter = 'family'

Some of these expressions are built up to convey a sense of the superlative or of picturesque speech.

- (105) *ŋki? to:p phe:p ja:k kuci:t tu:n u:n tuuk luuk thaŋ, tɿ: rɿ:j*
 thus group PN die **in.large.numbers** completely not able
nə:p nɿ:ŋ.
 count any.more
 'So the Phep Nhad died in great numbers, so many they couldn't be counted.' [192]
- (106) *ŋki? jɛ:ʔ tu:h kuaj pu:n palə:ŋ kutru? nɿj se:ʔ saʔɿ:n lə:h*
 thus all all person **under sky** **covering this** request thank again
 'So all the people under the covering of the sky want to thank...' [223]
- (107) *cɿ: heʔ tampe:ʔ mah mu:n cin tapu:n raləʔ*
 then 1PL(excl) divide **PLUR inheritance completely**
mah sanok ntun e:n jo:n an ki:
PLUR wealth in.addition give 3SG that
 'Then we will share all our inheritance and property with him...' [205]
- (108) *cɿ: alɿj tɿ:r aco:n ku: sadɿw tanaj kasaj kumə:*
 then 3PL big lift.up **each night day month year**
kɿ:t tata:m.
 become young.man
 Then they grew day by day, month by month, year by year until they became young men. [132]

Within the text are a few examples of poetic or formulaic speech. This feature often is found in Bru folk literature. This is illustrated, for example, in the ditties sung by the craftsmen forging swords for the two brothers.

- (109) *ŋki? se:m aj nɿj jo:n acujah pul lul nɿj de:r. cɿ: acujah*
 thus brothers this give grandfather PN this forge then grandfather
nɿj luʔ de:r, kɿp an pa:j nɛ:ʔ, "pul lul arjah klɿn
 this truly forge and 3SG say like.this **curved like anus**
kupest, pe:t le:t arjah klɿn kapah."
tick flat.&.thin like anus small.insect
 'So the brothers allowed Grandfather Pul Lul to forge swords. This man forged, and as he worked he sang this ditty, "Curved like a tick's anus, flat and thin like a bug's anus.'" [145]

Some phrases seem to be put together with an ear for stylistic elegance. These are found most frequently in the transition portions between episodes where the narrator is setting the scene. Several different types of alliteration or sound consonance are involved. One noted for

Rengao by Gregerson (pers. comm.) is of the type where the middle two of four words are made to rhyme in some way. Watson (1977:312-313) calls this 'rhetorical underlining'. This is seen in the poetic phrase used for forest by the narrator in example (110) and in the phrase in (111) which describes the wanderings of the brothers. Example (112) shows a series of three phonologically similar pairs used in one expression to describe the sustenance found by the brothers. Example (113) shows verb reduplication in the common phrase for 'upriver and downriver' for dramatic effect.

- (110) *ŋki? bu:n taŋaj ki: an saŋ jɔ:k e: pɬ:ʔ pupan la:h tɬŋ*
 thus have day that 3SG feel want want go hunt again in
aru:ʃh aruŋ saruŋ ma:ŋ sia,
 forest deep forest only as.before
 'Then the day came when he wanted to go hunting again in the forest.' [225]
- (111) *co:n kɔh sajɔh puh, ɬt tɬŋ leŋ seŋ tɬŋ le:h.*
 go.up mountain prance side be.at in pasture go.down in slope
 'They climbed mountains up and down the sides; they went through pastures and down hillsides.' [40]
- (112) *tam:ŋ ɲɔ: aluŋ alə:j pasɬj jua? palɬj palɔ:t*
 live because.of tree small.bamboo full because fruit plentiful
tɬŋ saruŋ sarja?,
 in forests
 'They were able to live because of the vegetation, able to eat because of the fruit in the forests.' [39]
- (113) *cɔ: an to:j? i:t kɔp palah atuŋ palah atɔ:k*
 then 3SG get take and turn.toward upriver turn.toward downriver
ɲe:ŋ sali:ʔ.
 look look
 'So she picked it up and turning first one way and then the other looked inside.'
 [114]

7 Indicators of emotion and attitude

Much of the flavour of the narrative is found in the words used either by the narrator or the characters themselves to indicate emotion or attitude. This feature is referred to as 'mood' by Grimes, who notes that this category includes 'communication options, logical status, and attitude of the speaker toward the content of the message' (1978:235). Levels of politeness, certainty, familiarity, surprise and affirmation are indicated by use of various markers in the text under study. Many of them are found only in direct address of which this text has a great deal.

7.1 Confirmation request

One of the most obvious characteristics of the speech found within the text is the range of particles used by speakers to request confirmation or agreement. In some cases the agreement is expected to be in the form of action. In these cases the markers are more of command than question.

It seems to us that of all the modal particles used in the text, these markers which request verification most clearly show the spectrum of formality to informality and politeness to rudeness demonstrated within the story.

Table 2: Bru modal particles for requesting verification arranged along more formal/polite to less formal/polite continuum

More formal/polite					Less formal/polite
<i>nɣ:</i>	<i>nɯ:</i>	<i>nɒŋ</i>	<i>bɣ:</i>	<i>dɣ:</i>	<i>bɑ:</i>

The form /nɣ:/ is used in any setting and level of formality. We note that /nɯ:/ was used by the brothers when speaking to each other or to the young women in the story, and that /nɒŋ/ also was used by the brothers when talking to each other. The others were mostly used by the ‘bad guys’ to show a certain level of disdain, though the younger brother when aroused by the taunts of the Tiau Calang Cloc responded in kind in example (119).

- (114) *sanɯɑ mɔxj ɒ:t ndoŋ nɣ:. kuw? nɒ:ŋ e: pɒ:ʔ lɑ:h sia.*
 now 2SG be.at house **huh** 1SG still want go again as.before
 ‘Now you stay at home, okay? I still want to go again as before.’ [3]
- (115) *ŋkiʔ haj pɒ:ʔ tɒ:p te: ntoʔ nɒj, nɯ: kɒj.*
 thus 1PL(incl) go right.away from place this **huh** friend
 ‘Then we would leave immediately, wouldn’t we friend?’ [169]
- (116) *cɣ: kuw? luw? hu:m ɒ:t kuci:t tɒŋ saruŋ nɒj, nɒŋ aʔe:m.*
 then 1SG truly see be.at die in young.forest this **right** younger.sib.
 ‘Then I saw you here dead in the forest, didn’t I, younger brother.’ [270]
- (117) *ntɣ:n awɣ:j ma: ntɣ:n blɔh, “ntrɔw ɒŋjɑ e:
 simultaneously hold.out but simultaneously ask what 2PL want
 taʔ bɣ: cɒw?”*
 do **huh** grandchild
 ‘They held it out and asked, “What do you want to do, huh grandchildren?”’ [57]
- (118) *khɒn jɔ:k e: tamɔ:ŋ,
 if want want to.live
 ki: la:ŋ ase:ŋ dajh te: ntoʔ ɒnɒj dɣ:.*
 that get.down take.down EXCL from place this **okay**
 ‘If you want to live, then you’d better get down from here, okay?’ [243]
- (119) *kɔp heʔ tɒ: bu:n lɣ:jʔ lakuəj kuaj aɛʔ dɣ:.*
 and 1PL(excl) not play wife person whoever **okay**
 ‘I don’t fool around with anyone’s wife, okay?’ [244]
- (120) *ma: tɒ: kɒ:t cu: nɒ:ŋ te: ntoʔ ɒki: bɑ:. tɒ: daŋ bɑ:,
 but not become return any.more from place there huh not know huh
 heʔ tɒ:ʔ jɔ:k e: lɣ:jʔ e: lɣ: kɔp sampuəʔ heʔ ki:ʔ
 1PL(excl) arrive want want play want play with fiancée 1PL(excl) that
 ‘How come this guy doesn’t leave, huh? Doesn’t he know we want to visit with our
 fiancée?’ [238]*

7.2 Vocative

In a similar way, the vocative expressions show a range of formality and politeness. The form /ɣ:j/ (or its variant form /ɣ:jh/) is used for anyone. Others seem to be used only in informal and in some cases even rude exchanges.

Table 3: Bru vocative expressions arranged along more formal/polite to less formal/polite continuum

More formal/polite			Less formal/polite
ɣ:j/ɣ:jh	ɣ:	Λw	o:

- (121) *kɔ:n ɣ:j sanuə it pjar nɔj da:ʔ kacat tɔŋ ratɔŋ doŋ.*
 child VOC now take flower this put hang.up on wall house
 ‘Child, you take this flower and hang it on the wall of the house.’ [24]
- (122) *cɣ: alɣj pa:j nne:ʔ, “ɣ:! kɔ:n sarɪh! khɔn jɔ:k e:*
 then 3PL say thus Oh child illegitimate.birth if want want
tamɔ:ŋ, kɔj la:ŋ ase:ŋ dajh te: ntoʔ anɔj dɣ:.”
 to.live that get.down take.down EXCL from place this okay
 ‘Then they said, “Hey you bastard, if you want to live, then you’d better get down from here.”’ [243]
- (123) *Λw! anɔj kɔ:n lakuəj mɣj ba:ʔ*
 oh! this child wife 2SG huh
 ‘Oh! So this is you wife, huh?’ [245]
- (124) *ŋkiʔ to:p tɔaw kala:ŋ klɔ:k nɔj thre:k an nne:ʔ,*
 thus group PN this insult 3SG thus
“o:, nɔ:w kɔj ɣ:! naʔ ma: daŋ seʔ tanaj nɔj cɣ:,
 oh someone that VOC how but know how.many day this already
ma: tɔ: kɔ:t cu: nɔ:ŋ te: ntoʔ akɔj ba:.”
 but not become return any.more from place there huh
 ‘So the Tiau Calang Cloc taunted him, “Hey, you! How many days are you going to hang around before you go back where you came from?”’ [238]

7.3 Terms of address

The terms used by the various characters in the story to address each other are a major indicator of social distance and respect on the part of the speaker. They also evoke emotion on the part of the hearer or reader. Terms of address range from the extremely polite /nɔ:w karsa:j / for ‘you’ and /abɔh/ for ‘me’ through the polite plural pronouns to the more familiar singular pronouns. Kinship terms vary in level of respect, but the addition of the third singular pronoun after the kin term makes the form much more polite. Note example (125) below.

Table 4: Bru address terms arranged along more formal/polite to less formal/polite continuum

	More formal/polite			Less formal/polite	
1SG	<i>abah</i>	<i>he?</i>	<i>he?</i>	kinship term	<i>ku?</i>
2PL	<i>nə:w karsa:j</i>	<i>nə:w an</i>	<i>aŋja</i>	kinship term (+) 3SG	<i>mΛj</i>

The contrast between the extremely polite heroes and the brusque and impolite Grandfather and Grandmother Soc Catuc comes through very clearly in the way the animals are addressed as these characters meet them blocking the way. The Bru value placed on politeness comes through frequently in the speech of the two brothers. As the two brothers meet the buffalo fighting in the road, their request is respectful and urgent, as shown in example (125).

- (125) *ŋki? ba:r na? se:m a:j nΛj se:? nne?*, “*ʔj! bak an ʔj!*
 thus two person brothers this request thus VOC **uncle** 3SG VOC
nunŋ, he? se:? ba:r lam bak an co:j ta? ŋki? juah.
 REQ 1PL(excl) request two CL **uncle** 3SG don't do thus yet
he? se:? tamɔ:ŋ se:? bak an pa:ŋ rana: jo:n he?
 1PL(excl) request live request **uncle** 3SG make.way road for 1PL(excl)
luh mahoj wə:j.
 run moment beforehand
 ‘The two brothers asked, “Oh respected uncles, we ask you two uncles to not do this yet. We ask to live. We ask you uncles to make way for a few moments so we can run by.”’ [77]

Coming along a bit later, Grandfather and Grandmother Soc Catuc speak to the pigs brusquely in example (126):

- (126) *alɔj ʔn ba:r lam ali:k tɔ:l nΛj pa:ŋ rana: jo:n alɔj pΛ:?. alɔj*
 3PL order two CL pig male this make.way road for 3PL go 3PL
pa:j nε:?, “Λw! aŋja pa:ŋ rana: he? e: pΛ:?. nΛj.”
 say like.this oh! 2PL make.way road 1PL(excl) want go this
 ‘They ordered the two male pigs to get out of the way so they could go. They said,
 “Hey, you get out of the way. We want to go here.”’ [79]

Even in face-to-face encounters with the Tiau Calang Cloc, the brothers continue to speak politely until goaded beyond what they can tolerate. The difference can be clearly seen in the following exchange where Tasuo has come to look for his younger brother’s body. They speak rudely to him, calling him by the shocking term /kɔ:n aɔ:/ ‘child of a dog’. But he answers with extremely polite speech, addressing them with the polite /nə:w karsa:j / and referring to himself as /abah/.

(127) *ŋki? tu: to:p t̪iaw kala:ŋ klɔ:k nɔj hu:m tasuə, a:j tanu: ki:,*
 thus when group PN this see PN older.bro. PN that

cɣ: alɣj thre:k sia, “ʌw! n̄a:w n̄a:ŋ ʌ:t tʌŋ, nto? ki: ɣ:?”
 then 3PL insult as.before oh! who still be.at in place that oh

na? aɾiŋ pa:j an ka: kɔn aɔ: ki: sia ba:? lu? pla:
 why like say 3SG who **child dog** that as.before huh truly powerful

ɣ:j?. na? tʌ: bu:n la:p juah? n̄a:ŋ jɔ:k e: kaŋeŋ kane:ŋ
 EMPH why not concede.defeat yet still want want threaten

sia ba:? kuci:t ɣ:j?! lu? kɔ:ŋ n̄a:ŋ kaɟ, mah ka: haj
 as.before huh die EMPH very daring still friend like who 1PL(incl)

khɔ:j? ta? an cɣ:, ma: n̄a:ŋ an tʌ: sʌŋ ŋkah sia.”

PST attack 3SG already but still 3SG not feel afraid as.before
 ‘When the Tiau Calang Cloc group saw Tasuo, the older brother of Tangu, they taunted him, “Hey, are you still around? Looks like that child of a dog is still here. He is surely powerful. How come you don’t give up? Do you still want to threaten us? How amazing! After what we did to him, he still isn’t afraid.” [263]

ŋki? ŋe:? na? to:p t̪iaw kala:ŋ klɔ:k nɔj canɣm pa:j la: tanu:
 thus all person group PN this think say be PN

sia, ma: alɣj tʌ: bu:n daŋ pa:j a:j tanu:. cɣ: tasuə,
 as.before but 3PL not know say older.bro. PN then PN

a:j tanu: nɔj, taʔʌ:j alɣj t̪iaw kala:ŋ klɔ:k nne:?, “ɣ:?,
 older.bro. PN this reply 3PL PN thus yes

n̄a:w karsaj ɣ:j! abah tʌ: bu:n tʌ:? jɔ:k ce:ŋ
 someone adult VOC 1SG(polite) not arrive want contest

sampuə? n̄a:w karsaj ma: ma: abah tʌ:? pa:
 fiancée someone adult NEG-EMPH but 1SG(polite) arrive direction

nɔj la: abah tʌ:? cɣa? la:h a?e:m abah,
 this be 1SG(polite) arrive look.for again younger.sib. 1SG(polite)

an ka: khɔ:j? pu:t ki:.”
 3SG who PST destroy that

‘All of the Tiau Calang Cloc group thought that this was Tangu. They didn’t know it was Tangu’s older brother. Then Tasuo, Tangu’s older brother, answered the Tiau Calang Cloc, “Oh gentlemen! Your servant has not come wanting to vie for your fiancée, but I have come to look for my younger brother. He has been lost.” [264]

Shock and anger are evoked by the use of terms such as ‘child of a dog’ in example (127) and ‘bastard’ in example (122). And when Tangu has finally had all the abuse he can take, he refers to himself defiantly as /ntre:h/, using simply the classifier for long slender objects. But the threat implied by his use of this term in example (128) to refer to himself is obvious to a Bru audience.

- (128) *khlan ntrɛ:h nɔj ma: nɔ:ŋ tamɔ:ŋ, ki: co:j ni:ʔ pampla:*
 if CL this but still live that don't compare powerful
mahlɛʔ.
 how.much

'As long as I am still living then don't try to compare who is most powerful.' [278]

The singular pronoun forms /kɔʔ/ 'I, me' and /mɔj/ 'you' are typically used between peers and those who are well acquainted with each other. For example, at the beginning of the story, the husband and wife used these terms of address with each other. The tiger mother and cow mother addressed each other with singular pronouns, as did also Grandfather and Grandmother Soc Catuc.

- (129) *ŋkiʔ kaja:k nɔj pa:j kɔp lakuəj an nne:ʔ,*
 thus husband this say with wife 3SG thus

“sanua mɔj taʔ doŋ saraj muəj na:ʔ mɔj saŋ.
 now 2SG make house field one person 2SG only

kɔʔ la: pɔ:ʔ ceʔ ca:ŋ. du:n naʔ kɔʔ bu:n cu:.”
 1SG be go sell buy long.time how 1SG have return

'So the husband said to his wife, "Now you stay here and mind the house and field by yourself. I am going to buy and sell. After a while I'll come back.'" [1]

More polite and less familiar are the plural pronoun forms /heʔ/ 'we (exclusive)' and /aŋja/ 'you (plural)'. The tiger having been caught in a compromising position by the husband speaks ingratiatingly to him.

- (130) *bɔ: an se:ŋ te: saru:ŋ ki:, cɔ: an pa:j ne:ʔ, “heʔ*
 when 3SG go.down from loft that then 3SG say like.this 1PL(excl)

khɔ:jʔ bu:n lo:jh cɔʔ aŋja cɔ:, kala: ɔ:j. aŋja khɔ:jʔ
 PST have sin against 2PL already friend VOC 2PL PST

pupɔ:ʔ te: doŋ, ma: heʔ co:n tɔŋ saru:ŋ doŋ aŋja ŋkiʔ.”
 walk from house but 1PL(excl) go.up in loft house 2PL thus

'When he went down from the loft he said, "I have wronged you, friend. You were gone from home, but I went into the loft of your house like this.'" [10]

Kin terms are used between the brothers, between the brothers and the young women and with the older people. Sometimes a combination of pronoun and kin term is used. While the two boys usually addressed each other with the sibling terms, they also used the familiar first and second person pronouns. With the various grandparent figures they usually used the term 'Grandfather' or 'Grandmother' and were addressed as 'Grandchild.'

Only once did Tangu address his older brother formally. This was when he told his brother he was leaving his wife and family for Miss Creq Phuom and life in the underworld. His brother was not happy about the choice but agreed not to abandon him. Tangu accepted responsibility for his decision but again asked his brother not to reject him, using the formal term of 'uncle' plus the third person singular pronoun to speak to his brother.

- (131) *cɣ:* *aʔe:m* *an taʔʌ:j* *lɑ:h* *an ka: a:j* *nne:ʔ*, “*te: ranaʔ*
 then younger.sib. 3SG reply again 3SG who older.bro. thus about thing
nʌj, *ta:m* *kuʔ*. *se:ʔ* ***bak an*** *co:j* *tah* *tɔ:j* *kuʔ*
 this it’s.up.to 1SG request **uncle** 3SG don’t discard throw.away 1SG
aʔe:m *nʌj*, *la: kuʔ aʔe:m* *saʔʌ:n* *tʌ:p*.”
 younger.sib this be 1SG younger.sib. thank right.away
 ‘Then the younger brother replied to his older brother, “As to this matter, I’ll take
 full responsibility. Please don’t turn away from me, your younger brother, and I
 will be very grateful.”’ [234]

8 Conclusion

While we feel we’ve learned a great deal in processing and studying this narrative text, much more could profitably be done. We would like to get further input from Bru readers and editors. Grimes notes that ‘the texts that yield the most consistent analysis are edited texts’ (1978:33). Well-formed texts which have been carefully edited are those most easy for hearers and readers to understand. The development of a written style is an emerging area of expertise within the Bru community. Further editing by Bru readers will likely improve the quality of an already well-formed text.

We’d also like to be able to investigate more fully the nuances of the terms of address and styles of speaking as well as to chart more carefully the time progression and logical progression of the story. And finally, we’d like to encourage Bru writers to follow the example of Mpoaq Saniq and bring into written form more of their rich oral tradition.

References

- Anderson, Stephen R., 1985, Typological distinctions. In Timothy Shopen, ed. 1985:3–56.
- Comrie, Bernard, 1981, *Language universals and linguistic typology: syntax and morphology*. Chicago: University of Chicago Press.
- Gregerson, Kenneth and Martha Martens, 1986, Perfective *mi* in Uma discourse. In *Languages in global perspective: papers in honor of the 50th anniversary of the Summer Institute of Linguistics 1935–1985*, 163–176. Dallas: SIL.
- Grimes, Joseph E., ed., 1978, Introduction. In *Papers on discourse*, p.viii. Dallas: SIL.
 1984, *The thread of discourse*. The Hague: Mouton Publishers.
- Howard, Olive M., 1978, The paragraph in Gagou (Gban) narrative. In *Papers on discourse*, 273–282. Dallas: SIL.
- Levinson, Stephen, ed., 1990, *Thematic continuity and development in languages of Sabah*. Canberra: Pacific Linguistics.
- Longacre, Robert E., 1985, Sentences as combinations of clauses. In Timothy Shopen, ed. 1985:235–286.
 1996, *The grammar of discourse*. New York and London: Plenum Press. (Second edition.)
- Miller, Carolyn P., 1987, Assessing levels of complexity in Kadazan texts. *Word* 1:38: 27–45.

- Miller, John D. and Carolyn P. Miller, 1990, Thematic continuity and development in Coastal Kadazan narratives. In Stephen H. Levinson, ed. 1990:105–135.
- Pickering, Wilbur, 1980, *A framework for discourse analysis*. Dallas: SIL and The University of Texas at Arlington.
- Shopen, Timothy, ed., 1985, *Language typology and syntactic description* Vol.III, *Grammatical categories and the lexicon*. Cambridge: Cambridge University Press.
- Thompson, Sandra A. and Robert E. Longacre, 1985, Adverbial clauses. In Timothy Shopen, ed. 1985:171–234.
- Watson, Richard, 1977, Discourse elements in a Pacoh narrative. *Mon-Khmer Studies* VI:279–322.
- 1980, *A grammar of two Pacoh texts*. Arlington: The University of Texas at Arlington.

8 *The temporal movement of the Hlai (Li) origin myth*

SOMSONGE BURUSPHAT

1 Introduction to the Hlai (Li) language

The Hlai (Li) language belongs to the Kadai language family. Its genetic relationship with other relatives is as follows:

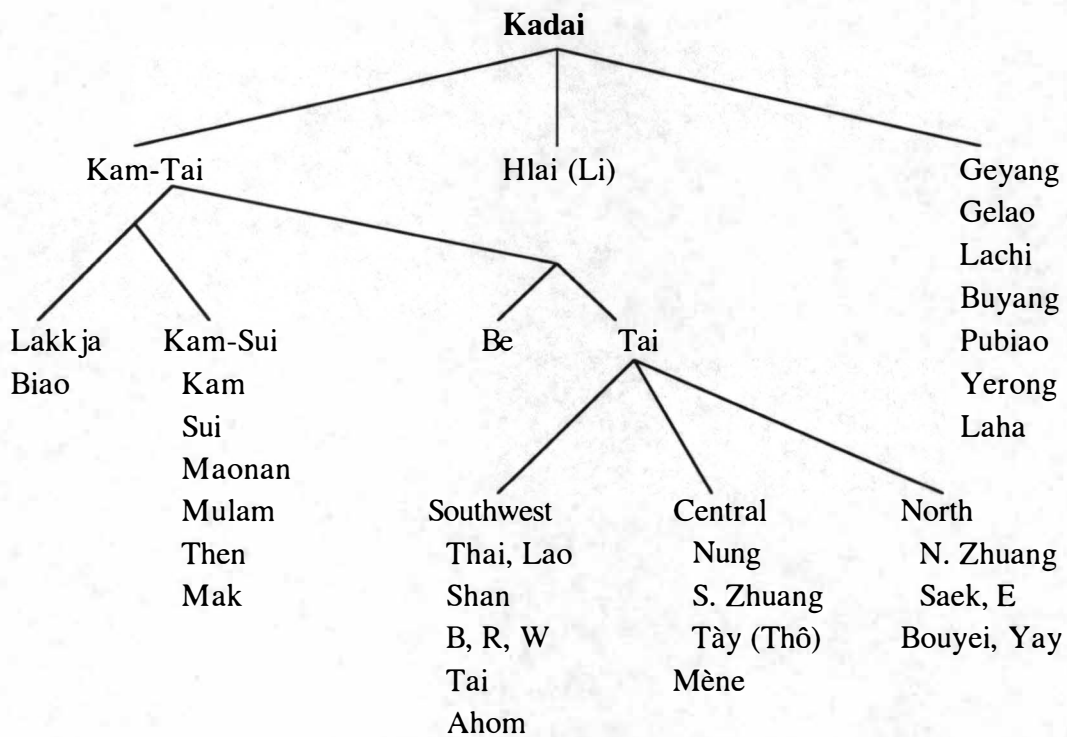


Figure 1: Classification of Kadai languages (Edmondson & Solnit 1997:2)

The Hlai speakers have a population of 1,110,900 living mainly in Hainan province, China (Zhou 1996). The Hlai language (the speech of Baoding in Hainan province) has three tones on open syllables (with tone contours of 53, 55, and 11) and includes the following initial consonants:

Table 1: Hlai initial consonants (Zhou 1996:59)

<i>p</i>	<i>ph</i>	<i>pl</i>	<i>ʔb</i>	<i>m</i>	<i>ʔw</i>	<i>f</i>	<i>v</i>
<i>t</i>	<i>th</i>		<i>ʔd</i>	<i>n</i>		<i>l</i>	<i>ʔ</i>
<i>ts</i>	<i>tsh</i>						<i>r</i>
				<i>n</i>	<i>ʔj</i>		<i>z</i>
<i>k</i>	<i>kh</i>		<i>g</i>	<i>ŋ</i>			
<i>ʔ</i>						<i>h</i>	
	<i>khw</i>		<i>gw</i>	<i>ŋw</i>		<i>hw</i>	<i>hj</i>

The Hlai data used for this discourse analysis is the origin myth taken from *Kam-Tai oral literatures* (Burusphat & Sinnott, eds 1998). To understand the sample given in this paper, the overall content of the myth arranged by episodes in §5, Temporal connectives, should be studied first.

2 Narrative timeline and temporal movement

The narrative timeline is defined as the main line of development, i.e., the line of sequential, punctiliar happenings (Longacre 1996). In addition to the chronological order of happenings, Dry (1983:33) points out that, 'for a structure to move time, it must present new information'. A perception of temporal movement along a narrative timeline is triggered by clauses which make reference to temporal points, rather than spans of time, and that these points are usually the endpoints of situations. The clauses which refer to sequenced points on a timeline and present new information, rather than referring anaphorically to happenings that have already occurred, constitute the foreground of a narrative. The foregrounded clauses are thus defined as clauses in which the action of the narrative takes up and time begins to move forward.

This paper aims to present the grammatical constructions of Hlai origin myth which are relevant to time-movement, namely, the main clauses which recount events on a timeline, the anaphoric temporal clauses which make reference to the final endpoint of a happening, the temporal connectives which establish temporal movement, and the topic subject of time-moving clauses which is related closely to the sequenced happenings.

3 Temporal movement in main clauses

It is the main clauses rather than the dependent clauses which appear to move time. The time-moving main clauses convey new information. The idea of old (given) information, i.e. already activated information, and new information, i.e. newly activated information, is defined by Chafe (1976:30) as follows:

Given (or old) information is that knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance. So-called new information is what the speaker assumes he is introducing into the addressee's consciousness by what he says.

The main clause which triggers the listener's perception of time movement presents a new happening which is sequenced with the preceding happening. The clause which refers to a happening which is overlapping with or simultaneous with the happening in the main clause is background information and does not move time. The clause which is stative, descriptive, or expository does not propel narrative time forward.

The Hlai origin myth is a lengthy text consisting of 170 main clauses. Based on the conditions on chronological ordering and new information, there are 134 time-moving main clauses. The rest of the main clauses do not move time but characterise the supportive materials such as setting (example 1), repetitive activities (example 2), and background situation which does not refer to sequenced temporal points but to indefinite spans of time (example 3).

- (1) *noŋ¹¹ gwau¹¹ tsau⁵⁵ tsu⁵⁵ laŋ ʔu⁵⁵ ʔa:u⁵³*
 once.upon.a.time have one CLF people give.birth
ʔwe:ŋ⁵³ ʔau¹¹ laŋ⁵³ lu:k⁵⁵
 to two CLF child
tsu⁵³ laŋ⁵³ ʔu:k⁵⁵ pu⁵³ k^hau⁵⁵ tsu⁵⁵ laŋ⁵⁴ ʔu:k⁵⁵ p^ha¹¹ ma:n⁵³
 one CLF child girl one CLF child boy
 'Once upon a time, there was a man who had two children, a girl and a boy.'
- (2) *pa¹¹ pa:n⁵³ hau⁵⁵ ʔai¹¹ la:i¹¹ ʔu:k⁵⁵ ʔuma:n⁵³*
 new.woman that discriminate.against child old
ran⁵⁵ hwan⁵³ ta⁵³ t^ha:r⁵⁵ kom¹¹ ts^ha¹¹
 every day not beat then curse
 'The stepmother treated the children badly. Everyday she beat and cursed them.'
- (3) *ʔu:k⁵⁵ k^hau¹¹ la:i¹¹ ʔu:k⁵⁵ gu:ŋ⁵³ ŋai¹¹ na⁵³ ŋan⁵³ ŋai¹¹*
 elder.sister see younger.brother cry she also cry
ʔu:ŋ⁵⁵ t^ho:ŋ¹¹ ŋai¹¹ ŋai¹¹ ʔau¹¹ tsu:n⁵³ du¹¹ hau⁵⁵
 together cry cry two people stay there
 'When the elder sister saw her younger brother cry, she also cried. Together they cried there.'

Of all time-moving main clauses, 117 clauses are introduced by the sequential indicator *fan⁵³* 'then' which first appears in the story when the narrative timeline actually gets underway. The major function of *fan⁵³* is to provide chronological narrative timeline cohesion to the origin myth and thus it gives prominence to the main line of development. Example (4) illustrates the movement of the timeline triggered by the sequential indicator *fan⁵³*.

- (4) *ʔu:k⁵⁵ k^hau¹¹ fan⁵³ ri:n⁵³*
 elder.sister then say
ga⁵³ gwa⁵³ pu⁵³ ra⁵⁵ ro:i¹¹ hau⁵⁵
 we plant in place.where.the.deer.died that

*ba:i*¹¹ *gwa*⁵³ *ta*⁵³ *ta:u*¹¹ *fan*⁵³ *la:i*¹¹ *ʔu:t*⁵⁵
 finish plant not grow **then** see sprout

*ʔu:t*⁵⁵ *fan*⁵³ *vu:k*⁵⁵ *tʰe:ŋ*¹¹ *kʰa:n*⁵⁵ *ga:i*¹¹ *fan*⁵³ *kʰa:n* *tsʰo:n*⁵³
 sprout **then** make frame climb gourd **then** climb.up grow.fruit

‘The elder sister then said, “Let’s plant them in the place where the deer died.” Not long after they planted the rice and gourd, a sprout grew out. They made a frame for the vine of the gourd to climb up.’

However, not all time-moving clauses are accompanied by *fan*⁵³; 17 time-moving clauses do not have this sequential indicator because the overt reference to sequenced points seems to be unnecessary. They are clauses of direct and indirect quotation, clauses expressing similar situation to the previous one, and clauses featuring simultaneous events.

Clauses of direct and indirect quotation occasionally are not preceded by *fan*⁵³ because the chronological order of the acts of speaking can be interpreted by the pragmatic knowledge of the situations. In example (5) our knowledge of the world tells us that the elder sister’s act of speaking is followed by God’s speech. It should be noted that in the absence of the sequential indicator *fan*⁵³ the adverb *lom*¹¹ ‘again’ is used to indicate that the act of speaking occurs more than once.

(5) *ʔu:k*⁵⁵ *kʰau*¹¹ *lom*¹¹ *ri:n*⁵³ *ʔi*¹¹ *tsʰi* *hau*⁵⁵ *dorj*⁵³ *ra*¹¹
 elder.sister **again** say dear me that how

*fa*⁵³ *ne*⁵⁵ *ʔu:k*⁵⁵ *kʰau*¹¹ *kʰu*⁵³ *ʔu:k*⁵⁵ *gu:ŋ*⁵³
 we this elder.sister and younger.brother

*na:m*¹¹ *lom*¹¹ *ri:n*⁵³ *han*¹¹ *ne*⁵⁵ *ʔu:ŋ*¹¹ *na:i*¹¹ *ʔa:u*⁵³
 God **again** say now this no other people

*meu*⁵³ *ta*⁵³ *kom*¹¹ *peu*⁵³ *tʰo:ŋ*¹¹
 you should marry companion

‘The elder sister again said, “How can we, brother and sister, get married?” God again said, “Now there are no people in the world, so you have to get married.”’

Example (6) below illustrates two clauses which characterise a similar activity to the previous one which is overtly indicated by the sequential indicator *fan*⁵³. The similarity of activities is marked by the adverb *lom*¹¹ ‘again’. The sequential indicator is not needed here since the chronological order of activity can be predicted from the previous similar activity. These two clauses also represent two simultaneous events on the narrative timeline. Though it is emphasised that a structure that moves time must present a situation in chronological order, two situations or events may overlap or take place at the same time and yet they develop the narrative timeline.

(6) *ʔu:k*⁵⁵ *gu:ŋ*⁵³ *lom*¹¹ *teu*¹¹ *ʔe:ŋ*⁵⁵ *ko:ŋ*⁵³ *tʰa*⁵⁵ *ri:n*⁵³ *kʰwei*¹¹ *la*⁵⁵ *tʰa*⁵⁵
 younger.brother **again** point.to jar food say want eat food

‘The younger brother again pointed to the jar of food and said he wanted to eat some food.’

While the occurrence of the sequential indicator *fan*⁵³ is tied to the temporal movement of the origin myth, there is one instance in which a clause preceded by *fan*⁵³ does not advance the narrative timeline. This is the case when *fan*⁵³ appears in a quote in which past happenings are reported in chronological order by a participant. In example (7) the elder sister told the

Thunder God how their stepmother brought her and her younger brother to the mountain, left them there, and went away. The past happenings are told in chronological order indicated by *fan*⁵³.

- (7) *ʔu:k*⁵⁵ *k^hau*¹¹ **fan**⁵³ *t^hen*⁵³
 elder.sister **then** answer
*pa*¹¹ *za*⁵³ *pe*⁵³ *gu*¹¹ *fə*⁵³ *pu:n*⁵⁵ *da:n*¹¹ *ne*⁵⁵ **fan**⁵³ *gou*⁵⁵ *ba:r*⁵⁵ *fə*⁵³
 step.mother bring us come arrive here **then** run.away from us
 ‘The elder sister then answered, “Our stepmother brought us here and then ran away from us.”’

4 Anaphoric temporal clauses

Longacre (1996) points out that an action or any category of happening may be demoted from a primary happening on the narrative timeline to the happening of lesser importance by subordinating a clause whose verb reports the action, especially in preposed adverbial clauses and in relative clauses.

The subordinate clauses found in the Hlai origin myth are clauses which are demoted to encode ongoing activities, script-predictable happenings, and anaphoric happenings. All of these clauses do not advance the narrative timeline but they are discussed here because the anaphoric temporal clauses are relevant to the temporal movement of the main clauses. Before proceeding to the discussion of the anaphoric temporal clauses, I believe a distinction of this clause type and others should be made.

A subordinate clause expressing an ongoing activity does not refer to sequenced points but makes reference to a span of time indicated by the temporal connective *ts^hɿ*⁵³ ‘while’ as in example (8).

- (8) *ts^hɿ*⁵³ *la*⁵⁵ *t^ha*⁵⁵ *fan*⁵³ *la:i*¹¹ *tsu*⁵⁵ *tsuen*⁵³ *ʔa:u*⁵³ *pu:n*⁵³ *ga:m*⁵³
while eat food then see one CLF person come ask
*meu*⁵³ *ta*⁵³ *vu:k*⁵⁵ *me*¹¹ *he*¹¹ *du*¹¹ *ne*⁵⁵
 you for what stay here
 ‘While they were eating, a stranger came and asked, “Why are you here?”’

Another kind of subordinate clause encodes a script-predictable happening which is preparatory to or resultant from happenings reported on the primary narrative timeline as in example (9).

- (9) *tsau*⁵⁵ *tsu*⁵⁵ *hwan*⁵³
 have one day
*fan*⁵³ *deu*⁵³ *t^ha*⁵⁵ *du*¹¹ *ʔe:ɲ*⁵⁵ *ko:ɲ*⁵⁵ *gui*¹¹ *ʔau*¹¹
 then take food put jar bring two
*laɲ*⁵³ *ʔu:k*⁵⁵ *ʔu*¹¹ *ma:n*⁵³ *he*⁵³ *k^ha:n*⁵³ *hwau*¹¹
 CLF child old go up mountain
*da:n*¹¹ *hwau*¹¹ *ma*⁵⁵ *fan*⁵⁵ *ga:m*⁵³ *ʔu:k*⁵⁵
arrive mountain **that** then ask child
*meu*⁵³ *ta*⁵³ *k^hu:ɲ*⁵³ *gweu*⁵³ *ne*⁵⁵ *ts^ha*¹¹ *ʔum*¹¹
 you know place this not know

'One day, the stepmother brought out a jar full of rice, and brought the two children into the mountain. When they arrived at the mountain, the stepmother then asked the children, "Do you know this place?"'

The subordinate clause which is an anaphoric temporal clause makes reference to a temporal point which is the endpoint of situation. In referring anaphorically to happenings that have already occurred, the anaphoric temporal clause functions to link the previous happening to the next one. It identifies two sequenced points on the timeline, that is, the completion of the previous happening and the beginning of the next happening. Therefore, the main clause preceded by the anaphoric temporal clause is sequenced with the previous time-moving clause and thus triggers the perception of temporal movement. It should be noted also that the anaphoric temporal clauses are always accompanied by the completive marker *ba:i'* which has a variation of occurrence, either preceding or following a verb. There is one instance in which an anaphoric temporal clause is simply reduced to the completive marker *ba:i'* functioning as the main verb meaning 'finish' as in example (10).

- (10) *na⁵³ fan⁵⁵ tak⁵⁵ ba:i' fan⁵³ fo:n⁵⁵*
 she then chop **finish** then scatter
 'She then chopped the baby into pieces. After chopping, she then scattered the pieces.'

Example (11) illustrates the frequent use of anaphoric temporal clauses accompanied by the completive marker *ba:i'*.

- (11) *na:m' fan⁵³ ?i:k⁵⁵ nam' na⁵³ ?an⁵³ lom' ?ai' ts^hin⁵³*
 God then strike water she also again not believe
na:m' fan⁵³ peu⁵³
 God then go.back
na:m' ba:i' he⁵³ ba:r⁵³ ?u:k⁵⁵ gu:ŋ⁵³ fan⁵³ ?u:t⁵⁵ ha:i'
 God already go leave younger.brother then defecate faeces
ba:i' ha:i' ba:i' fan⁵³ ga:m⁵³ ?u:k⁵⁵ k^hau'
 already defecate over then ask elder.sister
khwei' pha:r⁵³ ha:i' phai' ra'
 want wipe faeces put where
?u:k⁵⁵ k^hau' fan⁵³ ri:n⁵³
 elder.sister then say
p^ha:r⁵³ pu⁵³ ts^har⁵³ ?u' ren⁵⁵ ma⁵⁵ re⁵³
 wipe put wood PRT upright that PRT
?u:k⁵⁵ gu:ŋ⁵³ p^ha:r⁵³ ba:i' fan⁵³ k^hom⁵³ ha:i' ts^her⁵³
 younger.brother wipe finish then itch anus
 'God then struck water. She again did not believe him. God then went back. After God left, the younger brother went to defecate. When finished, he then asked his elder sister, "What can I use to clean my anus?" The elder sister then said, "Use a tree branch." When the younger brother finished cleaning his anus, he felt an itch on his anus.'

5 Temporal connectives

The sequentiality of narrative timeline is also indicated by temporal connectives most of which function to mark episode boundary. These temporal connectives chronologically link the following happenings to the previous ones. They are illustrated with a free translation of the story in example (12) below.

(12) Episode 1

‘Once upon a time, there was a man who had a son and a daughter. Later, his wife died. He then married a new woman who treated the boy and the girl very badly.’

Episode 2

tsau⁵⁵ tsu⁵⁵ hwan⁵³ ‘One day’

‘One day, the stepmother cheated the boy and the girl out of their home and left them in the wild woods. A stranger came to ask what happened to them and told them to stay where they were to pass the night.’

Episode 3

fā¹¹ ba:i¹¹ ts^hap⁵⁵ ‘It’s already dark.’

‘The boy and the girl were frightened by the bears so they went to beg the vines to protect them from the bears. The white vine and red vine stretched out to form a thick wall to surround the two of them. The boy and the girl then passed the night safely.’

Episode 4

tsu⁵⁵ dau¹¹ fā¹¹ lom¹¹ den¹¹ ‘It was bright again.’

‘Another stranger came to ask the boy and the girl why they were there and told them to get married. The boy and the girl refused to do so because they were afraid of the Thunder God. So the stranger told them that he himself was the Thunder God and tried to prove to them that he really was the Thunder God but they did not believe him. So he left. After the Thunder God left, they found a dead deer. With the help of a group of crows and an ape, they were able to cook the deer for food. While they were eating, a turtledove asked them to get rice and gourd seeds from its own body. The younger brother was able to shoot the turtledove with a bow and arrow. They cut open the turtledove, took out the rice and gourd seeds and planted them. Later they found rice seeds inside the gourd so they had that rice as their food. They used one of the gourds which was very big as a granary.’

Episode 5

tsau⁵⁵ tsu⁵⁵ hwan⁵³ ‘One day’

‘One day, the brother and sister went to see their father. They picked bananas and caught a poisonous snake. When they arrived at their father’s house, they gave their stepmother a gourd with the snake inside. Their stepmother was bitten by the snake and died. On the way back, it was raining hard. When the brother and sister arrived at the place where they lived, they saw the place had almost been submerged by the flood. So they put their belongings in the big gourd and

hid themselves inside the gourd. Later the flood disappeared. People in the world died because of the flood and only the two of them were alive. The Thunder God then appeared again and asked them to be husband and wife but they resisted. During the night the Thunder God made them roll into each other's bed. They finally became husband and wife.'

Episode 6

tau' 'long time'

'After a period of time, a baby shaped like a pile of cattle dung was born. Following the Thunder God's advice, the elder sister chopped the baby into small pieces and scattered them in the field. However, she did not say "More Li people, fewer Han people" as told by the Thunder God. On the contrary, she said the opposite.'

Episode 7

ʔu'' hau⁵⁵ 'the next day'

'The next day they went to the field and saw the small pieces of flesh had become human beings. So they gave them names. They named the smaller group "Li" and the bigger group "Han." They told them to remember their own names.'

Besides the temporal connectives displayed above, there are few temporal connectives which behave as a micro-device linking successive happenings within the same episode. They are *ts^hɿ⁵³ neɿ⁵⁵* 'just then' which occurs twice in example (13) and *puɿ⁵³ ts^hap⁵⁵* 'at night' which occurs once in example (14).

- (13) a. *ts^hɿ⁵³ neɿ⁵⁵ paɿ'' pa:ɿ⁵³ faɿ⁵³ thua'' lu:k⁵⁵*
just.then new.woman then cheat child
 'Just then the new woman cheated the children.'
- b. *ts^hɿ⁵³ neɿ⁵⁵ faɿ⁵³ la:i'' veɿ⁵³ tsiu'' kha⁵³ tshaɿ⁵³*
just.then then see ape climb tree.branch
 'Just then they suddenly saw an ape climbing up a tree.'
- (14) *puɿ⁵³ ts^hap⁵⁵ ʔu:k⁵⁵ khau'' ʔu:ɿ⁵⁵ ʔu:k⁵⁵ gu:ɿ⁵³ ʔa'' ra''*
at.night elder.sister and younger.brother who
kau⁵³ the:ɿ⁵⁵ ʔa'' ra'' kau⁵⁵ laɿ⁵³ laɿ⁵³
 sleep bed who sleep far.away
 'At night the elder sister and the younger brother who lay in different beds and were far away from each other...'

6 Topicality of the subject of the time-moving clause

Another important feature of the time-moving clause is its topic subject, i.e. the noun phrase designating the topic of the discourse used as the subject of a clause. Once the topic subject has been established in an initial clause, it will convey given information which is subject to pronominalisation in the following clauses as far as the subject continues. Chafe (1976:31) mentions that pronominalisation 'can be applied only to items that convey given

information, but it tends not to be applied when the speaker is aware that ambiguity would result (when there are two or more given items competing equally for the same pronoun)'.¹

Givón (1983:8) points out the relationship between the topic subject and the sequenced clauses as follows:

Within the thematic paragraph it is most common for one topic to be the continuity marker, the leitmotif, so that it is the participant most crucially involved in the action sequence running through the paragraph; it is the participant most closely associated with the higher-level 'theme' of the paragraph; and finally, it is the participant most likely to be coded as the primary topic — or grammatical subject — of the vast majority of sequentially ordered clauses/sentences comprising the thematic paragraph. It is thus, obviously, the most continuous of all the topics mentioned in the various clauses in the paragraph.

When the topic subject continues, it constitutes a topic chain which is defined as, 'where the topic established in the first clause serves as the referent for the unrealised topics in the chain of clauses following it' (Li & Thompson 1979:313). In the Hlai origin myth, the topic subject is first introduced by a noun phrase. And when the topic subject continues, a pronoun or zero anaphor is used. Of all the time-moving main clauses, 88 of them have either a pronoun or zero anaphor as the subject. The rest of the clauses have subjects represented by noun phrases which are usually used when there is a switch of topic subject. The following example illustrates a topic chain.

- (15) *ʔu:k⁵⁵ k^hau¹¹ fan⁵³ heɿ⁵³ kiu⁵³*
 elder.sister then go see
∅ fan⁵³ la:i¹¹ ka:t⁵⁵ mo:k⁵⁵ duɿ¹¹ ploŋ¹¹ ka¹¹ hau⁵⁵
she then see frog pronate in cage that
ro:ŋ⁵³ pe:t¹¹ pe:t¹¹ pe:t¹¹
 cry pee pee pee
∅ fan⁵³ ri:n⁵³ tiŋ¹¹ la⁵⁵ pai¹¹ ka:t⁵⁵ neɿ⁵⁵
she then say ghost eat big frog this
ri:n⁵³ ba:i¹¹ ∅ fan⁵² po:k⁵⁵ pai¹¹ ka:t⁵⁵ t^ha:ɿ⁵⁵ ʔa:u⁵⁵
 say over **she** then catch big frog beat die
ka:t⁵⁵ ba:i¹¹ ʔa:u⁵⁵ na⁵³ fan⁵³ peu⁵³ ʔu:ŋ⁵⁵ ʔu:k⁵⁵ gu:ŋ⁵³
 frog already die **she** then back accompany younger brother
 'The elder sister went to see what it was. She saw a frog crying, "Pee! pee! pee!"
 She said, "Ghost, eat this big frog." After saying that, she killed the frog. After
 the frog was killed, she went back to take care of her younger brother.'

In this example, only the subject of the first clause is marked by an overt form of the topic. The following clauses in the series have subjects expressed by zero anaphors and the pronoun *na⁵³* 'she'. All main clauses in this topic chain propel the narrative time forward.

7 Conclusion

This paper presents a discourse analysis of one piece of lengthy oral literature in the Hlai language, i.e. the Hlai origin myth, based mainly on the framework of Longacre (1996) and Dry (1983). The notion of new/given information comes from the work of Chafe (1976) and that of topic continuity from the works of Givón (1983) and Li and Thompson (1979).

This paper has focused on the grammatical features that trigger perception of temporal movement in the Hlai origin myth. A narrative discourse consists of an intricate interweaving of noun phrases for participants, props and verb phrases for events and many non-events. The elements that move time are the main clauses whose verb phrases carry new information and are sequenced with the previous time-moving clauses. The chronological ordering of happenings encoded by the foreground verb phrases is signalled by the sequential indicator *fan*⁵³ which gives prominence to the happenings on the narrative timeline as well as anaphoric temporal clauses which mark the end point of the preceding happening. In addition to these sequential indicators, the origin myth is also broken up into episodes which are chronologically linked together by temporal connectives. While all verb phrases of time-moving clauses convey new information, the subject noun phrases are topical and carry given information. They are occasionally realised by a noun phrase, and frequently by a pronoun and zero anaphor. Over fifty per cent of the time-moving clauses have zero anaphor as their unrealised subject.

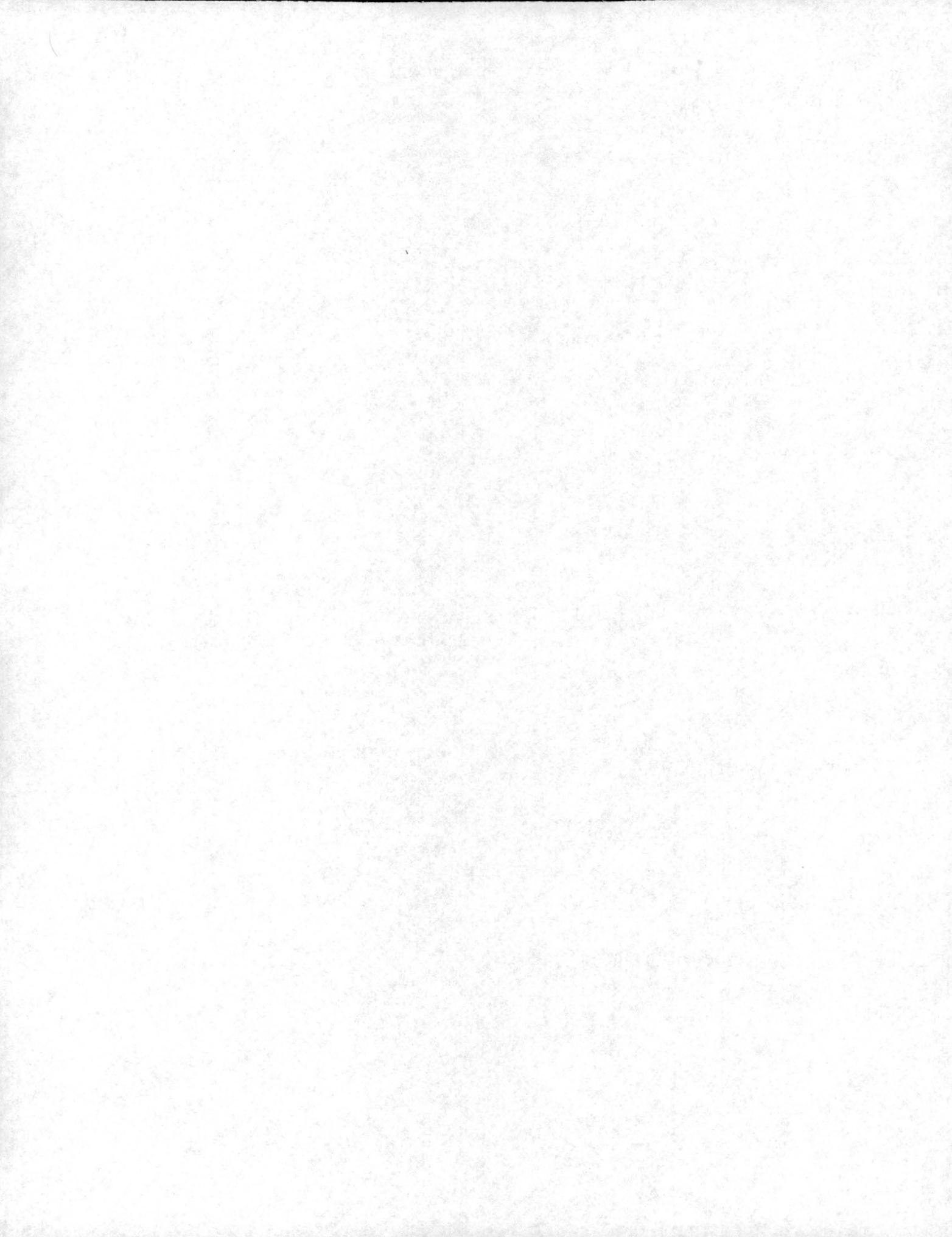
The discourse study of origin myth in the Hlai language has revealed that at the discourse level some discourse features of this language are similar to Bangkok Thai, its distant cousin language; for example, the use of *fan*⁵³ which is close to the preserial verbs *ko:*³ 'then' and *cuŋ*⁷ 'consequently', the frequent use of anaphoric temporal clauses tagged by *ba:i*¹¹ which is similar to *læ:w*⁶ 'already', and the regular use of the topic chain construction to propel time forward.

References

- Burusphat, Somsonge and Megan Sinnott, eds, 1998, *Kam-Tai oral literatures*. Bangkok: T&D Printing Ltd.
- Burusphat, Somsonge, Suriya Ratanakul, Jerold A. Edmondson and Zhou Guoyan, eds, 1996, *Languages and cultures of the Kam-Tai (Zhuang-Dong) group: a word list*. Bangkok: Sahadhammika Co. Ltd.
- Chafe, Wallace, 1976, Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Charles N. Li, ed. 1976:25–56.
- Dry, Helen, 1983, The movement of narrative time. *Journal of Literary Semantics* 12/2:19–53.
- Edmondson, Jerold and David Solnit, eds, 1997, *Comparative Kadai: the Tai branch*. Dallas: Summer Institute of Linguistics and the University of Texas at Arlington Publication in Linguistics, Publication 124.
- Givón, Talmy, ed., 1979, *Syntax and semantics 12: discourse and syntax*. New York: Academic Press.
1983. *Topic continuity in discourse: a quantitative cross-language study*. Amsterdam: Benjamins.
- Li, Charles N., ed., 1976, *Subject and topic*. New York: Academic Press.
- Li, Charles N. and Sandra Thompson, eds, 1976, Subject and topic: a new typology of language. In Charles N. Li, ed. 1976:457–490.
- 1979, Third person pronouns and zero anaphora in Chinese discourse. In Talmy Givón, ed. 1979:311–335.

Longacre, Robert, 1996, *The grammar of discourse*. New York and London: Plenum Press.

Zhou Guoyan, 1996, An introduction to the Kam-Tai (Zhuang-Dong) group of language in China. In Somsong Burusphat, et al., eds 1996:1–65.



IV

Sociolinguistics

9 *The future of Nyah Kur*

SUWILAI PREMSRIRAT

1 Introduction

Nyah Kur is a Mon-Khmer language of the Austroasiatic language family and is spoken in central Thailand. It belongs to the same sub-branch as Mon which is spoken in Thailand and Myanmar by about two million people. Nyah Kur and Modern Mon are too different to be dialects of the same language, yet, Nyah Kur is very similar to the Old Mon language which is found in inscriptions from the Dvaravati period. Diffloth (1984) in his book on Dvaravati, Old Mon, and Nyah Kur has proposed that Nyah Kur is an offshoot of Dvaravati Mon spoken about 2000 years ago. Payau (1979), Theraphan (1984), Subhab (1986) and Sudsawad (1990) have described various dialects of Nyah Kur spoken by the best speakers from various Nyah Kur villages in Chaiyaphum and Petchabun. All these dialects present typical Mon-Khmer characteristics.

A few years ago I happened to meet a Nyah Kur woman about twenty years of age who was working in a cafeteria at Mahidol University. She told me that her parents were Nyah Kur and that she spoke *nakun* instead of *nahkur*. This phonetic variation struck me as interesting and inspired me to investigate the current language situation in the Nyah Kur villages.

According to the *Ethnolinguistic map of Thailand* (Suwilai et al. 1998) the Nyah Kur people live in twenty-three villages in three provinces of Thailand: Nakhorn Rachasima, Chaiyaphum and Petchabun (see Figure 1). The population is between 4,000 and 6,000 speakers. However, all villages contain a mix of ethnic Thai Korat, Lao and others, as well as Nyah Kur. The degree of mixing between the Nyah Kur and other ethnic groups varies. The Nyah Kur population in Nakhorn Rachasima has almost died out with only a few Nyah Kur speakers in each of the three villages. They believe that after their generation Nyah Kur will certainly disappear from that area. In Petchabun the situation is a little better except for Ban Thaduang where there are more speakers. However, the Nyah Kur language has undergone some rapid changes. Although the linguistic situation in the Nyah Kur villages of Chaiyaphum is better, the language in each village is at a different stage of decline.

This paper describes the current Nyah Kur language situation and ventures a prediction of whether or not Nyah Kur will survive this age of mass communication. The investigation was carried out in 1997 and 1998 in four Nyah Kur villages in Chaiyaphum where the majority

of the population are Nyah Kur speakers. The purpose was to find out whether there is a chance of survival or any possibility of reversing the situation. This paper deals with two main aspects of the language situation:

- (1) Phonetic changes in Nyah Kur final consonants *r*, *l*, *c* and *ɲ* in the speech of speakers from three age groups; and
- (2) Attitudes toward the Nyah Kur language and community and Nyah Kur language usage and ability.

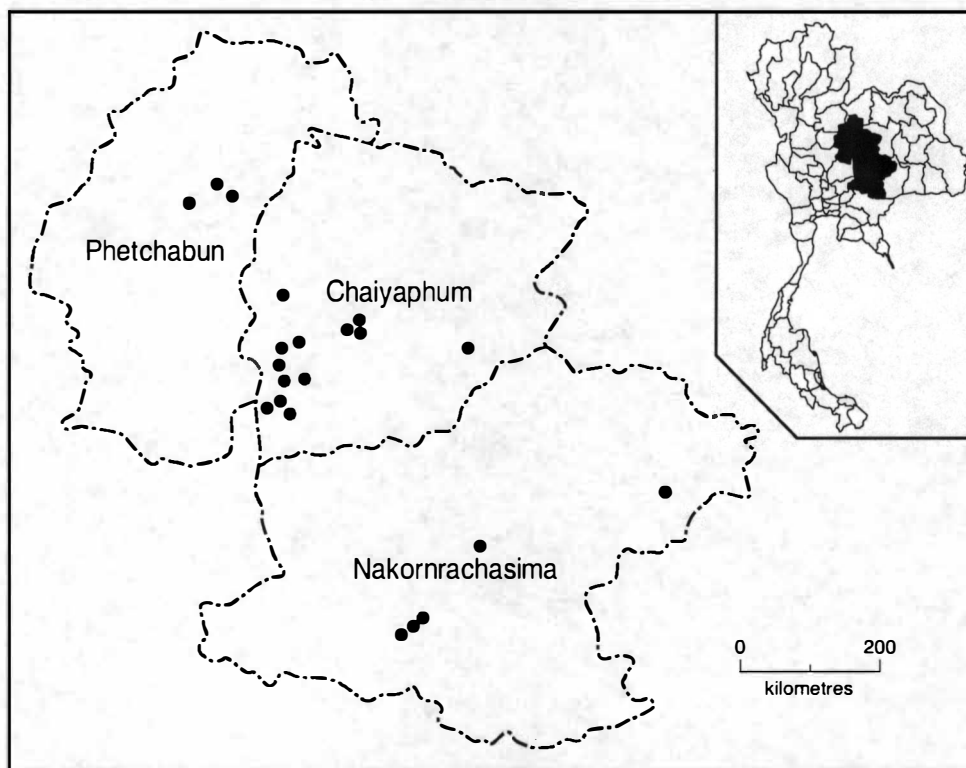


Figure 1: Nyah Kur settlements in Thailand

2 Phonetic changes in Nyah Kur final consonants

Final consonants *r*, *l*, *c* and *ɲ* are distinctive in Mon-Khmer languages. To investigate these final consonants the author examined the speech of three Nyah Kur age groups: (1) over fifty years of age, (2) between twenty-one and fifty, and (3) under twenty-one years of age. Five speakers from each age group in each of the four villages were investigated. In general this study revealed that younger Nyah Kur speakers appear to be strongly influenced by the Thai phonological system.

2.1 Final *r*

The investigation of the use of final *r* shows that the change starts with Nyah Kur speakers who are aged between twenty-one and fifty years. Some speakers use the flap *r* instead of trill *r*, and lateral *l* is used after long vowels by some people. However, all Nyah

Kur speakers under 21 use either *n* or *j* (but not both in variation) instead of final *r*, or totally omit the final *r* (this omission is represented as 0 ‘zero’ in Table 1 below).

Table 1: Phonetic variants of final *-r* in Nyah Kur for three age groups

English gloss:	Ages:	Over 50	21–50	Under 21
‘dog’		<i>chûr</i>	<i>chûr/r</i>	<i>chûn/j/0</i>
‘pumpkin’		<i>mpî:r</i>	<i>mpî:r/r/l</i>	<i>mpî:n/j/0</i>
‘ear’		<i>katûær</i>	<i>katûær/r/l</i>	<i>katûæn/j/0</i>
‘salt’		<i>paʔû:r</i>	<i>paʔû:r/r/l</i>	<i>paʔû:n/j/0</i>
‘to roll’		<i>tamâr</i>	<i>tamâr/r</i>	<i>tamân/j/0</i>
‘short’		<i>chêr</i>	<i>chêr/r</i>	<i>chê:n/j/0</i>
‘mat’		<i>nhâ:r</i>	<i>nhâ:r/r</i>	<i>nhâ:n/j/0</i>
‘to fly’		<i>phâr</i>	<i>phâr/r</i>	<i>phâr/j/0</i>
‘dove’		<i>pûr</i>	<i>pûr/r</i>	<i>pûr/j/0</i>
‘to crawl’		<i>mama:r</i>	<i>mama:r/r/l</i>	<i>mama:n/j/0</i>
‘tomorrow’		<i>ŋu:r</i>	<i>ŋu:r/l/l</i>	<i>ŋu:n/j/0</i>
‘to carry by fingers’		<i>jəjə:r</i>	<i>jəjə:r/r/l</i>	<i>jəjə:n/j/0</i>
‘to turn upside down’		<i>kaŋâ:r</i>	<i>kaŋâ:r/r/l</i>	<i>kaŋâ:n/j/0</i>
‘mountain’		<i>kûr</i>	<i>kûr/r</i>	<i>kûn/j/0</i>
‘lime’		<i>mô:r</i>	<i>mô:r/r/l</i>	<i>mô:n/j/0</i>

2.2 Final *l*

The findings show that Nyah Kur speakers under twenty-one years of age use final alveolar nasal *n* instead of lateral *l*, whereas Nyah Kur speakers of other age groups still use final *l*.

Table 2: Phonetic variants of final *-l* in Nyah Kur for three age groups

English gloss:	Ages:	Over 50	21–50	Under 21
‘lazy’		<i>ncʰî:l</i>	<i>ncʰî:l</i>	<i>ncʰî:n</i>
‘cotton’		<i>tûal</i>	<i>tûal</i>	<i>tûan</i>
‘to smooth’		<i>châ:l</i>	<i>châ:l</i>	<i>châ:n</i>
‘shallow’		<i>kadâ:l</i>	<i>kadâ:l</i>	<i>kadâ:n</i>
‘to plant’		<i>tâl</i>	<i>tâl</i>	<i>tân</i>
‘needle’		<i>nchul</i>	<i>nchul</i>	<i>nchun</i>
‘gourd’		<i>lul</i>	<i>lul</i>	<i>lun</i>
‘to give’		<i>kûl</i>	<i>kûl</i>	<i>kûn</i>
‘mortar’		<i>nû:l</i>	<i>nû:l</i>	<i>nû:n</i>
‘olive’		<i>kaʔî:l</i>	<i>kaʔî:l</i>	<i>kaʔî:n</i>
‘howl’		<i>khâl</i>	<i>khâl</i>	<i>khân</i>
‘leg’		<i>thɬ</i>	<i>thɬ</i>	<i>thɬn</i>
‘stump’		<i>takhâl</i>	<i>takhâl</i>	<i>takhân</i>
‘gibbon’		<i>juljûl</i>	<i>juljûl</i>	<i>juljûn</i>
‘heel’		<i>nə:l</i>	<i>nə:l</i>	<i>nə:n</i>

2.3 Final *c*

Findings from this study show that some Nyah Kur speakers under twenty-one years of age use final alveolar stop *t* and some use final palatal stop *c* (but not both in variation), whereas older speakers still use final *c*.

Table 3: Phonetic variants of final -*c* in Nyah Kur for three age groups

English gloss:	Ages:	Over 50	21–50	Under 21
'to fear'		<i>phlî:c</i>	<i>phlî:c</i>	<i>phlî:c/t</i>
'torn'		<i>tíc</i>	<i>tíc</i>	<i>tíc/t</i>
'small'		<i>nê:c</i>	<i>nê:c</i>	<i>nê:c/t</i>
'pretty'		<i>mê:c</i>	<i>mê:c</i>	<i>mê:c/t</i>
'cloth'		<i>ne:c</i>	<i>ne:c</i>	<i>ne:c/t</i>
'to tear'		<i>chê:c</i>	<i>chê:c</i>	<i>chê:c/t</i>
'sweet'		<i>tadâ:c</i>	<i>tadâ:c</i>	<i>tadâ:c/t</i>
'pig'		<i>khî:c</i>	<i>khî:c</i>	<i>khî:c/t</i>
'to swallow'		<i>ɲu:c</i>	<i>ɲu:c</i>	<i>ɲu:c/t</i>
'bald'		<i>talâ:c</i>	<i>talâ:c</i>	<i>talâ:c/t</i>
'to steal'		<i>khác</i>	<i>khác</i>	<i>khác/t</i>
'to cut'		<i>kéc</i>	<i>kéc</i>	<i>kéc/t</i>
'to dive'		<i>múc</i>	<i>múc</i>	<i>múc/t</i>
'a kind of vegetable'		<i>phrû:c</i>	<i>phrû:c</i>	<i>phrû:c/t</i>
'chisel'		<i>phənʔû:c</i>	<i>phənʔû:c</i>	<i>phənʔû:c/t</i>

2.4 Final *ɲ*

The findings show that some Nyah Kur speakers over twenty-one years of age still use the final palatal nasal *ɲ*, and some use *n*, whereas some Nyah Kur speakers under twenty-one use alveolar nasal *n* and some use the palatal nasal *ɲ* (but one speaker does not use both nasals in variation).

Table 4: Phonetic variants of final -*ɲ* in Nyah Kur for three age groups

English gloss:	Ages:	Over 50	21–50	Under 21
'ginger'		<i>pawî:ɲ</i>	<i>pawî:ɲ</i>	<i>pawî:ɲ/n</i>
'to swim'		<i>bî:ɲ</i>	<i>bî:ɲ</i>	<i>bî:ɲ/n</i>
'long'		<i>khlî:ɲ</i>	<i>khlî:ɲ</i>	<i>khlî:ɲ/n</i>
'grandfather'		<i>pê:ɲ</i>	<i>pê:ɲ</i>	<i>pê:ɲ/n</i>
'to lick'		<i>khla:ɲ</i>	<i>khla:ɲ</i>	<i>khla:ɲ/n</i>
'to shoot'		<i>pêɲ</i>	<i>pêɲ</i>	<i>pêɲ/n</i>
'to buy'		<i>ra:ɲ</i>	<i>ra:ɲ</i>	<i>ra:ɲ/n</i>
'yellow'		<i>pacû:ɲ</i>	<i>pacû:ɲ</i>	<i>pacû:ɲ/n</i>
'tiger'		<i>mpê:ɲ</i>	<i>mpê:ɲ</i>	<i>mpê:ɲ/n</i>
'red'		<i>phê:ɲ</i>	<i>phê:ɲ</i>	<i>phê:ɲ/n</i>
'elephant'		<i>cî:ɲ</i>	<i>cî:ɲ</i>	<i>cî:ɲ/n</i>
'fishing net'		<i>nêɲ</i>	<i>nêɲ</i>	<i>nêɲ/n</i>
'kite'		<i>khlîɲ</i>	<i>khlîɲ</i>	<i>khlîɲ/n</i>

According to oral histories related by elderly Nyah Kur speakers, the Nyah Kur were hunters and gatherers. They lived in a mountainous area that was malarial. About twenty years ago a road was built to this area, and the big trees were cut down by Thai businessmen from other areas. With the connection of electricity came television, and a primary school was opened. The Nyah Kur were then exposed to the outside world. More and more people from the outside, mainly Thai Korat and Lao speakers, started to settle in the same area as the Nyah Kur. The timing in this part of the Nyah Kur history corresponds well to what we have found in the speech of the Nyah Kur speakers. The younger generation of Nyah Kur speakers, especially those under twenty-one years, demonstrate clearly the change in the final consonants towards the Thai phonological system. The typical Mon-Khmer finals *r*, *l*, and *ŋ* have been changed to *n* and also to *j* in case of final *r*. Final *c* is changed to *t*. Alongside these changes in the speech of the younger generation are some changes in initial consonants, such as *ŋ* changing to *n* and deletion of final *h* and *ç*. For example, the word *ŋah kûr* 'people-mountain' which refers to their own language and ethnic group is pronounced *nakûn*.

It is also quite obvious that besides the changes towards the Thai phonological system, younger Nyah Kur speakers have lost about half the vocabulary which is still known to the older generation. Therefore, it is very clear that Nyah Kur is gradually changing to become more like Thai. However, it is also obvious that the younger speakers of Nyah Kur still keep some aspects of the general phonological system, such as voice quality, that is, the distinction between lax and tense voice. It is likely that in the future Nyah Kur will lose all Mon-Khmer final consonants and adopt the Thai system of finals and will also probably lose its original vocabulary. Nonetheless, it will keep the general phonological characteristics, so that in the distant future, when reflexification process is completed, this language will become a language that looks very much like Thai but has a special accent.

3 Language attitudes, language use and language ability

This section focuses on the Nyah Kur speakers' attitudes towards their language and ethnic groups, including Nyah Kur language use and ability. The investigation was conducted by the author in March 1999, in four villages — Wang Ajpho, Saliengthong, Wang Ajkhong and Namlat, Chaiyaphum provinces — by the use of questionnaires. The questionnaires are adapted from Benjamas Khamsakul (1998) and Wimon Thanapaes (1998). The research subjects were 120 Nyah Kur speakers of various backgrounds from four villages where Nyah Kur speakers are in the majority. The study also recorded each subject's age, sex, education, occupation, duration of time working outside the village, the ethnicity of parents and spouse, and so on.

3.1 Language attitudes

The questionnaire for investigating language attitudes is based on the theory of Ethnolinguistic Vitality Beliefs proposed by Allard and Landry (1986, 1992), Giles, Bourhis and Taylor (1977), and Kreidler and Kreidler (1972, 1976); this theory states that by determining the beliefs of a group regarding their language and other languages in the same area it is possible to predict a broad range of ethnolinguistic behaviours, including language attitude and language choice.

The questionnaire consisted of twenty-three questions designed to investigate the Nyah Kur attitudes towards their own language and ethnic group. The questionnaire examined two

types of beliefs: Non-self beliefs and Self-beliefs. The Non-self beliefs section consisted of eight questions that involved the feelings of the native speakers of Nyah Kur toward the things that influence the vitality of the Nyah Kur group, such as demographic, economic, and cultural factors. The Self-beliefs section consisted of fourteen questions that focus on the feelings of the native Nyah Kur speakers about what is important in their minds as to the vitality of the group and what they think is appropriate for their group. Respondents were asked to indicate their degree of agreement or disagreement with particular statements. The questions and responses according to age groups are presented below. The scoring of the answers to these questions was as follows: strongly agree 5; agree 4; neutral 3; disagree 2; strongly disagree 1.

3.1.1 *Nonsel-belief*

As indicated in Table 5, the responses show that Nyah Kur speakers over fifty years old have a neutral attitude (23.70) whereas the Nyah Kur speakers between twenty-one and fifty years of age (26.00) and under twenty-one years of age (29.77) have a positive attitude in regard to these statements.

Table 5: Nonsel-belief about Nyah Kur language and Nyah Kur ethnic group

Statements (5 points each):	Ages:	Responses:		
		Over 50	21–50	Under 21
1.1 The Nyah Kur in this village are the majority.		3.58	4.27	4.21
1.2 The Nyah Kur in this village are wealthy.		2.36	2.35	2.04
1.3 Most leaders in this village are Nyah Kur people.		4.07	4.36	4.22
1.4 In 50 years there will still be Nyah Kur people in this area.		3.93	3.67	2.01
1.5 Most Nyah Kur people prefer to use Nyah Kur with Nyah Kur friends.		4.39	4.74	4.78
1.6 The school in this village has influence in promoting the Nyah Kur language among Nyah Kur students.		2.88	1.80	1.61
1.7 The government (local) has influence in supporting Nyah Kur culture.		3.68	2.63	2.61
1.8 The school here teaches Nyah Kur students to be proud of their ethnicgroup.		3.61	2.48	2.15
		23.70	26.00	29.77

3.1.2 *Self-belief*

As indicated in Table 6, the responses show that the Nyah Kur of all age groups are quite positive about their ethnic language.

Table 6: Self-belief about Nyah Kur language and Nyah Kur people

Statements (5 points each):	Responses:			
	Ages:	Over 50	21–50	Under 21
2.1 If I could choose to be reborn, I would choose to be born Nyah Kur.		3.71	3.70	3.71
2.2 It is better to use Nyah Kur with children than other languages.		4.35	4.02	4.35
2.3 We should only use Nyah Kur in our family for its preservation.		3.77	3.63	3.77
2.4 It is not shameful that Nyah Kur does not have a writing system.		3.77	3.63	3.77
2.5 The government should support the culture of Nyah Kur		4.55	4.51	4.55
2.6 It would be better if there were only Nyah Kur people in this village.		4.55	4.26	4.55
2.7 Reducing Nyah Kur language use will have the effect of reducing the Nyah Kur identity.		4.46	3.05	4.46
2.8 It is shameful to be a part of the Nyah Kur group.		2.96	3.05	2.69
2.9 I feel more confident to tell people that I am Lao or Thai Isan than to tell them that I am Nyah Kur.		3.27	3.17	3.27
2.10 Knowing Nyah Kur language can help me get a good job.		2.80	1.92	1.98
2.11 It would be better to marry within the Nyah Kur group than with others.		4.23	3.58	4.23
2.12 This village is a place where Nyah Kur people will live forever.		4.64	4.46	4.64
2.13 Nyah Kur should have a writing system as do other languages.		4.49	4.67	4.49
2.14 If there is an election for leaders in this village, the Nyah Kur will elect Nyah Kur people.		4.76	4.68	4.76
		62.76	59.40	59.43

3.2 Language use

The use of the Nyah Kur language in various domains was studied to find out how often the Nyah Kur people use their own language. It is interesting to see the use of the Nyah Kur language in intra-group contact and inter-group situations. The intra-group situations include the nuclear family domain, extended family domain or relative domain, and neighbours in the same village domain. The inter-group situation includes the neighbours outside the village, in school, in trading and in government offices such as district offices. The nuclear family domain (24) includes father, mother, spouse, sibling, children over twenty-one and under twenty-one years of age. The extended family or relative domains (12) include older

relatives, younger relatives, and their children. The neighbours in the same village (12) are considered in terms of older neighbours, younger neighbours, and the children. The neighbours outside the village (12) are also considered in terms of older neighbours, younger neighbours and the children. The school domain (12) includes friends in school, friends out of school and teachers. Other domains are speaking with merchants (8) who come to sell things in the village and government officials (8) such as district officers, health personnel, and so on. The figures below show the use of Nyah Kur in different domains according to the age variable.

The responses show that the Nyah Kur still mostly use their ethnic language in their homes and among their family members. They also use Nyah Kur with relatives and friends in the village, but less with friends in other Nyah Kur villages. However, at school and in the local government office Nyah Kur is used very little. Speakers aged over twenty-one also speak Nyah Kur more than those under twenty-one years of age.

Table 7: Language use in seven different social domains of three age groups

Social Domains:	Ages:	Over 50	21–50	Under 21
Family domain (24)		17.79	14.51	9.94
Relatives domain (12)		9.52	9.85	8.32
Neighbour domain				
a. in the village (12)		9.51	9.65	7.91
b. out of the village (12)		7.75	7.07	2.88
School (12)		1.83	3.98	3.36
Trade (8)		3.04	8.47	3.12
Government officials (8)		0.08	0.42	0.06

3.3 Language ability

The questionnaire section on language ability of the Nyah Kur consists of two parts which both involve self-evaluation and can give us some insights into the whole situation of this preliterate or semi-literate society.

The first part of the questionnaire asked about ability to speak and understand the languages they come across in their daily life. These are Nyah Kur, Lao, Thai Korat and Standard Thai. The Nyah Kur speakers were asked to self-evaluate their ability to understand and speak each language: no understanding, a little, enough, good, and very good. The percentages for the ability at the 'very good' level according to the Nyah Kur of different age groups are presented below.

The responses show that the Nyah Kur speakers understand and speak Nyah Kur more than other languages found in their villages. The speakers over fifty years can understand and speak Nyah Kur the most (98%), whereas speakers 21–50 years of ages can understand and speak more (79.63%) than those under twenty-one years of age (42.4%). It is noticeable that the Nyah Kur can understand Lao, Thai Korat and Standard Thai more than they can speak them. Nyah Kur speakers under twenty-one years of age can understand and speak Standard Thai more than older people.

Table 8: Ability to understand and speak Nyah Kur, Lao, Thai Korat and Standard Thai for three age groups

Knowledge of languages:	Ages:	Ability:		
		Over 50	21–50	Under 21
1. Understanding Nyah Kur		98	79.63	42.4
2. Speaking Nyah Kur		98	79.1	51.5
3. Understanding Lao		21.95	16	6.1
4. Speaking Lao		15.5	11	6.1
5. Understanding Thai Korat		28.2	31.7	5.4
6. Speaking Thai Korat		10.6	20.03	5.9
7. Understanding Standard Thai		30.1	26.23	36.4
8. Speaking Standard Thai		22.75	22.53	33.3

The second part of the questionnaire on language ability of Nyah Kur speakers of different age groups asked them to self-test whether a Nyah Kur can speak in Nyah Kur on certain topics. The Nyah Kur speakers were asked to answer ‘yes’ or ‘no’. ‘Yes’ indicated a greater degree of ability. ‘No’ indicated a lesser degree of ability (although some ‘No’ answers may indicate a greater degree of ability).

The responses in Table 9 show that the Nyah Kur people still have a good ability in speaking the Nyah Kur language: 80% can speak in Nyah Kur on various topics.

Table 9: Self-test of speaking in Nyah Kur on various topics

Language ability:	Responses:	
	Yes	No
1. Can you say how to go from this village to the district office?	90.8	10.2
2. Can you ask and tell the time in Nyah Kur?	92.8	8.2
3. Can you tell the history of the settlement of your family in Nyah Kur?	90.2	10.8
4. Can you explain how to grow rice in Nyah Kur?	92.2	8.8
5. Can you give the details of the members in your family in Nyah Kur?	98.9	2.1
6. Can you tell someone to do something and understand when someone asks you to do something in Nyah Kur?	96.3	4.7
7. Can you explain about climate and vegetation in Nyah Kur?	93.22	6.8
8. Are there times when you cannot remember Nyah Kur words?	49.6	51.4
9. If Thai speakers talk with you about politics, can you talk about this?	73.6	27.4
10. If the doctor explains to you about disease and how to protect yourself from disease will you be able to understand?	92.5	7.5
11. Can you sing a song or tell jokes in Nyah Kur?	58.1	42.9
12. If you quarrel with a Nyah Kur, can you abuse him in Nyah Kur?	94.1	6.9
13. Can you voice your opinion about various things in Nyah Kur?	94.9	6.1
14. Do you often make mistakes speaking Nyah Kur?	62.4	38.6
15. Can you speak Nyah Kur like a native speaker?	87.5	13.5
16. Do people know that you are not ethnic Thai when you speak Thai?	32.1	67.9
	80	20

4 Conclusion

In the present era of globalisation, the mass media — especially television — are very powerful. TV can reach right into the homes of the villagers. The languages of wider communication are pushing out the smaller minority languages. Nyah Kur cannot escape this situation. The younger generation speaks Nyah Kur less than the older people. In the mixed family where the spouse is not a Nyah Kur speaker it is very likely that the children do not speak Nyah Kur. This type of family is becoming more and more common in Nyah Kur villages. However, this research investigated only the pure Nyah Kur race in four Nyah Kur villages in order to see what the situation is like and to find out whether there is any chance to maintain this language in the future.

A sample of 120 Nyah Kur speakers of different age groups were interviewed about their use of final consonants *r*, *l*, *c* and *ɲ* which are the typical final consonants of the Mon-Khmer language family, in order to see how much Nyah Kur speakers can still keep the characteristics of their ethnic language. I have found that the language of the young generation increasingly favours the Thai phonological system. This is actually a normal development and has happened to many minority languages in Thailand. However, it is also noticeable that the vocabulary of the younger generation has also decreased to about half that of the older generation. This shows that the Nyah Kur are going to gradually lose their language in the future if nothing is done to maintain it. This is also a natural process.

The results of the questionnaire survey of attitudes toward the Nyah Kur language and identity are still very positive. The Nyah Kur people have Nyah Kur people as their leaders and will elect Nyah Kur people to be their leaders. Most of them use their ethnic language at home with their relatives and neighbours in the villages. They think that the Nyah Kur language and culture give them their identity, and they want the government to promote their culture and identity. They also want to have an orthography for their language. This desire is very interesting as it shows that if the government officials and other authorities support efforts to encourage and help the Nyah Kur people, then the language should be able to survive in the 21st century.

References

- Allard, R. and R. Landry, 1986, Subjective ethnolinguistic vitality viewed as a belief system. *Journal of Multilingual and Multicultural Development* 7 (supplement 1), 1-12.
- 1992, Ethnolinguistic vitality beliefs and language maintenance and loss. In W. Fase, K. Jaspaert and S. Koon, eds 1992:171-195.
- Benjamas Khamsakul, 1998, The attitude of So (Thavung) ethnic group toward their own language and its use: case study of Nong Waeng village, Pathumwapi sub-district, Songdao district, Sakon Nakorn province. MA thesis, Institute of Language and Culture for Rural Development, Mahidol University.
- Diffloth, Gérard, 1984, *The Dvaravati Old Mon language and Nyah Kur*. Bangkok: Monic Language Printing House Studies, Chulalongkorn University.
- Fase, W., K. Jaspaert and S. Koon, eds, 1992, *Maintenance and loss of minority languages*. Philadelphia: John Benjamins Publishing Company.
- Giles, H., ed., 1977, *Language, ethnicity and intergroup relations*. London: Academic Press.
- Giles, H., R.Y. Bourhis and D.M. Taylor, 1977, Toward a theory of language in ethnic group relation. In H. Giles, ed. 1977:307-348.

- Kreitler, H. and S. Kreitler, 1972, The model of cognitive orientation: towards a theory of human behaviour. *British Journal of Psychology* 63:9–30.
- 1976, *Cognitive orientation and behaviour*. New York: Springer.
- Maher, B. and W. Maher, eds, 1982, *Progress in experimental personality research*, Vol.II. New York: Academic Press.
- Payau Memans, 1979, A description of Chaobon: an Austroasiatic language in Thailand. MA thesis, Institute of Language and Culture for Rural Development, Mahidol University.
- Subhab Phiukhou, 1986, A phonological description of Nyah Kur at Ban Nam Lat, Chaiyaphum province. MA thesis, Institute of Language and Culture for Rural Development, Mahidol University.
- Sudsawad Chuasuwan, 1990, A phonology of Nyah Kur at Ban Tha Duang, Petchabun province. MA thesis, Institute of Language and Culture for Rural Development, Mahidol University.
- Suwilai Premsrirat, 1998, On language maintenance and language shift in minority languages of Thailand: a case study of So (Thavung). In Kazuto Matsumura, ed. *Studies in endangered languages: papers from the International Symposium on Endangered Languages, Tokyo, November 18-20, 1995*. 149–175. Tokyo: Hitzui Syoho.
- Suwilai Premsrirat, et al., 1998, *Ethnolinguistic map of Thailand*. Institute of Language and Culture for Rural Development, Mahidol University.
- Theraphan L.Thongkum, 1984, *Nyah Kur (Chaobon)-Thai-English dictionary*. Bangkok: Chulalongkorn University Press.
- Wimon Thanapaes, 1998, Language use of the Isan group in Palan village, Tambol Pong-noi, Mae-Chan district, Chiangrai province. MA thesis, Institute of Language and Culture for Rural Development, Mahidol University.

V

Historical linguistics

10 *A comment on Gedney's proposal for another series of voiced initials in Proto Tai*

GRAHAM THURGOOD

1 Introduction

In a paper entitled 'Evidence for another series of voiced initials in proto-Tai', Gedney (1989 [1979]) posited a second series of voiced initials for Proto Tai (PT) to account for roughly fifty-one forms with a pattern of initial and tonal correspondences otherwise unique within Tai. Although the anomalous nature of these forms must also have been apparent earlier to Li, in Li's *Handbook* (1977) the majority of these forms are simply reconstructed with voiceless aspirated stops for Southern Tai (that is, the Southwestern (SWT) and Central Tai (CT)), with Li frequently appending a note that, in Northern Tai (NT), the forms are reconstructed, not with a voiceless onset, but with a voiced onset.

Table 1 illustrates the patterns noticed both by Li and by Gedney. The first column lists the five tone classes reconstructed for Proto Tai (PT), which are indicated by A, B, C, DS, and DL. The final syllable in tone classes A, B, and C ends either in a vowel or a sonorant, constituting what Tai scholars term 'live' syllables. The final syllables in tone classes DS and DL end in a stop, constituting what Tai scholars term 'dead' syllables, with the DS occurring with short vowels and the DL with long vowels.

The primary focus of interest in Table 1 is the pattern of tonal variation for **B* forms. In the Southern Tai dialects, the tones of these forms pattern as if they had had PT voiceless obstruent initials; in the Northern Tai dialects, these forms pattern as if they had had PT voiced obstruent initials. Thus, in Siamese, a dialect of Southwestern Tai, and in Lungming, a dialect of Central Tai, tones of the **B* forms pattern tonally with the voiceless aspirated **ph*-series, while in Saek and Yay, both Northern Tai dialects, the **B* forms pattern tonally with the **b*-series. This tonal alternation was obvious to Li, who reconstructed the majority of these forms with voiceless aspirated stops in his *Handbook* (1977), but with an appended note that indicated that in NT these forms were reconstructed with voiced obstruents.

Table 1: Gedney forms: the tones and the initials

PTai tone	PTai Initial	Siamese (SWT)	LM (CT)	Saek (NT)	Yay (NT)
*A	* <i>ph-</i>	<i>ph- 5</i>	<i>ph- 1</i>	<i>ph- 2</i>	<i>p- 1</i>
	* <i>B-</i>	<i>ph- 5</i>	<i>ph- 1</i>	<i>ph- 4</i>	<i>p- 4</i>
	* <i>b-</i>	<i>ph- 1</i>	<i>p- 4</i>	<i>ph- 4</i>	<i>p- 4</i>
*B	* <i>ph-</i>	<i>ph- 2</i>	<i>ph- 2</i>	<i>ph- 6</i>	<i>p- 2</i>
	* <i>B-</i>	<i>ph- 2</i>	<i>ph- 2</i>	<i>ph- 5</i>	<i>p- 5</i>
	* <i>b-</i>	<i>ph- 3</i>	<i>p- 5</i>	<i>ph- 5</i>	<i>p- 5</i>
*C	* <i>ph-</i>	<i>ph- 3</i>	<i>ph- 3</i>	<i>ph- 3</i>	<i>p- 3</i>
	* <i>B-</i>	<i>ph- 3</i>	<i>ph- 3</i>	<i>ph- 6</i>	<i>p- 6</i>
	* <i>b-</i>	<i>ph- 4</i>	<i>p- 6</i>	<i>ph- 6</i>	<i>p- 6</i>
*DS	* <i>ph-</i>	<i>ph- 2</i>	<i>ph- 3</i>	<i>ph- 4</i>	<i>p- 2</i>
	* <i>B-</i>	<i>ph- 2</i>	<i>ph- 3</i>	<i>ph- 6</i>	<i>p- 1</i>
	* <i>b-</i>	<i>ph- 4</i>	<i>p- 4</i>	<i>ph- 6</i>	<i>p- 1</i>
*DL	* <i>ph-</i>	<i>ph- 2</i>	<i>ph- 2</i>	<i>ph- 6</i>	<i>p- 2</i>
	* <i>B-</i>				
	* <i>b-</i>	<i>ph- 3</i>	<i>p- 5</i>	<i>ph- 5</i>	<i>p- 5</i>

It remained for Gedney (1989a), however, to fully recognise the problem that these forms presented. He carefully documented the evidence, noting as Li had that the tones of the Southern Tai reflexes suggest earlier voiceless obstruents, while the initial and tonal patterns of the Northern Tai reflexes suggest earlier voiced obstruents. Then, he addressed the problem, which is, of course, that PT already has, for instance, a well-documented series of voiceless unaspirated stops, another series of voiceless aspirated stops, a series of voiced stops, and so on. If the assumption is made that the new series is inherited, the problem is how to account for the additional correspondence patterns. Gedney (1989a) tentatively suggested accounting for this new correspondence pattern by positing another set of PT voiced initials. As PT already has a rather rich array of obstruents, the phonetics of this new series are a problem, as Gedney acknowledges by using the phonetically indeterminate symbols **B*, **G*, **D*, **Z*, and so on to represent the various members.

In place of Gedney's internal account, this paper proposes that the fifty-one forms in question were borrowed. The evidence is complex but compelling. Within the Tai dialects, there is internal evidence that these forms are borrowed. Beyond Tai, there is evidence that many of these same forms occur in other language subgroups; in those related to Tai, the evidence often suggests that the corresponding forms are borrowed. At least two-thirds of the forms have counterparts outside of Tai-Kadai and in many cases the counterpart is not just present but reconstructs in the other language family; this is evidence that the corresponding forms are native to those language groups. Further, the fact that the

preponderance of forms occurs outside of the Tai-Kadai family is itself an indication that the Tai forms are borrowed into Tai.

2 The data

It is important to note that no attempt has been made to be definitive. There is no doubt whatsoever that, for example, more Tibeto-Burman counterparts could be discovered with additional work. Similarly, it is likely that more Chinese counterparts could be discovered by searching with greater care. The paper would certainly have been improved in terms of clarity had the Karlgren forms been systematically replaced by the Baxter forms. Nonetheless, the evidence is overwhelming.

The data for this paper come from a multitude of sources. For the Tai languages, the base comes from Gedney (1989). Anyone who has read Gedney's work is familiar with the care and meticulousness that he assembles his data. That material is supplemented by other works by Gedney, specifically his dictionaries of Lungming (1991a), Yay (1991b), and Saek (1993). The reconstructions of PTai follow Strecker (1983), who first suggested combining Li's (1977) initials with Sarawit's (1973) vowels and then critiqued both analyses. For the Tai-Kadai, the reconstructions of the various subgroups are examined: for Proto Be, the basic sources are Hansell (1988) and my own notes; for Proto Kam-Sui (PKS), the basic source is Thurgood (1988), a work that incorporates Li (1965), supplemented by other forms that needed to be examined specifically for this paper; for the Proto Hlai, that is, the Li languages of Hainan, the basic source is Thurgood (1991), a work that owes a great deal to Matisoff (1988); for the Chinese component of Sino-Tibetan, the main source has been Karlgren (1923, 1957), but also examined were Baxter (1992) and Li (1976). And, for the Tibeto-Burman component of Sino-Tibetan, the major source was Benedict (1972), with various other sources used for subgroups: for Proto Lolo-Burmese (PLB) Burling (1967), Matisoff (1972), Bradley (1979), and Thurgood (1977) were used, for so-called Northern Naga French (1983) was used. For Proto Austronesian, Blust's files and reconstructions were used, sometimes supplemented by personal communication. For the Malayic subgroup, Adelaar (1992) was used and for Proto Chamic (PC), Thurgood (1999) was used. For the Mon-Khmer (MK) branch of Austroasiatic several reconstructions of subgroups were used: Proto South-Bahnaric (PSB: Blood 1967, 1968, 1974 [who reconstructs Proto Mnong, but uses a data base broad enough to make it equivalent to PSB most of the time]; Efimov 1987), Wa (Diffloth 1980), Proto North-Bahnaric (PNB: Smith 1972), and Proto Katuic (Thomas 1967; Peiros 1996).

Each of the sets has been assigned a number with the numbers in the text corresponding to the numbers in the tables.

3 The internal evidence

The most obvious internal evidence that these forms are borrowings is the very alternations in tones and initials to identify this particular set of forms. *Unless another series of obstruents can be reconstructed for PT*, these forms are irregular in PT: these forms do not fit with any of the series reconstructed thus far. That is, if they are reconstructed with voiceless initials, they are irregular in NT; if they are reconstructed with voiced initials, they are irregular in ST. It was because of this problem, of course, that Gedney proposed reconstructing a new series of obstruents for these forms in the first place.

At this point, it is imperative to emphasise that just the fact these fifty-one forms manifest a certain regularity in their correspondence patterns does not ensure that they are native forms. It is common for borrowings to exhibit regular correspondence patterns within the recipient language. To cite an obvious example (and ignoring other evidence that the more Latinate forms are borrowed), within Germanic the initials of the English borrowed forms *penta-*, *pater-*, *Pisces*, and *pecuniary* correspond as nicely as do the inherited *five*, *father*, *fish*, and *few*, at least upon first inspection. Thus, the existence of regular correspondence patterns within a subset of the vocabulary is interesting but by itself tells little about the ultimate origins of the forms, aside from indicating that, if the forms are borrowed, most of them must have been borrowed before the breakup of ST and NT.

Beyond the simple alternation in the tonal patterns that correlates with the split between ST and NT, there are numerous other pieces of evidence within Tai suggesting the forms are loans, some relatively weak, some much stronger. One weak indicator is the fact that the pattern involves only a small number of forms, some fifty-one or so. However, the small number by itself proves little. After all, the so-called third tone reconstructed for Lolo-Burmese is only manifested by some fifty or so forms and is an internal development within Tibeto-Burman.

More telling are the indications of irregularities within the examples themselves. Given that many of the forms manifesting these patterns are also found outside Tai and, for that matter, outside of the Tai-Kadai family, irregularities in the forms within Tai are what one might associate with borrowing, particularly late borrowing. It is simply a basic part of methodology to assume that, if a form occurs in more than one language family and corresponds regularly within one family but irregularly within the other, it was borrowed into the family in which it behaves irregularly. While in many cases it remains to be proven that these forms behave regularly outside of Tai-Kadai, it is evident that a number of these forms behave irregularly within Tai-Kadai.

Both Li and Gedney raised questions about the internal correspondences of a number of the forms. In Table 2, the forms within Tai itself show various irregularities, all of them serious enough to disqualify the form as an example of the pattern. (39) The form **Goin* A/B -t, in addition to minor problems, has a basic tonal discrepancy; it is tone A in Southern Tai, but tone B in NT. (40) The form **Fia* A, with the range of meanings including 'braid; harrow; part of loom' apparently has irregular initials, tones, and vowels in its various manifestations. The picture is complicated by both its unclear semantics and the fact that it is not widely attested. Gedney (1989a:247) suggested this etymon may actually be two etyma which have fallen together in some dialects. (41) The form **Giaw* B is too weakly attested to be considered a serious exemplar of the pattern. (42) On semantic grounds, the form **Zii* A is a likely borrowing but it is so marginally attested that, in any case, it is a very weak exemplar of the pattern. (43) The form **ZaY* A -it shows irregularities in the vowel length of its Siamese form (signified by -l), and in its initials (signified by -i), in its tones (signified by -t). Further, as Li (1977:151) noted, the 'NT forms seem to go back to PT **gw-* and may not be directly related'. (44) The form for 'line; row' has an irregular tone in Saek, as Gedney notes, but is otherwise okay. (45) The form **Zuam* C is not even attested in NT. It is only included by Gedney because it appears to have had an originally voiced initial in Kam-Sui. (46) In a similar way, the form **Piaa* B/C -t is tonally irregular, having tone B reflexes in ST but tone C reflexes in NT. (47) The form **Gooj* A is irregular in both its tones and its vowels. (48) Finally, the form **Zaaw* C -i for 'pole' has an irregular initial; the ST forms reconstruct with a **z-*, but the NT forms in part reflect PT **ʃ-*

and in part *z-. (49) The form *Gwaa A -t has, as Li (1977:238) noted, 'irregularity in the initial'; it also has, as indicated by my -t, irregularities in its tonal reflexes (The superscripted x indicates that, quite apart from this analysis, I had determined that the form was originally a loan into PT. (50) The form for 'ripe; cooked' has irregularities in its initial correspondences. (51) Aside from the word being marginally attested, the final of the form for 'pole; staff' is irregular.

Table 2: Internal problems within Tai

	PTai	Siamese (SWT)	Lungming (CT)	Saek (NT)	Yay (NT)	
39.	*Goin A/B -t	khɔɔn ⁵	—	—	kuan ⁵ -t	'log'
40.	*Fia A	fia ⁵	—	via ¹ -t / ⁴ ; phia ² -t; phwa ² -t	—	'to braid; loom, part of; to harrow'
41.	*Giaw B	—	—	—	—	'cross the arms'
42.	*Zii A	sii ⁵	—	—	si ⁴	'fifth earthly branch'
43.	*ZaY A -it	saay ⁵	saay ¹ -l	svɣ ⁴ 'time'	saay ⁶ -t 'afternoon'	'late; in the morning'
44.	*Deew A	thɛɛw ⁵	tweew ⁴	thɛɛw ² -t	—	'line; row'
45.	*Zuam C	suam ³	suum ³	—	—	'room; compartment'
46.	*Piaa B/C -t	phaa ³	—	phia ⁵	pua ⁵	'cloth, clothing'
47.	*Goon A	(khraŋ ¹)	—	reeŋ ¹ -t	koŋ ⁴ -v	'to moan'
48.	*Zaaw C -i	saw ³	saaw ³	yaaw ⁶ -i	θaaw ⁶ -i	'pole'
49.	x*Gwaa A -t	khwaa ⁵	saa ¹	khwaa ⁴	kwa ⁴	'right [hand]'
50.	*Zuk D -i	suk ²	sok ³	suk ⁶	suk ¹	'ripe; cooked'
51.	xDij C -v	—	—	—	—	'pole; staff'

These purely internal considerations eliminate roughly a quarter of the proposed sets from consideration. Remembering that there were only fifty-one or so manifestations of the pattern, this reduction in the number of attested sets is of concern. It is also instructive that most of these irregularities were noted by Li, by Gedney, or by both.

4 The wider Tai-Kadai evidence

An examination of counterparts in other subgroups of Tai-Kadai (Proto Be, Proto Kam-Sui, and Proto Hlai) shows that many of the corresponding forms are, within one or more of these subgroups, demonstrably borrowed, as is evident either from tonal patterns, initial patterns, or both.

The evaluation of wider Tai-Kadai evidence depends heavily on the reconstructions which have already been done of the various subgroups. Although much remains to be done before the reconstructions of any of these languages can be said to be fully understood, the basic relationships are now understood. For a large number of the Gedney forms, wider Tai-Kadai shows evidence that the forms are borrowings. (5) The form 'to rake' has a reconstruction in KS that makes it evident that it is a loan. (14) The form for 'cooked rice' has an irregular initial, final, and tone, making it a loan. (15) The form 'sharpen; to grind' has an irregular

tone and initial in PBe. (19) The form for ‘sugar’ is unreconstructable in PHLai, due to its irregularities. In Hlai, it has an irregular final and tone; in fact, the final is a final that is restricted to loans. (22) The form for ‘ten’ has an initial in PKS that is restricted to loans. (23) The form for ‘thick; dense’ has an irregular final and is not reconstructable in KS. (26) The form for ‘carry; hold’ occurs in KS, but the irregular vowel correspondences suggest that it was borrowed into KS. (27) For ‘arrive, reach’, aside from the vowel of the PBe form, the forms in the individual languages look fine, but any cross-language examination will run into serious problems, that is, the initials vary across languages in a way that it is unlikely to be reconstructable at a higher level. (28) The form for ‘bean’ is a borrowing in KS. (29) The form for ‘bowl; cup’ has an irregular final in PBe. (30) The second form for ‘bundle’ has an initial in KS that appears to be restricted to borrowed words. (31) The form for ‘chopsticks’ within both PBe and PHLai has as of yet no obvious internal problems, but the problem will come in attempting to fit these forms into a reconstruction of Proto Tai-Kadai. (32) The form for ‘eggplant’ is a borrowing into KS and Hlai. (33) The KS reflexes of the form **Duaŋ* B ‘weigh’, **daŋ⁵*, indicate a voiced onset while the tonal reflexes indicate a voiceless onset. Thus, the form was borrowed into KS. (34) The form for ‘bank (paddy)’ has variant tones in PBe, but otherwise has no problems. (35) For the form **Gut* D -t ‘dig, hollow out’, the PKS form is a borrowing and the final of the PHLai suggests it is a borrowing. Gedney speculated that this word would sometimes fall together with a distinct word meaning **khuut* D ‘scrape, grate’ and that two words are involved. Gedney is clearly right, with both words showing up in PHLai and in Sino-Tibetan as distinct etymon. (36) The Kam-Sui (KS) reflexes of the form **χəy* C ‘excrement’, **ke⁴* -ti, indicate a voiced initial but the tonal reflexes indicate a voiceless initial. (37) The KS reflexes of the form ***Giaay* B, **ki^{2/6}* -tif, indicate a voiceless initial but the tonal reflexes indicate a voiced initial. Thus, the word is borrowed. The form ‘ride on horseback’ is irregular throughout Tai-Kadai. (38) The word for ‘ear’ is unique, both in PTai and elsewhere. Thus, it has no value in establishing these patterns. In fact, its only potential value is as an exemplar of an already established pattern. That is, if the analysis is established, the analysis might be used to explicate ‘ear’, but not the other way around. (47) The form for ‘to moan’ is quite irregular in PHLai. (51) The form for ‘pole’ is very irregular in KS. (50) The form ‘ripe; cooked, done’ has a PKS initial that is only found with loanwords.

Table 3: Tai-Kadai counterparts

	PTai	PBe	PKS	PHlai	
2.	<i>*Gum</i> A	<i>*xɔm²</i>	—	—	‘pit; ditch’
3.	<i>*həəŋ</i> B	—	<i>*hŋla:ŋ⁵</i>	—	‘young [chicken]’
5.	<i>*Graay</i> B	—	<i>**khra:i⁵</i>	—	‘to rake’
8.	<i>**Gun</i> -t	—	—	<i>*kho:n³</i>	‘to bundle’
12.	<i>*Dək</i> D t?	<i>*hɔk⁸</i>	<i>*dak⁸</i>	—	‘male [animal]’
14.	<i>*χaw</i> C	—	<i>*hau⁴</i> -fit	—	‘cooked rice’
15.	<i>*Fan</i> A	<i>*vɔn^{1/3}</i> -it	<i>*gwan²</i>	<i>*ka:n²</i>	‘sharpen, to; grind’
19.	<i>*Diaŋ</i> A	<i>*l/han²</i>	<i>*da:ŋ²</i>	<i>*tha:ŋ³</i> -tf	‘sugar’
20.	<i>*χəm</i> A -t	<i>*kam²</i>	<i>*kam¹</i>	<i>*xuam¹</i>	‘bitter’
21.	<i>*Fon</i> B	<i>*pfoŋ⁴</i> f?	—	<i>*pu:ŋ³</i>	‘dust; dusty’

22.	*Zip D -i	*tvp ⁸ v?	*zup ⁸	—	'ten'
23.	*Dii B	—	^x dai ⁴ -f	—	'thick, dense'
24.	*Dəm A	*həm ²	*thlam ¹	*thuam ³	'pond, pool; dam'
25.	*χəp D	*kap ⁷	—	—	'bite; chew'
26.	*D# A -i	—	^x day ² -v	—	'carry; hold'
27.	*Dəŋ A	*ɕəŋ ¹	*taŋ ¹	*ɕa:n ³	'arrive, reach'
28.	*Dia B	*həu ⁴	^x dau ⁶ -f	*hñ ² au ¹	'bean'
29.	*Duay C	*ɕəu ¹ -f	*du:i ⁴	—	'bowl, cup'
30.	*Zuk D -i	—	*dzuk ⁸	—	'to bundle'
31.	*D# A	*sú ⁴	—	*thi:p ⁷	'chopsticks'
32.	*Gia A	*kio ⁴	^x gia -fit	*ku- -it	'eggplant'
33.	*Duaŋ B	*səŋ ³	^x daŋ ⁵	—	'weigh'
34.	*Faŋ B	*pwiaŋ ^{1/4}	*pwaŋ ⁵	—	'bank [paddy]'
35.	*Gut D -t	*ku ⁸	^x kwət ⁷	*hn ² ut ⁷ ?	'dig, hollow out'
36.	*χəy C	*kai ⁴	^x ke ⁴ -ti	*xa:i ³	'excrement'
37.	*Giaay B	*xi ⁴ -ft	^x ki ^{2/6} -tif	—	'ride on horseback'
38.	*h̄rwa A	*sa ¹	*khra ¹	*ly ² ai ¹	'ear'
41.	*Giaw B	*khiew ⁴	—	—	'cross the arms'
44.	*Deew A	—	—	*ɕ ² o:i ²	'line, row'
46.	*Piaa B/C -t	*pfa ³	—	—	'cloth, clothing'
47.	*Gooŋ A	*tsaŋ ²	—	^x kaŋ ¹ -iv	'to moan'
48.	*Zaaw C -i	*həu ¹	—	—	'pole; pillar'
49.	*Gwaa A -t	—	*hwa ¹	—	'right [hand]'
50.	*Zuk D -i	—	*zuk ⁸	—	'ripe; done; cooked'
51.	^x D# C -v	*həŋ ³	^x gjuŋ ⁴ -iv	—	'pole; staff'

5 Counterparts in other language families

Finally, when one looks beyond Tai-Kadai, an inordinate number of the fifty-one or so forms are found in Chinese (and, often, in Tibeto-Burman as well), in Austronesian, or in Mon-Khmer languages. In fact, around two-thirds of the suggested forms displaying the critical correspondence patterns in Tai have counterparts in such non-Tai-Kadai languages.

The Old Chinese forms are of particular significance. The preponderance of forms from Chinese indicates that Chinese was a major donor to the Tai languages, a finding that is certainly not new. It is significant that, of the fifty-one forms being considered, thirty-four have Chinese counterparts. Notice that the direction of borrowing must overwhelmingly, although not necessarily exclusively, be from Chinese into Tai. The possibility, of course, exists that some forms were borrowed in the other direction, that is, from Tai into Chinese.

Table 4: Old Chinese (OC) forms

	PTai (GWT)	GSR	Karlgren OC	Baxter OC	Mandarin	
2.	* <i>Gum</i> A	AD 378	* <i>k'âm</i>	—	—	'pit; ditch'
8.	* <i>Gun</i> -t	231j	* <i>d'ïwan</i>	—	<i>kǔn</i>	'to bundle'
9.	* <i>Zək</i> D	851a; 877q	* <i>djĕk</i> ; * <i>d'iek</i>	—	—	'enemy; war'
10.	* <i>Dwak</i> D	—	—	Li (1976) * <i>thjuk</i> 'to butt'	<i>jǐzhòng</i>	'hit target; correct; cheap'
11.	* <i>DII</i> A -i	315a	* <i>tâd</i>	* <i>tats</i>	<i>dài</i>	'hold; carry'
12.	* <i>Dək</i> D t?	961h'	* <i>d'ək</i>	* <i>dIk</i>	<i>gōngde; tè</i>	'male, young [animal]'
13.	* <i>Puu</i> C	101a; 102a	* <i>pĭwo</i> , * <i>b'ïwo</i> 'father'	—	<i>gōng[jǐ]</i>	'person; male'
14.	* <i>χaw</i> C	AD 601	* <i>χâu</i> 'a kind of rice'	—	<i>mǐ</i>	'rice'
17.	* <i>Paa</i> B	25o	* <i>p'wâ</i> 'break'	* <i>phajs</i>	<i>pò</i>	'split, hew'
18.	* <i>Duam</i> B -1	614c; 643g	* <i>g'əm</i> 'submerge; flood'	—	<i>zhǎngcháo</i>	'flood; submerge; soak'
19.	* <i>Diaŋ</i> A	AD 973	* <i>d'âng</i>	—	<i>táng</i>	'sugar'
20.	* <i>χəm</i> A -t	49u	* <i>k'o</i>	* <i>kha?</i> 'bitter'	<i>kǔ; suān</i>	'bitter'
21.	* <i>Fon</i> B	374a	* <i>d'jĕn</i>	* <i>drjĭn</i>	<i>chén</i>	'dust; dusty'
22.	* <i>Zip</i> D -i	686a	* <i>d'əp</i>	—	<i>shǐ</i>	'ten'
23.	* <i>Dii</i> B	10831	* <i>d'jōg</i>	—	<i>chóumǐ</i>	'thick, dense'
24.	* <i>Dəm</i> A	AD 969	* <i>d'âm</i>	—	<i>tán</i>	'pond; pool'
25.	* <i>χəp</i> D	660f	* <i>tsəp</i> 'bite'	—	<i>yǎo; jiáo;</i> <i>dīng</i>	'bite; chew' cf. 'hold in jaws'
28.	* <i>DIIa</i> B	118ac	* <i>d'u</i>	* <i>dos</i>	<i>dòu</i>	'bean'
29.	* <i>Duay</i> C	116a	* <i>tu</i>	* <i>to?</i>	<i>dǒu</i>	'bowl, cup'
30.	* <i>Zuk</i> D -i	AD 900	* <i>sĭwok</i>	—	<i>kǔn</i>	'to bundle'
31.	* <i>Dii</i> A	45j	* <i>d'jō</i>	—	<i>zhù; kuàizi?</i>	'chopsticks'
32.	* <i>GIIa</i> A	AD 342	* <i>g'ia</i>	—	<i>qiézi?</i>	'eggplant'
33.	* <i>Duaŋ</i> B; * <i>jəŋ</i> B	894g	* <i>t'jəŋ</i>	Li (1976) * <i>thjəŋ</i> ; * <i>thjəŋh</i>	—	'weight; weigh down'

34.	* <i>Faŋ</i> B	181k; 389j; 390a	* <i>b'wân</i> ; * <i>piěn</i> ; * <i>b'ïěn</i>	—; * <i>pjin</i> ; * <i>bjin</i>	<i>bîn</i> ; <i>pîn</i>	'bank of paddy; shore'
35.	* <i>Gut</i> D -t	496p	* <i>k'wət</i>	* <i>khut</i>	<i>jé</i> ; <i>wā</i>	'dig, hollow out'
36.	* <i>χəy</i> C	561d	* <i>χiar</i>	* <i>xJi</i> [?]		'excrement'
37.	* <i>Giaay</i> B	1u	* <i>g'ia</i> ; * <i>g'jię</i>		<i>qí</i>	'ride on horseback'
38.	* <i>Hrwa</i> A	981a	* <i>niəg</i>	* <i>nj</i> ?	<i>ěr</i> ; <i>èrduo</i>	'ear'
46.	* <i>Piaa</i> B/C -t	102j	* <i>pwo</i>	—	<i>jǐn</i> ; <i>bù</i>	'cloth; clothing'
47.	* <i>Gooŋ</i> A	AD 69	* <i>ng</i>	—	<i>shēnyín</i>	'to moan'
48.	* <i>Zaaw</i> C -i	129h	* <i>d'iu</i>	—	<i>zhùzǐ</i>	'pole'
49.	* <i>Gwaa</i> A -t	995i	* <i>giüg</i>	—	<i>you</i>	'right [hand]'
50.	* <i>Zuk</i> D -i	1026a	* <i>dìòk</i>	—	<i>shú</i>	'ripe; cooked'
51.	* <i>Dij</i> C -v	722b	* <i>d'iang</i>	—	<i>dòng</i>	'pole, staff'

Several of the forms in Table 4 merit further comment. (18) The form for 'flood; submerge; soak' is apparently a member of a word family. In any case, it looks to be borrowed. (20) The form for 'bitter' is intriguing because of the final *-m*. As Table 4 and Table 5 (Tibeto-Burman counterparts) both show, the Sino-Tibetan form for 'bitter' ends in a vowel, not with a final *-m*. Precisely where the final came from is not clear. (25) The forms for 'bite; chew' and various other related etyma are probably not, despite their presence in Chinese, Tibeto-Burman, and Tai, original to any of these groups, but rather, as Table 6 (Mon-Khmer counterparts) shows, are probably ultimately Mon-Khmer borrowings. This should not be too surprising as Austroasiatic speakers (Mon-Khmer) inhabited much of mainland Southeast Asia before the arrival of the Sino-Tibetan or Tai speakers. Various other forms in the table might also be originally Mon-Khmer, rather than Chinese, but in most cases it was probably from Chinese that the Tai speakers borrowed the forms. (24) The form for 'pool; pond' has a Chinese counterpart, but it also seems to have a Malay counterpart in *kolam*, which apparently has a Tamil origin.

Just from the forms in Table 4, it is evident that although some of the borrowing might have gone the other way, the bulk of the borrowings went from Chinese to Tai. The argument that a form went from Chinese to Tai is particularly strong when a form is also reconstructed for Tibeto-Burman subgroups. (25) For the forms for 'bite; gnaw', it is possible that these are ultimately of MK origin. (35) For the form for 'dig; hollow out' there is a good PTB cognate, and, as Gedney suspected, a separate form for 'scrape; grate' with the shape **khuut*?

A caveat is in order: only a handful of the Tibeto-Burman forms were examined here. Thus, no particular significance should be attached to gaps in the Tibeto-Burman data. If the direction of potential borrowing were more of a question, more time and energy would have gone into identifying TB counterparts.

The oldest layer of borrowings should be MK as Austroasiatic speakers are the oldest speakers found in the region (Table 6). (5) The form for 'to rake', if it is related, has undergone some changes in Tai. (19) The form for 'sugar' is also found in Chinese. Thus, it is not clear what the donor language for the Tai form was. What is clear is that 'sugar' was borrowed into Tai-Kadai. (25) MK is the apparent source for 'to bite'. It reconstructs with

basically the same form in four subgroups of MK. (28) The form for ‘bean’, which also occurs in Chinese, appears to be ultimately of MK origin. (31) The form for ‘chopsticks’ appears in MK, although it is quite possible that this reflects an ultimately Chinese origin.

Table 5: Tibeto-Burman counterparts

	PTai	PNN (French)	WT	PLB	PTB (PKB)	OC (Karlgrén)	
7.	* <i>Gɿay</i> A	—	—	—	* <i>krwiy</i>	—	‘son-in-law’
20.	* <i>χəm</i> A -t	* <i>C-kha</i>	<i>kha</i>	* <i>ka</i> ²	* <i>ka</i>	* <i>k'o</i>	‘bitter’
21.	* <i>Fon</i> B	<i>pan</i> ¹ (Konyak)	—	—	—	—	‘dust, dusty’
22.	* <i>Zip</i> D -i	—	—	—	* <i>gip</i>	* <i>d':əp</i>	‘ten’
23.	* <i>Dü</i> B	—	<i>mt hug</i>	* <i>tu</i> ¹	* <i>tow</i>	* <i>d':iôg</i>	‘thick’
24.	* <i>Dəm</i> A	—	—	* <i>tum</i> ²	—	* <i>d':âm</i>	‘pond; pool’
25.	* <i>χəp</i> D	* <i>gək</i> -f	—	* <i>m-gwap</i> ; * <i>C-kwap</i>	<i>hap</i>	* <i>tsəp</i>	‘bite; gnaw’
35.	* <i>Gut</i> D -t	—	—	—	* <i>r-ko-t</i>	* <i>k:wət</i>	‘dig, hollow out’
36.	* <i>χəy</i> C	—	—	* <i>kyiy</i> ²	* <i>kliy</i>	* <i>χiər</i>	‘excrement’
37.	* <i>Gɿay</i> B	—	—	* <i>dzi</i> ²	—	* <i>g':ia</i> ; * <i>g':jię</i>	‘ride [horse]’
38.	* <i>ɦrwa</i> A	* <i>na</i>	<i>rna</i>	* <i>s-na</i> ² -	* <i>g-na</i>	* <i>ni̯əg</i>	‘ear’
51.	* <i>Dij</i> C -v	* <i>thuŋ</i>	—	—	—	* <i>d:iang</i>	‘post; house’

Table 6: Mon-Khmer counterparts

	PTai	PNB (Smith)	PSB (Efimov)	Wa (Diffloth)	PKatuic (Peiros)	PKatuic (Thomas)	
5.	* <i>Graay</i> B	* <i>kuč</i>	—	—	—	—	‘to rake’
19.	* <i>Diaŋ</i> A	—	* <i>srəda:ŋ</i>	—	—	—	‘sugar’
25.	* <i>χəp</i> D	* <i>kăp</i> ‘eat’	* <i>kap</i>	* <i>kăp</i>	* <i>kăp</i>	* <i>kap</i>	‘bite’
28.	* <i>Dia</i> B	* <i>tòh</i>	* <i>tu:h</i>	* <i>tuh</i>	—	—	‘bean; pea’
31.	* <i>D#</i> A	—	* <i>duə</i> ¹ <i>h</i>	—	—	—	‘chopsticks’
34.	* <i>Faŋ</i> B	—	—	* <i>plak</i>	—	—	‘bank [river, paddy]; shore’

Several of the forms in the Gedney list have apparent Austronesian counterparts (Table 7). Bear in mind that Austronesian speakers were located south of the Yangtze roughly 8,000 years ago and that a number of the loans that they left behind have already been identified (Thurgood 1994). (6) The form for ‘house; hut; tent’ is reconstructed for PChamic (PC), but it is undoubtedly borrowed from the neighbouring MK languages. (18) The forms for ‘soak’ are found throughout Austronesian, but are, as mentioned already, part of a word

family, which I suspect to ultimately be MK in origin. (24) The form for 'pond; pool' is found in Malay, where it is a borrowing from Tamil. (25) The form for 'bite', despite being found in PC, is ultimately a MK borrowing. (34) The form for 'bank; shore' is found in Austronesian as well as in MK.

Table 7: Austronesian counterparts

	PTai	PChamic	Malay	PMalayic	PMP	PAn	
4.	* <i>Dian</i> B	* <i>huta:n</i>	<i>hutan</i>	* <i>hutan</i>	* <i>qutan</i>	* <i>quCaN</i>	'forest; jungle'
6.	* <i>Dian</i> C-t	* <i>sa:ŋ</i>	—	—	—	—	'house; hut; tent'
18.	* <i>Duam</i> B-1 'to flood'	* <i>tram</i>	<i>rëndam</i>	—	—	* <i>edem</i> ; * <i>Redem</i>	'soak'
24.	* <i>Dəm</i> A	—	<i>kolam</i>	—	—	—	'pond, pool'
25.	* <i>χəp</i> D	* <i>kăp</i>	—	—	—	—	'bite'
27.	* <i>Dəŋ</i> A	—	<i>datang</i>	* <i>datŋ</i>	* <i>dateŋ</i>	—	'come; arrive'
34.	* <i>Faŋ</i> B	—	<i>těbing</i>	—	* <i>te[m]biŋ</i>	—	'bank; shore'
45.	* <i>Zuam</i> C	* <i>rumah</i>	<i>rumah</i>	* <i>rumah</i>	* <i>Rumaq</i>	* <i>Rumaq</i>	'house'
51.	* <i>Dij</i> C-v	—	<i>tiaŋ</i>	—	* <i>tiaŋ</i>	—	'pole; post'

5 Conclusions

The obvious conclusion one reaches in evaluating the set of forms pulled together by Gedney is that, as a whole, the forms are overwhelmingly, if not exclusively, borrowed into Tai. An examination of Table 8 makes the case rather convincingly. It summarises the discussions of data in this paper, with the forms listed by the same numbers used earlier in this work. Each of the fifty-one forms is listed, along with two types of information: whether there is a problem with the reconstruction either within the Tai subgroup (Tai problem) or within Tai-Kadai (TK) and whether an apparent counterpart turns up outside of Tai-Kadai, that is, in Old Chinese (OC), in Tibeto-Burman (TB), in Mon-Khmer, or in Proto Austronesian (PAn).

The significance of the irregularities of the reconstructions within Tai-Kadai is that such irregularities are a potential indication that the forms were borrowed. It is worth pointing out that the converse proves little, that is, the fact that a form behaves regularly does not make it a native form, although it may indicate that it was borrowed into the proto-language, as forms borrowed into a proto-language before it breaks up into dialects behave just as regularly as do the native forms. Note that 13 forms show significant irregularities in their Tai correspondences and 12 show irregularities in their wider Tai-Kadai correspondences. The fact that a large number of the Tai forms do not show up at all in wider Tai-Kadai also raises some questions. Of the 51 forms in Gedney's list, 25 or roughly half have reconstruction problems in either Tai or in Tai-Kadai.

The second question summarised in the table is whether an apparent counterpart shows up outside of Tai-Kadai. Aside from Old Chinese, with a handful of exceptions, a form was only considered a counterpart if it is reconstructed in another family. Note that 34 forms occur in OC, with several more occurring in modern Mandarin, that is, roughly two-thirds of the forms show up in OC. It is unlikely that OC has borrowed two-thirds of the Tai

vocabulary. Another dozen show up in Tibeto-Burman, 6 more in Mon-Khmer, and 8 more in PAn, although 2 of the PChamic forms are obviously borrowed from MK. In total, roughly 40 of the 51 forms show up in other language families.

Table 8: An overall evaluation of Gedney's forms

	PTai	Tai problem	TK	OC	TB	MK	PAn	
1.	*Fəy A	—	—	—	—	—	—	'boil; ulcer'
2.	*Gum A	—	—	—	—	—	—	'pit; ditch'
3.	*həəŋ B	—	—	—	—	—	—	'young [chicken]'
4.	*Dian B	—	—	—	—	—	x	'forest; wild'
5.	*Grai B	—	—	—	—	x	—	'rake, to'
6.	*Diaŋ C -t	—	—	—	—	x	x MK	'hut, field'
7.	*Giay A	—	—	—	x	—	—	'son-in-law'
8.	*Gun -t	—	—	x	—	—	—	'to bundle'
9.	*Zək D	—	—	x	—	—	—	'enemy; war'
10.	*Dwak D	—	—	x	—	—	—	'hit target; correct'
11.	*Dɨ A -i	—	—	x	—	—	—	'hold; carry'
12.	*Dək D t?	—	—	x	—	—	—	'male, young [animal]'
13.	*Puu C	—	—	x	—	—	—	'person; male'
14.	*χaw C	—	—	x	—	—	—	'rice'
15.	*Fan A	—	—	x	—	—	—	'sharpen; grind'
16.	*Faay B	—	—	x	—	—	—	'side, part'
17.	*Paa B	—	—	x	—	—	—	'split; hew'
18.	*Duam B -l	—	—	x	—	—	x	'flood, to'
19.	*Diaŋ A	—	—	x	—	x	—	'sugar'
20.	*χəm A -t	—	—	x	x	—	—	'bitter'
21.	*Fon B	—	—	x	x	—	—	'dust; dusty'
22.	*Zip D -i	—	—	x	x	—	—	'ten'
23.	*Dii B	—	—	x	x	—	—	'thick, dense'
24.	*Dəm A	—	—	x	x	—	x	'pond; pool'
25.	*χəp D	—	—	x	x	x	x MK	'bite'
26.	*Dɨ A -i	—	x	—	—	—	—	'carry; hold'
27.	*Dəŋ A	—	x	—	—	—	x	'arrive, reach'
28.	*Dia B	—	x	x	—	—	—	'bean'
29.	*Duay C	—	x	x	—	—	—	'bowl, cup'
30.	*Zuk D -i	—	x	x	—	—	—	'to bundle'
31.	*Dɨ A	—	x	x	—	—	—	'chopsticks'
32.	*Gia A	—	x	x	—	—	—	'eggplant'
33.	*Duaŋ B	—	—	x	—	—	—	'weight; weigh down'
34.	*Faŋ B	—	x	x	—	x	x	'bank (paddy); shore'

35.	*Gut D -t	—	x	x	x	—	—	'dig, hollow out'
36.	*χəy C	—	x	x	x	—	—	'excrement'
37.	**Giay B	—	x	x	x	—	—	'ride on horseback'
38.	*ɦrwa A	unique	—	—	x	—	—	'ear'
39.	*Goin A/B -t	x	—	—	—	—	—	'log'
40.	*Fia A	x	—	—	—	—	—	'braid; harrow; loom'
41.	*Giaw B	x	—	—	—	—	—	'cross the arms'
42.	*Zii A	x	—	—	—	—	—	'fifth earthly branch'
43.	*ZaY A -it	x	—	—	—	—	—	'late; in the morning'
44.	*Deew A	x	—	—	—	—	—	'line, row'
45.	*Zuam C	x	—	—	—	—	x	'room; compartment'
46.	*Piaa B/C -t	x	—	x	—	—	—	'cloth; clothing'
47.	*Gooŋ A	x	—	x	—	—	—	'to moan'
48.	*Zaaw C -i	x	—	x	—	—	—	'pole'
49.	**Gwaa A -t	x	—	x	—	—	—	'right [hand]'
50.	*Zuk D -i	x	—	x	—	—	—	'ripe; cooked'
51.	*Diŋ C -v	x	—	x	x	x	—	'pole; post'

Only 3 forms — the first three on the table — emerge untarnished. The remaining 48 show either internal problems within Tai-Kadai or have a counterpart in another language family. Thus, while there may be room for argument about details of individual forms, the overall conclusion seems irrefutable: these forms were borrowed into Tai after the breakup of PT. It is important to note that this conclusion is not so much a refutation of Gedney's paper as an extension and clarification of it. It was Gedney who assembled the forms and it was Gedney who pointed out that they posed a problem for the reconstruction of PT. He then examined several potential internal solutions but he clearly found fault with each of them. The contribution of this paper is to account for the problem through an external solution — borrowing.

The conclusion opens up three obvious areas for future work. First, as Gedney noted in his paper (1989a:254), Li (1977:193) observed that White Tai has *kh-* 'chiefly for words with tone alternations and *x-* for words with no tone alternations'. The forms with the tone alternations are borrowed, with the *kh-* reflex simply a further indication of this fact. The question is to what degree does the White Tai *kh-* versus *x-* distinction correlate with a borrowed versus native distinction? Second, to what degree will the recognition of these words as borrowings help simplify the reconstruction of PT, particularly the vowels. Third, and far more speculatively, will Gedney's puzzle (1989b[1971]), which lays out irregularities in the vowel correspondences of Tai also turn out to be the result of borrowing? If, as I suspect, the answer is yes, then it should be possible to reconstruct a much less esoteric vowel inventory for Tai, while at the same time clarifying not only that a large number of the forms that exist in both Chinese and Tai are borrowed, not inherited, but also documenting the directionality of the borrowing.

References

- Adelaar, K.A., 1992, *Proto-Malayic, the reconstruction of its phonology and parts of its lexicon and morphology*. Canberra: Pacific Linguistics.
- Baxter, William H., 1992, *A handbook of Old Chinese phonology*. Berlin: Mouton de Gruyter.
- Benedict, Paul K., 1972, *Sino-Tibetan: a conspectus*. James A. Matisoff, contributing ed. Cambridge: Cambridge University Press.
- Blood, Henry F., 1967, 1968, 1974, *A reconstruction of Proto-Mnong*. MA thesis, University of Indiana; published 1968, SIL/University of North Dakota, 115pp.; republished 1974, 122pp.
- Blust, Robert A., 1990, 1991, 1992, 1993, 1994, 1995, Austronesian comparative dictionary. [work in progress].
- Bradley, David, 1979, *Proto-Loloish*. Oriental Monograph Series. No.23. Canberra: The Australian National University.
- Burling, Robbins, 1967, Proto-Lolo-Burmese. *IJAL* 33.2. Part II. Bloomington.
- Diffloth, Gérard, 1980, The Wa languages. *Linguistics of the Tibeto-Burman Area* 5/2:1–182.
- Efimov, Aleksandr Yu, 1987, Historical phonology of South Bahnaric languages. PhD thesis. Moscow: Academy of Sciences of USSR, Oriental Institute. 240pp. [Wordlist translated by William Gage].
- French, Walter T., 1983, Northern Naga: a Tibeto-Burman mesolanguage. Two volumes. PhD dissertation. CUNY.
- Gedney, William J., 1989a [1979], Evidence for another series of voiced initials in Proto-Tai. *Selected papers on comparative Tai studies*. Robert J. Bickner, John Hartmann, Thomas John Hudak and Patcharin Peyasantiwong, eds *Michigan Papers on South and Southeast Asian Studies*. The University of Michigan, No.29:229–270.
- 1989b, A puzzle in comparative Tai phonology. *Selected papers on comparative Tai studies*. Robert J. Bickner, John Hartmann, Thomas John Hudak and Patcharin Peyasantiwong, eds *Michigan Papers on South and Southeast Asian Studies*. The University of Michigan, No.29:271–280.
- 1991a, *The Tai dialect of Lungming*. Thomas J. Hudak, ed. Center for South and Southeast Asian Studies. No.39. The University of Michigan.
- 1991b, *The Yay language*. Thomas J. Hudak, ed. Center for South and Southeast Asian Studies, No.38. The University of Michigan.
- 1993, *The Saek language*. Thomas J. Hudak, ed. Center for South and Southeast Asian Studies, No.41. The University of Michigan.
- Hansell, Mark, 1988, The relation of Be to Tai; evidence from tones and initials. *Comparative Kadai: linguistic studies beyond Tai*. Jerold A. Edmondson and David B. Solnit, eds Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics, No.86:239–287.
- Karlgren, Bernhard, 1923 [1973], *Analytic dictionary of Chinese and Sino-Japanese*. [AD] Taipei: Ch'eng-wen Publishing Company.
- 1957, Grammata serica recens. [GSR] *Bulletin of the Museum of Far Eastern Antiquities* 29:1–332.
- Li, Fangkuei, 1965, The Tai and the Kam-Sui languages. *Lingua* XIV:148–179.

- 1976, *Sino-Tai. Genetic relationship, diffusion, and typological similarities of East and Southeast Asian languages*, 230–239. The Japan Society for the Promotion of Science. Tokyo.
- 1977, *A handbook of comparative Tai*. Honolulu: The University Press of Hawai'i.
- Matisoff, James A., 1972, *The Loloish tonal split revisited*. Research Monograph No.7. Center for South and Southeast Asian Studies. University of California: Berkeley.
- 1988, Proto-Hlai initials and tones: a first approximation. *Comparative Kadai: linguistic studies beyond Tai*. Jerold A. Edmondson and David B. Solnit, eds Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics No.86:289–321.
- Peiros, Ilia, 1996, *Katuic comparative dictionary*. Canberra: Pacific Linguistics.
- Sarawit, Mary Elizabeth Sautter, 1973, The Proto-Tai vowel system. PhD dissertation, University of Michigan.
- Smith, Kenneth D., 1972, *A phonological reconstruction of Proto-North-Bahnaric*. (Language Data, Asian-Pacific series, No.2). Santa Ana, CA, SIL. 109pp.
- Strecker, David, 1983, Proto-Tai vowels revisited: a comparison and critique of the work of Sarawit and Li. *Linguistics of the Tibeto-Burman Area* 7/2:33–74.
- Thomas, Dorothy M., 1967, A phonological reconstruction of Proto-East-Katuic. MA thesis, University of North Dakota. Dallas: Summer Institute of Linguistics. microfiche.
- Thurgood, Graham, 1977, Lisu and proto-Lolo-Burmese. *Acta Orientalia* 38:147–207.
- 1988, Notes on the reconstruction of Kam-Sui. *Comparative Kadai: linguistic studies beyond Tai*. Jerold A. Edmondson and David B. Solnit, eds Summer Institute of Linguistics and The University of Texas at Arlington Publications in Linguistics No.86:179–218.
- 1991, Proto-Hlai (Li): a look at the initials, tones, and finals. *Kadai: Discussions in Kadai and SE Asian Linguistics*, 3:1–49.
- 1994, Tai-Kadai and Austronesian: the nature of the historical relationship. *Oceanic Linguistics* 33/2:345–368.
- 1999, *From ancient Cham to modern dialects: two thousand years of language contact and change. With an appendix of Chamic reconstructions and loanwords*. Oceanic Linguistics Special Publications in Linguistics No.28. Honolulu: University of Hawai'i Press.

11 *The rise and fall and rise and fall of Proto Malayo-Polynesian*

STANLEY STAROSTA

1 Introduction¹

The career of the Malayo-Polynesian (MP) language family has had its ups and downs. Originally proposed as a name for the great language family spread across much of the Pacific Ocean, parts of mainland Southeast Asia, and as far west as Madagascar, the term ‘Malayo-Polynesian’ was later replaced by Wilhelm Schmidt’s term ‘Austronesian’, and was then subsequently rehabilitated by Robert Blust (1977:10) as a name for the Austronesian languages spoken outside Taiwan. The hypothesis that all Austronesian languages outside Taiwan belong to a single subgroup was first proposed by Otto Dahl (1973) (Ross 1995b:102, fn.25) and is generally accepted by Austronesian scholars:

The unity of the Malayo-Polynesian languages is probably not open to serious question, and Proto Malayo-Polynesian is readily reconstructible. (Ross 1995b:69)

...2. there is unambiguous evidence for a subgroup, which includes all extra-Formosan languages’... (Blust 1995a:586)

Blust regarded these languages as a first-order subgroup of Proto Austronesian, and while he gave explicit evidence for his claim that these languages formed a subgroup, he simply assumed by default that they were a first-order subgroup of the proto-language. In this paper I will try to show that this assumption is almost certainly incorrect, and I follow up a few of the consequences of this conclusion.

¹ This paper germinated while I was spending five months at the Centre Nationale de la Recherches Scientifiques sur l’Asie Orientale in Paris under a grant from the French Ministère de l’Éducation Nationale de la Recherche et de la Technologie, and owes much to extended discussions I had there with Laurent Sagart and Lawrence Reid. I would like to thank the Ministry for its support, the CRLAO for its hospitality, and Laurie and Laurent for their time and ideas, as well as acknowledge with thanks theirs and Woody Mott’s comments on an earlier draft of this paper. I hereby absolve them of all responsibility for what I have done with their suggestions in the final version.

2 Malayo-Polynesian as a first-order subgroup

The assumption that Malayo-Polynesian is a first-order subgroup of Proto Austronesian is well established in the literature:

I have argued for the following major divisions: 1) Austronesian splits into at least one Formosan subgroup and Malayo-Polynesian (MP = all extra-Formosan Austronesian languages.) (Blust 1990:232)

The Malayo-Polynesian (MP) hypothesis (that all extra-Formosan languages belong to a single first-order An subgroup, while the Formosan languages constitute one or more first-order subgroups) rests on the following phonological (and some non-phonological) innovations:... (Ross 1992:25)

Many scholars consider that the Austronesian language family has four highest subgroups. Three of these subgroups comprise languages confined to Taiwan. The fourth subgroup – Malayo-Polynesian – includes all of the Austronesian languages spoken outside Taiwan. (Bellwood, Fox & Tryon 1995:5)

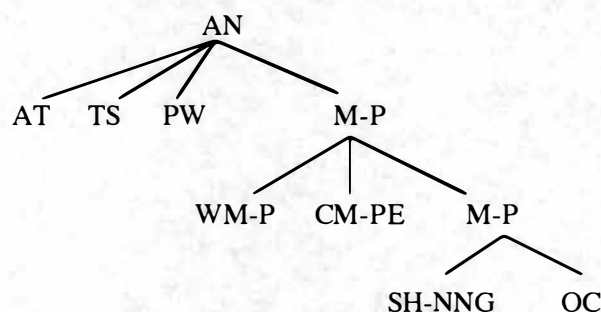
Ross (1992) has partly abandoned this reconstruction on the following grounds: 1) he regards the Formosan and Malayo-Polynesian (MP) languages as members of different highest-order Austronesian (An) subgroups... (Blust 1997:4)

Comparative linguists are now in general agreement about the basic shape of the Austronesian family tree. Most today use the classification developed by linguist Robert Blust (1977, 1978, 1982, 1993, 1995a). This classification...divides the Austronesian family into at least two major groups, of which one, Malayo-Polynesian, includes all Austronesian languages not located in Taiwan (see Figure 4.3). (Bellwood 1997:104)

Words and meanings can only be reconstructed for Proto-Austronesian if cognates are found in the languages of two or more of the primary subgroups (Formosan and Malayo-Polynesian) and if... (Bellwood 1997:110)

The following tree representations from the literature add a bit more detail to the first-order subgroup picture.

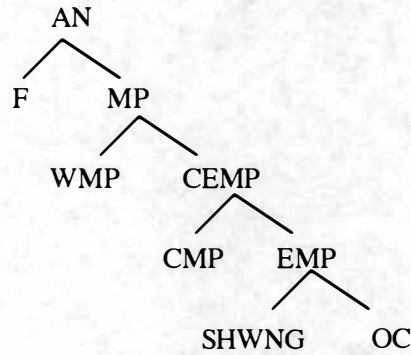
The following subgrouping of the AN language family will be assumed.



AT: Atayalic (Formosa), TS: Tsouic (Formosa), PW: Paiwanic (Formosa), M-P: Malayo-Polynesian (all AN languages outside Formosa)... (Blust 1977:2)

A subgrouping theory that is now accepted by many Austronesian specialists is depicted in Figure 1:

Figure 1: A higher order subgrouping of the Austronesian languages (after Dahl 1976; Blust 1974, 1977a, 1978a, 1982b, 1983/84a)

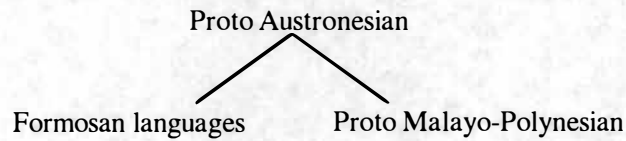


AN = Austronesian, F = Formosan: one or more primary subgroups in Taiwan, here treated for reasons of convenience as a genetic unit. ... (Blust 1988:16-17)

Essentially the same diagram or its upper branches is repeated in later works by Blust and others, for example in Blust (1980:208; 1988:16-17; 1995a:586); Lynch (1993); Tryon (1995:20); Li (1997:157); and Ross (1995b):

Figure 2 shows the major subgroups which developed from one daughter-language of PAN, namely PMP. (Ross 1995b:67)

Figure 10: Schematic diagram of the diversification of Austronesian languages



(Ross 1995b:68; first branching)

Alternative hypotheses are sometimes given a passing mention and then subsequently ignored. For example:

The major subgroups of Austronesian are therefore as follows:...

1.2. Malayo-Polynesian (all Extra-Formosan languages according to Blust, although Reid [1982] excludes some Northern Philippine languages from Malayo-Polynesian and places them in a separate subgroup intermediate between 1.1 and 1.2; see also Starosta [1995] for an even more complex tree at this level). (Bellwood 1997:105)

Assuming the validity of the first-order subgroup hypothesis, anything found in any MP language and in one Formosan language can be reconstructed to PAN, and features found only in Formosan languages cannot. Until quite recently, this was the cornerstone of most work on PAN reconstruction:

Blust (1977) proposed an An family tree which recognized three primary branches in Taiwan (Atayalic, Tsouic, Paiwanic), and a single Malayo-Polynesian branch comprising all the other An languages. In later publications (as Blust 1982; 1983/84a), the Formosan languages are treated for purposes of lexical reconstruction as constituting a single primary branch. The difference...reflected a view that no lexical reconstruction could be safely assigned to Proto-Austronesian (PAN) if its known distribution is confined to the Formosan languages, since the latter have been in close geographical proximity and hence in a potential borrowing relationship for perhaps six millenia [sic]. (Blust 1995a:587-588)

More recently, the MP cornerstone has begun to show some cracks. For example, Formosan languages now have a limited licence to promote their own candidates for PAN reconstruction:

Appendix 1 It differs from earlier reconstructions that I have made in allowing 'Formosan-only' distributions to count as evidence for PAN on the following conditions: 1) reflexes must appear in at least two primary subgroups in Taiwan, and show no phonological irregularities, 2) if reflexes appear in only two primary subgroups in Taiwan these subgroups (or the relevant constituent languages) in general must not be contiguous. Exceptions are made only where the phonological evolution of forms in the two witnesses is so different that a borrowing hypothesis is unlikely.' (Blust 1997:19)

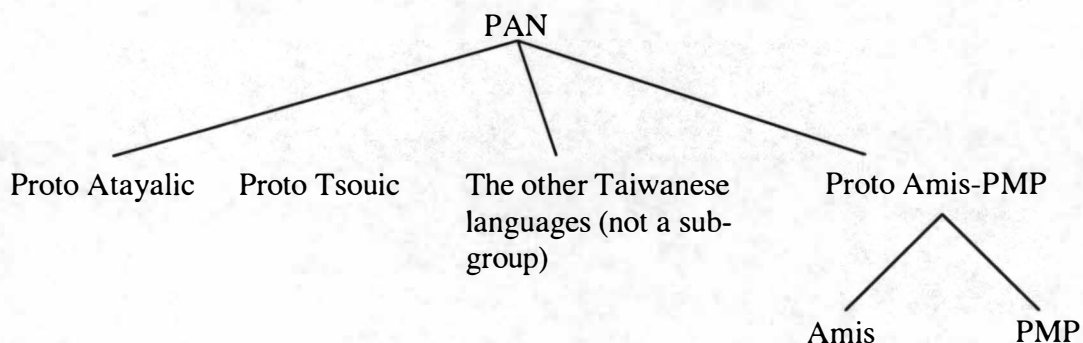
3 MP and the reconstruction of Proto Austronesian

Whether or not Proto Malayo-Polynesian (PMP) is a first-order subgroup of Proto Austronesian (PAN) is an important question. If the answer is positive, then facts about languages in the PMP family can be used directly in the reconstruction of PAN. That is, any property found in one MP language and one non-MP language can probably be reconstructed at the highest level. If the answer is negative, however, the reconstructed PAN language might look much less like PMP. Of course Blust cannot be faulted for failing to give evidence for the first-order subgroup status of PMP, because it is not possible even in principle to give such evidence. Instead, first-order subgroup status is arguably the default assumption until evidence is found for moving a node down the family tree:

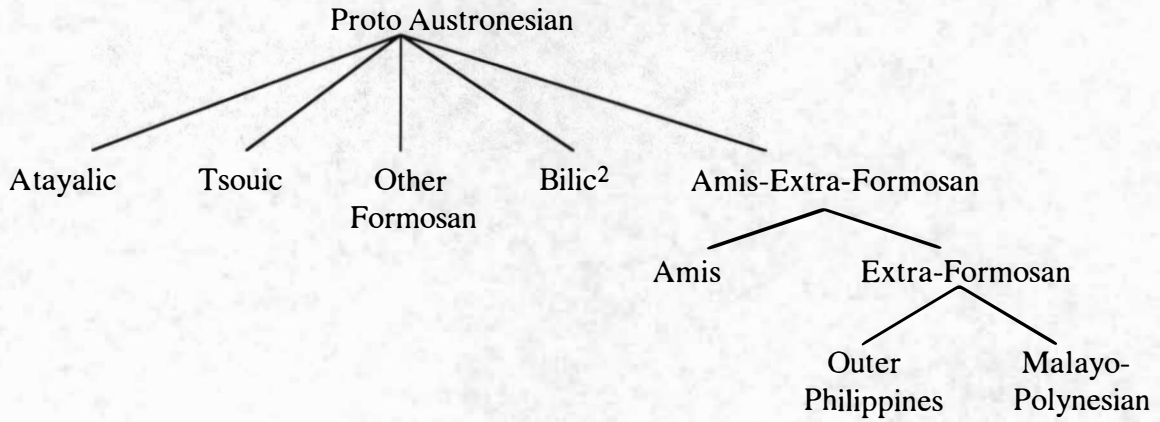
It follows that if no evidence of [exclusively shared innovations] is found for assigning a language or group to a subordinate node within a tree that language or group must be assigned directly to the highest node. Direct assignment to the highest node is, in effect, the default case, and can only be overridden by evidence that clearly favours assignment to a lower node (Figure 1): (Blust 1997:2)

Oddly, however, the requisite evidence for moving the PMP node farther down in the family tree has been available since the appearance of work by Mark Harvey (1979, 1982) and Lawrence Reid (1982) on phonology and continuing with work by Stanley Starosta on verbal morphology (1985, 1994, 1995, 1996). The subgrouping trees (1) through (3) that follow below all reflect the conclusion that PMP is not a first-order subgroup:

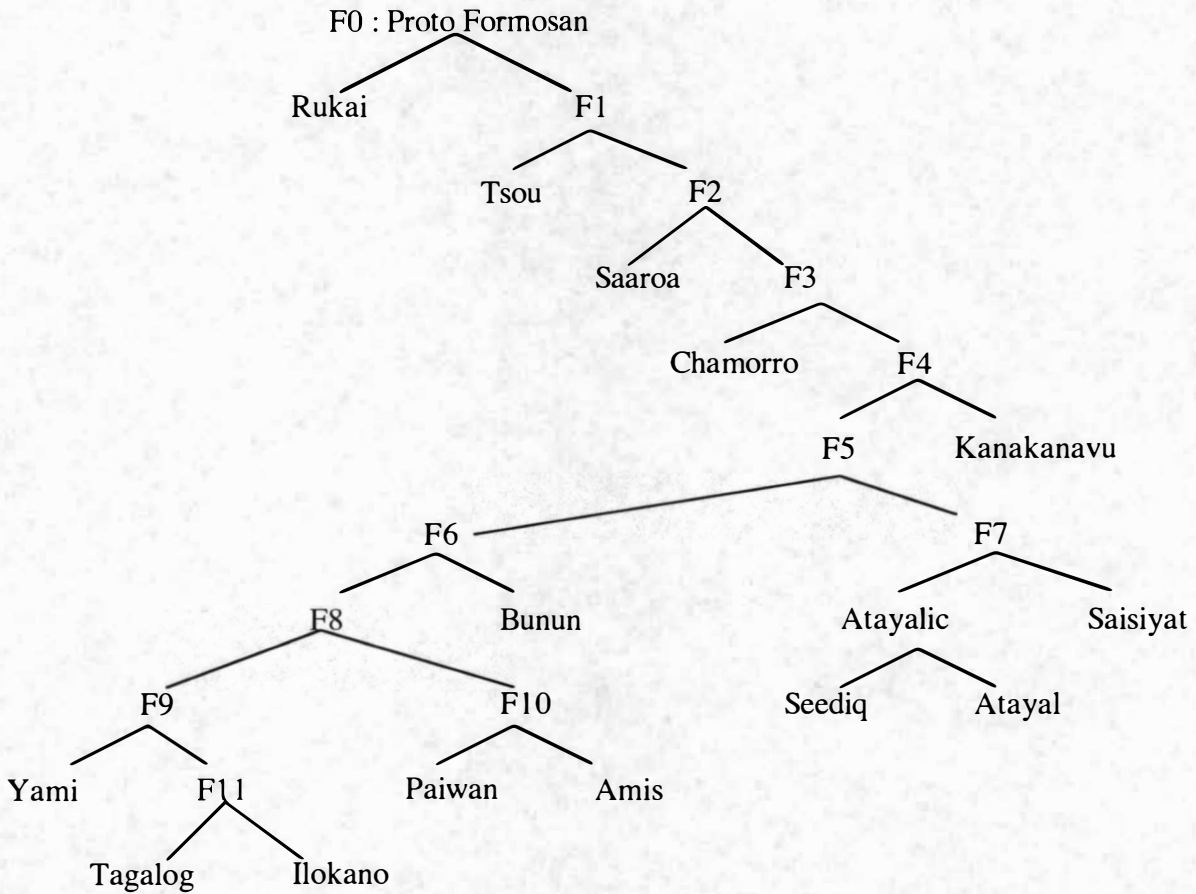
(1) Austronesian higher-order subgroups (Harvey 1979:104, 1982:93):



(2) Austronesian higher-order subgroups (Reid 1982:213):



(3) Grammatically based subgrouping of Formosan languages (adapted from Starosta 1995:691)³



² Laurie Reid informs me (pers. comm.) that he no longer considers Bilic to be a first-order subgroup of PAN.

³ Here F8 is comparable to Reid's Amis-Extra-Formosan (AEF), and F9 to Blust's Malayo-Polynesian.

In spite of the existence of this evidence, however, the PMP node has continued to dangle unconcernedly from the PAN node in the work of many linguists and prehistorians. Peter Bellwood should apparently also be included in the group who accept the lower-order subgroup status for PMP, although the quotes above indicate that earlier he had accepted the first-order subgroup position. Thus, his own recent subgrouping diagram (Bellwood 1997:103, below) shows PMP as subgrouping with one of the Formosan subgroups rather than branching off directly from PAN, in contrast to the claims of Blust and the other first-order subgroup proponents. In a recent email (Bellwood, pers. comm.) he states that he is in agreement with the lower-order position of PMP and attributes the apparent confusion to differences in the interpretation of the term 'first order'.

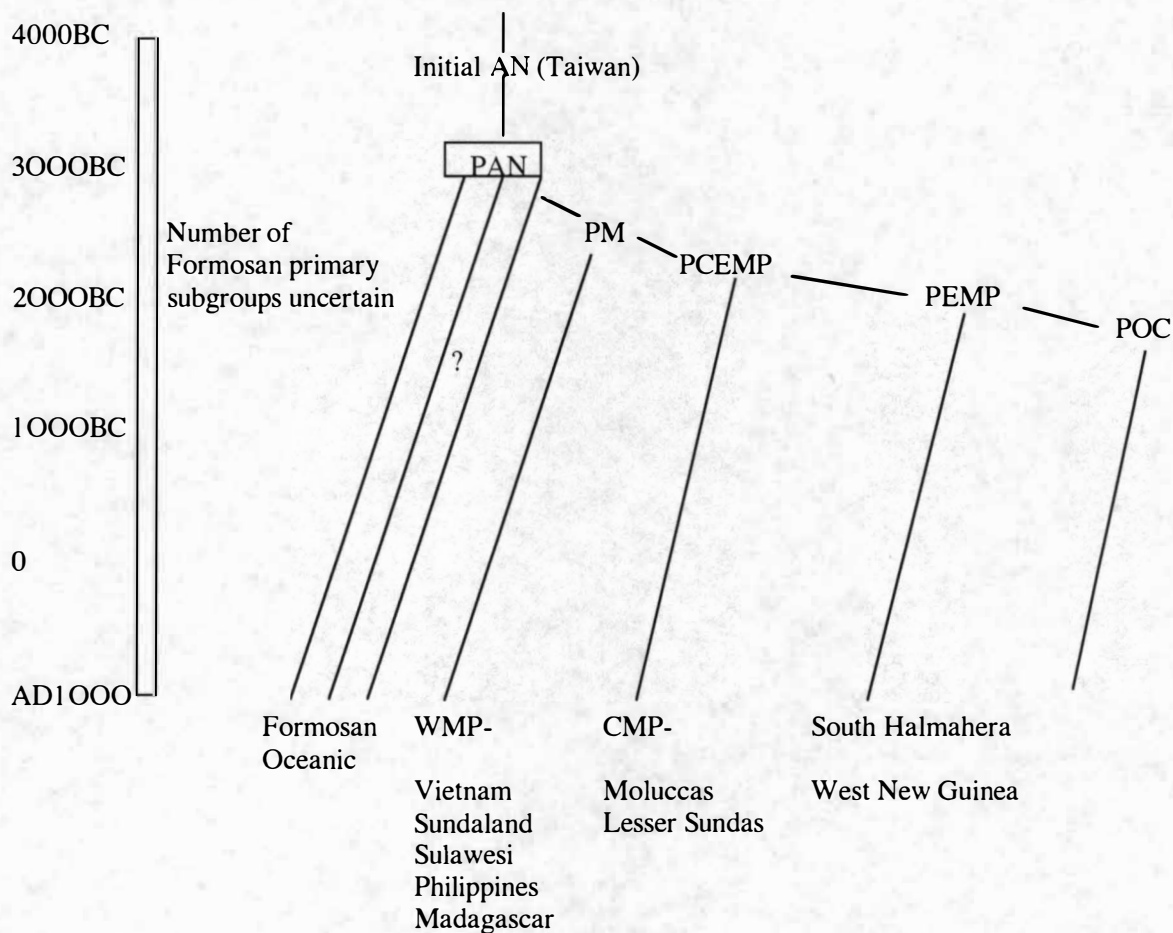


Figure 4.3: A 'family tree' for the Austronesian languages, derived from the subgrouping of Robert Blust

AN = Austronesian; MP = Malayo-Polynesian; WMP = Western Malayo-Polynesian; CEMP = Central-Eastern Malayo-Polynesian; OC = Oceanic.

4 Extra-Formosan as a non-subgroup

Malcolm Ross seems to have been the first person to introduce the concept of 'linkage' into Formosan linguistic studies (Ross 1995b:45ff.). A 'linkage' is a group of languages that has

arisen as a result of dialect differentiation rather than abrupt separation (see Ross 1995b:46), and he thinks that this is the model that best fits the Formosan situation:

If we were to attempt to redraw Figure 1 less ambiguously using the conventions above, we might arrive at something like Figure 9 below. This says that Proto Malayo-Polynesian (henceforth PMP) diverged from the Austronesian languages of Taiwan as the result of separation (when its speakers left Taiwan more than 4,500 years ago). This is a reasonable assumption. (Ross 1995b:47; see also Tryon 1995:20)

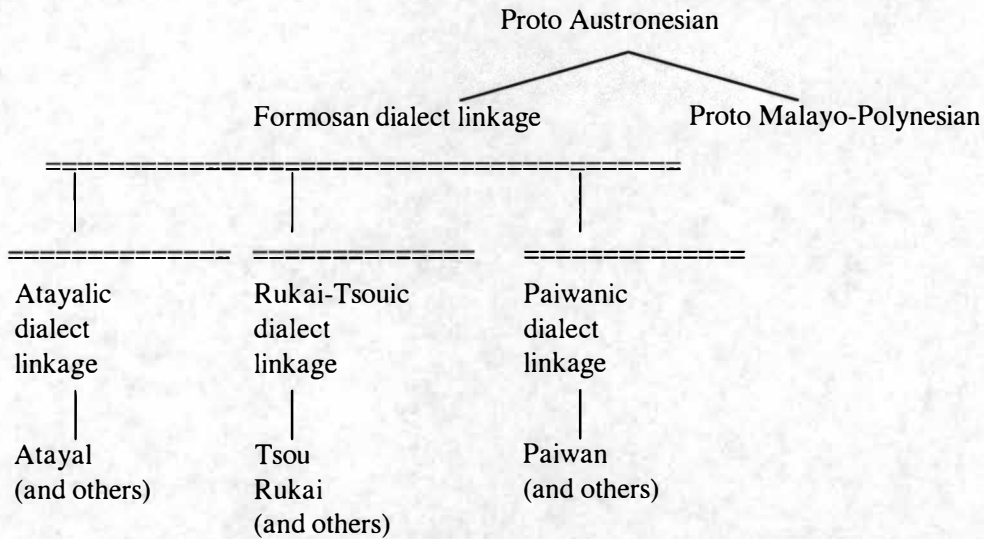


Figure 9: A hypothetical (but not necessarily correct) representation of the earliest branchings of the Austronesian language family. (Ross 1995b:48)

Ross's statement and his diagram raise a paradox that has been lurking in the shadows since the first-order subgroup was first proposed. The diagram and the work of previous first-order subgroup-oriented linguists up until this point assume that PMP broke off from the rest of the family by separation rather than by dialect differentiation. This assumption is crucial, since it is the only justification for regarding MP as a point of triangulation for PAN reconstruction. However, (i) if PAN broke up 5000 years ago (Bellwood, Fox & Tryon 1995:5), (ii) if 'Proto-Austronesian diversified into a linkage of dialects and/or languages before the speakers of what later became Proto-Malayo-Polynesian (PMP) left Taiwan' (Tryon 1995:23), and (iii) if pre-PMP left Taiwan from the east coast some 1000 years after the initial break-up (Li 1997:157–158), then how could the break-off of PMP have possibly been by separation rather than by dialect differentiation? Where did the PMPs spend that millenium while they were waiting for the southbound boat to Botel Tobago and the Batanes Islands? Is it conceivable that, conscious of their linguistic destiny and the needs of later historical linguists, they held themselves in splendid isolation from linguistic intercourse with their relatives for that entire period?⁴ Another quote by Ross indicates that he in fact did not hold such a belief:

4 'As the term implies, languages diverge by separation when two or more communities speaking the same language become sharply separated socially and/or geographically so that contact between them is more or less severed and as a result the two languages change in different ways and perhaps at different speeds.' (Ross 1995b:46).

Circumstantially, it is almost certain that PAN had diversified into a linkage of dialects and/or languages before speakers of what was to become PMP left Taiwan. It is therefore relevant to ask from which part of the linkage this pre-PMP broke off. (Ross 1995b:69)

This latter position, though, is not reconcilable with the 'clean break' scenario that Ross described at the beginning of his article:

...Proto Malayo-Polynesian (henceforth PMP) diverged from the Austronesian languages of Taiwan as the result of separation... (Ross 1995b:47)

and depicted in his Figure 9 above, nor with any of the reconstruction work based on the first-order subgroup hypothesis, all of which crucially assumed such an abrupt separation. That is, the second linkage scenario potentially invalidates any reconstruction based on one MP language and one or more Formosan languages.

To add to the worries of PMP, evidence has been accumulating from work on Formosan languages that even the claims of subgroup status for the extra-Formosan languages are questionable. In each case, properties proposed as evidence for this hypothesis have proven to be either not characteristic of the whole extra-Formosan group or shared with some but not all of the Formosan languages. The impression is becoming stronger that the PMP speech community, if there was one, would have had to be part of a dialect chain spoken in eastern Taiwan. I will refer to this chain, blending terms used by Malcolm Ross (Ross 1995b:48) and Robert Blust (Blust 1997:12, 13), as the East Formosan Linkage (EFL). If this view is correct, then evidence from an appropriately demoted MP grouping of languages is no more and no less privileged as a basis for PAN reconstruction than evidence from other members of this chain which are currently spoken in Taiwan.

4.1 Subgrouping criteria

Do the extra-Formosan languages form a subgroup, and if so, do they form a first-order subgroup? As Blust has stated (1997:2), a subgroup is a first-order subgroup if no evidence of exclusively shared innovations is found for assigning a language or group to a subordinate node within a tree; and, as Harvey (1979, 1982), Reid (1982), and Starosta (1985, 1994, 1995, 1996) have shown, there is evidence that the MP languages share exclusive innovations with languages of eastern Taiwan. Some of the best support for this claim has been provided recently by Blust himself (see §4.5 and §4.6 below), though he has not accepted the conclusion pointed to by his own evidence.

So, the MP languages are not a first-order subgroup. But then, does it follow that they are a lower-order subgroup (LOS)? If we can find a set of innovations shared by all the MP languages and none of the Formosan languages, they are; otherwise, they are not.

In his influential 1997 article, Robert Blust proposed a set of shared innovations which he claimed justified treating the MP languages as a subgroup. I would like to go over the list in light of some of the things we have learned in the past twenty years and show why at least 4 of the 5 criteria can no longer be accepted. Darrell Tryon (1995:22-23) conveniently summarises the innovations proposed by Blust as a justification for the MP subgroup, and I have further abbreviated them here. For 4 of the 5, I will give reasons why each should be rejected.

4.2 #1: The 'Second Austronesian Politeness Shift'

However, in all regions outside Formosa we find that the short form of the pronoun corresponding to **kamu* '2nd pl' is typically, although not exclusively used as a singular pronoun,⁵ explained as a "politeness shift"...Blust concludes that the change **-mu* '2nd pl' > *-mu* '2nd sg' is therefore taken as evidence for a non-Formosan (Malayo-Polynesian) subgroup of the Austronesian languages... (Tryon, ed. 1995:22)

The first problem with the adduction of this criterion is that if this shift is 'typical' but not 'exclusive' in Tryon's words, then it is not an innovation shared exclusively by MP languages. Thus it cannot be used to justify this subgroup. The second problem is that **mu* does in fact appear at least as a component of second person singular forms in Formosan languages, so this property is not coextensive with the MP language family.

Examples:

Rukai:

<i>musu</i>	2s Nom free form, Maga dialect (Zeitoun 1995:139)
<i>-moʔo</i>	2s Nom bound form, Mantauran dialect (Zeitoun 1995:140)
<i>muʔu</i>	2s Topic free form, Mantauran dialect (Li 1996)
<i>mosoa</i>	2s Obl, Tona, Labuan, and Tanan dialects (Zeitoun 1995:140)
<i>mosoanə</i>	2s Obl, Budai dialect (Zeitoun 1995:140)

Kanakanabu (Tsuchida 1976):

<i>musu</i>	2s Gen clitic pronoun
-------------	-----------------------

Saaroa (Tsuchida 1976):

<i>iimukasu</i>	2s Topic free form
-----------------	--------------------

4.3 #2: *-en* focus pronouns

Blust suggests that in languages outside Taiwan the **-en* suffix [on pronouns] indicates goal focus, while on the Proto-Austronesian level the only meaning that can be securely attributed to **a(N)ken* is that of absolute possession.... (Tryon 1995:22)

This claim is rather difficult to interpret because 'focus' is a morphological property of verbs, not of pronouns. Stated in a more precise and theoretically-founded way, the intended meaning may be that a set of pronominal forms ending in a reflex of **-en* occurs in Formosan and extra-Formosan languages to mark absolute possession, but that the same forms also occur in extra-Formosan (MP) languages marking the Agents of transitive ('goal focus') verbs. If that is the intention, the claim is not obviously true. First of all, the Formosan evidence seems rather thin. Of the ten Formosan languages I have personally worked on, only Saisiyat and Paiwan show probable reflexes of this form (Saisiyat *yakin*,

⁵ Blust does not state clearly whether the PMP form is supposed to have been genitive singular or just singular:

...a reflex of **mu* as a 2nd sg. genitive pronoun is unknown in any Formosan language.

Since **mu* is unambiguously reconstructible as a PAN plural pronoun (matching **kamu*), its use as a singular pronoun must have been an innovation...this innovation...took place in...'Malayo-Polynesian'... (Blust 1977:9-10).

Paiwan (*ti-yaken*), but the Saisiyat form is described as accusative (Huang et al. 1996:3) and the Paiwan form as nominative ('Fokus'; Egli 1990:154-155). Neither is characterised as marking absolute possession. Secondly, even if Blust's description of the properties of the **-en* form does characterise some MP languages, it certainly does not characterise all of them. To cite two MP examples, Yami *yaken* is only nominative and not limited to co-occurrence with transitive 'goal focus' verbs, while the Tagalog *akin* indicates absolute possession or the Locative case form, but not transitive agents. It would be interesting to find out just what Blust meant by this criterion, and which MP languages it is supposed to characterise.⁶

4.4 #3: Loss of *S

PAN preconsonantal and final *S disappears, the resultant final shewa [sic] merging with *a in all Malayo-Polynesian languages. (Tryon 1995:22)

I have found no counterexamples to this criterion.

4.5 #4: *maŋ-*

At the same time, another innovation which is reflected right across the Malayo-Polynesian region is the use of the PMP verbal prefixes **paŋ-*, and **maŋ-* to form verbs where the agent is the subject from verbs where the patient is subject... (Tryon 1995:22)

However, as demonstrated long ago by Shigeru Tsuchida (1976:168, 171, 257-258), this morphological property is also found in Formosan languages. The presence of reflexes of **maŋ-* in Formosan languages has also recently been conceded and exemplified by Blust himself:

Typologically most languages that have been called Western Malayo-Polynesian have a prefix reflecting **maN-* which is used in the formation of active verbs, agentive/instrumental nouns in **paŋ-*, and the phonological process of nasal substitution when these prefixes occur with stems that contain certain initial consonants [sic]. These features are not found as active parts of the grammar of any Formosan or CEMP language. However, traces of nasal substitution and of the prefixes **maN-* and **paN-* do appear in some Formosan and OC languages, and thus suggest that their appearance as productive features in WMP languages is a retention from PAN. (Blust 1997:31)

In Amis the pattern not only exists but is fairly productive:

(4) (Wu 1995:98; orthography regularised)

miadop ci aki to fafoy
 hunt Aki pig
 'Aki is going to hunt a pig.'

⁶ For a detailed and theoretically informed reconstruction of these and related forms, see Reid (1997:9-10, 15).

- (5) (M2010833057a;⁷ Chen 1987:83)
sapiadop kina koang no lomaqako
 use.for.hunting this gun of my.family
 'My family uses this gun for hunting.'
- (6) (M2010833058a; Chen 1987:83)
mamiadop kina waco no lomaqako
 assign.to.hunt this dog of my.family
 'My family uses this dog for hunting.'; 'This dog is used for hunting by my family.'
- (7) (M101017.1; Starosta's field notes)
miasik ko babahi i nacila i lomaq
 sweep – woman at yesterday at house
 'The woman swept the floor yesterday in the house.'
- (8) (M2010803055; Chen 1987:80)
mamiasik cira to adawang no cacodadan
 assigned.to.sweep that with front of school
 'He is assigned to sweep the front of the school.'

Amis *mi-* verbs are grammatically intransitive, and *mi-* in forms such as (4) *miadop* and (7) *miasik* are morphophonemically *m-pi-* in structuralist IA (item-and-arrangement) terms. In all cases I am familiar with, the *mi-* corresponds to *pi-* when preceded by another prefix, as shown by (4) *m-pi-adop* versus (5) *sa-pi-adop*. The *mami-* forms are thus not analysable as *ma-mi-*, but must rather be *maŋ-* plus *pi-*, with nasal assimilation of *-ŋ-* to the point of articulation of the following consonant, parallel to the nasal assimilation of *m-* to *pi-* in *mi-* verbs.⁸

4.6 #5: The merger of PAN *C and *t

Other phonological innovations upon which the PMP subgroup is based. These include the merger of PAN **t* and **ts* as PMP **t*. It should be noted, however, that of the languages of Taiwan both Amis and Bunun share this phonological merger. (Tryon 1995:22–23)

The invalidity of the **C*/**t* merger as a criterion for regarding MP as a subgroup has been known for many years. Thus Raleigh Ferrell (1969:64) used this merger to include Yami, an MP language, in his 'Paiwanic II' Formosan subgroup:

⁷ Numbers from Starosta's field notes data base.

⁸ 'I am not convinced yet about the evidence for *maŋ-* in Formosan languages. If it is found, what is the form of the underlying nasal, engma or *n* (as in Chamorro)? I think this is relevant to the whole history of this affix, because I think that the nasal is a frozen ligature, and it shouldn't be engma if it is found in Formosa. The possibility that the assimilation and deletion processes were independently innovated there must also be considered. Much more is still to be said on this topic...' (Laurie Reid, email, 9 October 1998)

Using a somewhat different approach, I have tentatively made a simple two-part division of Paiwanic according to whether the languages do or do not conserve the distinction between the proto-Austronesian phonemes **t* and **C* recently posited by Dyen (1965b). In Paiwanic I (see above, p.25) this distinction is maintained; in Paiwanic II, as apparently in all Austronesian languages outside Taiwan, this distinction disappears.

Mark Harvey (1979:103) used it as a criterion for his 'P-Amis-PMP' subgroup:

However, [PMP] does share the merger of **t* and **ts* to *t* with Amis, Bunun, Siraya and Kavalan. Of these languages, PMP appears to have the most similarities with Amis. Amis has the *ka* pronouns which the other three languages do not.

Finally, based on Blust 1997, Laurent Sagart in an email (pers. comm. 1998) notes that Kavalan, Basay and Trobiawan share this merger.

It is not the case that those linguists most heavily committed to the existence of the MP subgroup have been completely oblivious to the phonological evidence against it. For example, Ross, who crucially assumed the correctness of the MP hypothesis in his reconstruction of PAN verbal morphology (Ross 1995a; see Starosta 1994), was aware of the phonological counter-evidence to PMP. His reaction to it was to relegate it to a footnote and send it to committee:

As noted, the Formosan languages Amis and Bunun also share in the merger of PAN **C* and **t*, and may therefore form a subgroup with PMP. This is a hypothesis which needs further research. (Ross 1995b:101, footnote 11)

Blust (1997:13) gives Ferrell's 'Paiwanic II' even shorter shrift:

The closest previous approximation to East Formosan is Ferrell's 'Paiwanic II' which inexplicably appealed to the less distinctive **t/C* merger rather than the unique **j/n* merger in defining the group. As a result, Kavalan, Amis and Siraya were thrown together with Bunun, and more egregiously, Yami, in a heterogeneous collection of languages which concealed the core of a legitimate and important phylogenetic unit.

Four pages later, Blust (1997:17) recognises the problem with using the **t/*C* merger to support the first-order subgroup hypothesis:

First, as noted above, the merger of **C* and **t* is shared with East Formosan languages and with PMP.

His reaction to this long-known evidence against this last criterion is to simply dismiss it as a 'convergent merger' or 'convergent evolution', that is, as an accidental similarity:

All languages which share the merger of PAN **j* and **n* also share the merger of PAN **t* and **C*. Although the latter change has also taken place in Bunun and in PMP it is otherwise unknown in Taiwan. The simplest hypothesis is therefore to posit three convergent mergers of PAN **t* and **C*: one in Proto-East Formosan, another in Bunun, and a third in PMP. (Blust 1997:13)

Bunun shares certain phonemic mergers with several other languages. However, in every case these appear to be products of convergent evolution. First, as noted above, the merger of **C* and **t* is shared with East Formosan languages and with PMP. (Blust 1997:17)

While the evidence for an East Formosan Linkage (EFL)⁹ is becoming more and more credible, I am unable to follow Blust's reasoning for excluding PMP from membership in it. It is not clear to me how essentially the same set of phonological¹⁰ and grammatical properties can bring Amis safely inside the fence but place PMP beyond the pale.

5 MP languages as part of the EFL

Over the years, circumstantial evidence has been accumulating in support of something like a Greater East Formosan Linkage, which incorporates at least some MP languages. I will present some supportive examples here.

5.1 Linguistic support

1. General

It is claimed [by Wolff] that 'There is clearly much by way of phonological, morphological, syntactic, and lexical innovation that would support the proposition that the Austronesian languages of Taiwan are close to the Philippine languages, especially those of northern Luzon. (Blust 1997:29)

It seems likely that Proto Malayo-Polynesian, the language ancestral to all extra-Formosan languages, may subgroup with a small number of Formosan languages, probably in the south of Taiwan, and research is needed to identify innovations which may be shared by south Formosan languages and Proto Malayo-Polynesian. (Ross 1995a:771)

Interestingly, as noted in Starosta (1996:9), this is the same Ross whose reconstruction of PAN verb morphology in the same paper is completely dependent on the first-order subgroup hypothesis he calls into question in this quotation.

2. Lexicon

Similarities in vocabulary are more obvious across Philippine languages and certain languages of southeast Taiwan than across the rest of Taiwan. Structural similarities, and in particular the elaborate system of verbal 'focus' (in which a wide range of semantic roles may occur as the topic or subject of a clause with each role marked by a distinctive affix on the verb), occur across a range of Philippine languages and some (but by no means all) Formosan and western Indo-Malaysian languages... (Pawley & Ross 1993:9)

⁹ It is possible that the current appearance of intersecting circles of innovations could eventually be replaced by a classic Stammbaum once the innovations from PAN are re-reconstructed in accordance with the LOS hypothesis.

¹⁰ The **C/*t* merger is only one of the mergers that PMP shares with the EFL, but space limitations do not allow me to consider them all in detail in this paper.

5.2 Archaeological support

1. Ferrell's 'Paiwanic II' and the 'Littoral Culture'

Much of aboriginal Taiwan's coastal areas were occupied by a fairly uniform culture complex, which I shall call the Littoral Culture. This complex extended along the east coast, across the northern and southern extremities of the island and throughout the wide southwestern plain (see Figure 2, page 28). The geographical distribution of this culture coincides approximately with that of the coastal Paiwanic II languages (Kavalan/Ketagalan, Ami, Siraya). The Littoral Culture is noticeably more similar to cultures of the southwestern Pacific area than are the other Taiwan aboriginal cultures. (Ferrell 1969:27)

6 PAN that never was?

One conceptual problem that pervades the discussion of phonological and morphological reconstruction in the upper branches of the Austronesian language family tree is the problem of circularity. As is well known, claims about reconstructed forms and innovations depend on assumptions about subgrouping, and vice versa. Thus statements such as 'Language A preserves the PAN distinction between **x* and **y*' cannot be accepted if the PAN reconstructions themselves are based on invalid subgroupings, and a statement that 'morpheme **abc* must be reconstructed for PAN' is invalid if the evidence for reconstructing **abc* is confined to what turns out to be a lower-order subgroup or dialect linkage of PAN.

This problem shows up in the case of Malayo-Polynesian, where reconstructed PAN forms are used to determine innovations and decide whether the MP languages are a first-order subgroup, even though these same forms had originally been reconstructed already assuming that it was a first-order subgroup. If PMP was not a first-order subgroup but rather a member of the EFL, then some forms previously reconstructed at the PAN level based only on MP and EFL evidence might actually go back only as far as Proto EFL. This then undermines claims about what was present in PAN, what was retained, and what was innovated. The following citations illustrate arguments which must be re-evaluated in this light.

Many lexical items which are otherwise widely distributed in An are missing from particular languages. For example, the Atayalic languages show no trace of **maCa* 'eye', **laNiC* 'sky', **Sikan* 'fish', **Caqi* 'excrement', **ina* 'mother', **ama* 'father', **enem* 'six', **walu* 'eight' or hundreds of other words which belong to cognate sets that are represented both in Formosan and in extra-Formosan languages. Are we, therefore, to conclude that the Atayalic languages split off from all others before these lexical items were innovated? (Blust 1997:27–28)

To answer this rhetorical question, we need to go back and re-examine the basis for each of these reconstructions. If one or more of them is reconstructed on the basis of MP and EFL languages only, then it is indeed quite possible that 'the Atayalic languages split off from all others before these lexical items were innovated' in EFL.

One other inference about East Formosan can be justified: members of this group have been in continuous contact with the sea since the break-up of PAN. That is, we can be certain that they never underwent a retreat into the mountainous interior followed by a return to the coast. This inference is supported by the appearance of terms for marine fauna and other terms related to the sea that have cognates over a wide geographical area outside Taiwan, but which are absent from other Formosan languages. These terms include Basay /nutse/ 'squid, cuttlefish' (PAN *(*nN*)*usi*), Kavalan /bubur/ 'jellyfish'

(PAN **bubuR*), /penu/ 'sea turtle' (PAN **peñu*), /umaŋ/ 'kind of hairy hermit crab' (Tsuchida 1994) (PAN **qumaŋ*), and /rayar/ 'sail' (PAN **layaR*), Amis /ʔiso/ 'whale' (PAN **qiSu* 'shark'), and Siraya /pagig/ 'stingray' (PAN **paRiS*). (Blust 1997:13-14)

If these items turn out to have been reconstructed to the PAN level on the basis of eastern Formosan and MP evidence, then they may not have existed at all in PAN, and thus tell us nothing about how the eastern languages got where they are now after the breakup of PAN.

In the area of morphology, Blust states in the abstract for his 1995 paper on the semantics of the *-*an* locative:

The Proto-Austronesian (PAN) affixes **Si-*, **-um-*, **-in-*, **-an* and **-en* played a central role in the construction of utterances. In the more common view, as exemplified by Wolff (1973) all of these affixes except **-in-* functioned as voice markers (VM) in the verb, while **-in-* signaled tense or aspect, often in conjunction with a VM. Starosta, Pawley and Reid (1982), on the other hand, have expressed the view that these affixes had exclusively nominalising functions in PAN, and developed verbal functions at a later period in the history of the An languages. This paper attempts to show that both verbal and nominalising functions must be attributed to the VMs at all stages in their reconstructible history. In particular it argues that the syntactic functions and lexical semantics of **-an* only partially overlapped in PAN and many of its descendants, making it difficult to characterize this VM semantically. More generally the problem of reconciling the syntactic functions and lexical semantics of PAN **-an* raises questions about the discreteness of the inflectional/derivational distinction in morphology. (Blust 1995b)

'The more common view' referred to is of course the first-order subgroup view. If it is incorrect, then a number of these forms and their multiple functions probably did not exist at the PAN level. The necessity of reconstructing both verbal and nominal functions for these morphemes, the fuzzy lexical semantics that comes with them, and the possibly drastic consequences for morphological theory, are unavoidable only to the extent that the first-order subgroup hypothesis is valid, otherwise the problem may not exist. In my 1995 version of the LOS hypothesis, these problems do not arise, since only two of the forms cited by Blust, **-um-* and **-in-*, can be reconstructed at the PAN level at all, and their lexical properties are fairly straightforward.

Analogous problems pervade Malcolm Ross's reconstruction of PAN verbal morphology (Ross 1995a). If the first-order subgroup were correct, there would have been a lot of morphology that got lost on the way to some of the modern Formosan languages (see Ross 1992:12), especially in Rukai:

(ii) Le Rukai présente des caractéristiques syntaxiques que l'on ne trouve nulle part ailleurs. Aucun de ces dialectes n'a préservé le système flexionnel qui caractérise aussi bien les langues aborigènes de Taiwan que celles des Philippines. (Zeitoun 1995:101)¹¹

but if not, most of the Philippine-type verb morphology in Formosan languages turns out to be the result of post-PAN innovations (Starosta 1995).

If the first-order subgroup is wrong, then it will also affect work on prehistory which has assumed its correctness. For example,

¹¹ 'Rukai exhibits syntactic characteristic that one does not find anywhere else. None of these dialects preserved the inflectional system that characterises the aboriginal languages of Taiwan as well as those of the Philippines.' (Author's translation).

To date there is no trace of rice in the earliest (Ta-p'en-k'eng) Neolithic culture identified in Taiwan around 6,000 B.P., although rice appears at around 4,500 B.P. Nonetheless, linguistic evidence is unambiguous in supporting the inference that PAN speakers had rice. Moreover, chronologically earlier cultures in southern China which are arguably the most likely antecedents of the Neolithic populations on Taiwan, also had rice. So, to date at least, the rice which PAN speakers must have cultivated during the initial settlement of Taiwan has, indeed, been lost from the archaeological record. (Blust 1997:27)

If the unambiguous linguistic evidence referred to were based crucially on the assumption that MP is a first-order subgroup of PAN, then it would be only as valid as the increasingly dubious first-order subgroup hypothesis. In this particular case, fortunately, John Wolff (1994) has presented data that confirms the reconstruction of at least the word **pagey* 'rice plant' at the PAN level on anybody's subgrouping, without the need to refer crucially to MP evidence.

7 Conclusion

So, what does the future hold for Proto Malayo-Polynesian, and for the analyses that are predicated on its existence and first-order status? I am afraid the prospects do not look very good for the first-order subgroup hypothesis at least. Robert Blust's 'Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics' (Blust 1997) may be an indication of the shape of things to come. Although the topic of this paper is 'Austronesian comparative linguistics', and although its scope ranges temporally and geographically from supposed Austronesian speech communities on the China mainland to Austronesian migrations from Taiwan to the Philippines and out into eastern Indonesia, there is almost nothing substantive said about Malayo-Polynesian. We do not find the family tree diagram which is normally included in such papers and which would have immediately revealed the author's current position on the status of MP. The table of phonological changes covers only Formosan languages, though adding an MP column would have been easy enough, and would have made the table much more revealing. Table 3, 'A classification of the Formosan languages based on shared innovations in phonology', would also have been much more useful if it had included an entry for MP. I found only one statement that I could interpret as a restatement of the author's old first-order subgroup position:

CONCLUSION: At least nine primary subgroups of the AN language family appear to be represented among the extant and extinct languages of Taiwan. These groups are residual in the sense that they show no positive evidence of exclusively shared innovations by which they could be joined with other languages under a subordinate node within the AN family tree. In this respect they differ from *all* other AN languages, *most* of which exhibit innovations attributable to PMP... (Blust 1997:18; italics mine)

The 'all other An languages' presumably refers to the MP subgroup, but if so, it is difficult to interpret the quantifiers 'all' and 'most'. If MP is a subgroup, there must be one or more shared innovations that define it ('all'). But if some of them do not exhibit the characteristic MP innovations ('most'), what is the basis for assigning these languages to the MP group in the first place? The overall effect one gets from the paper is that of a language family going out of favour at court. One is reminded of successive editions of Soviet history books in which discredited revolutionary comrades have mysteriously disappeared from the old photographs.

To end, let me remind the reader that I have considered two questions in this paper: (1) Is MP a subgroup? and (2) Is it a first-order subgroup? The answer to the second is almost surely negative, but in considering the first, I have addressed only the 5 points originally presented by Blust in support of an MP subgroup and rejected only 4 of the 5. It is quite possible that other shared innovations will be presented which reconfirm the existence of this subgroup. For example, Laurent Sagart (pers. comm.) notes the use of reflexes of **dilaq* to mean 'lick' in Formosan languages but 'tongue' outside. It might be possible to combine this and other lexical evidence with Blust's original criterion #3: Loss of **S*, and perhaps to revise criterion #1: The 'Second Austronesian Politeness Shift' so that it referred unambiguously to monosyllabic genitive second person singular reflexes of **mu* and dealt with the exceptions hinted at by Tryon. This might restore some credibility to the claim that the MP languages constitute a lower-order subgroup. So, MP could rise again someday, but it will almost certainly never again ride as high as it once did.

References

- Baldi, Philip, ed., 1990, *Linguistic change and reconstruction methodology*. Trends in Linguistics Studies and Monographs 45. Berlin, New York: Mouton de Gruyter.
- Bellwood, Peter, 1997, *Prehistory of the Indo-Malaysian Archipelago* (Revised edition). Honolulu: University of Hawai'i Press.
- Bellwood, Peter, James J. Fox and Darrell Tryon, 1995, *The Austronesians: historical and comparative perspectives*. Canberra, Australia: Department of Anthropology, Research School of Pacific and Asian Studies, The Australian National University.
- Blust, Robert A., 1977, The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *University of Hawai'i Working Papers in Linguistics* 9/2:1–15.
- 1980, Early Austronesian social organization: the evidence from language. *Current Anthropology* 21/2:205–247.
- 1988, *Austronesian root theory: an essay on the limits of morphology*. Studies in language companion series, Vol.19. Amsterdam, Philadelphia: John Benjamins.
- 1990, Patterns of sound change in the Austronesian languages. In Philip Baldi, ed. 1990:231–267.
- 1995a, The position of the Formosan languages: method and theory in Austronesian comparative linguistics. In Paul J-K. Li, Dah-An Ho, Ying-kuei Huang, Cheng-hwa Tsang and Chiu-yu Tseng, eds 1995:585–650.
- 1995b, Notes on the semantics of PAN **-an* 'locative'. Abstract, 28th Annual meeting, Workshop 6: Voice in Austronesian. Leiden: Societas Linguistica Europaea.
- 1997, Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. *Papers for Eighth International Conference on Austronesian Linguistics*. 1–54. Taipei: Institute of Linguistics, Academia Sinica.
- Chen, Teresa M., 1987, *Verbal constructions and verbal classification in Nataoran-Amis*. Canberra: Pacific Linguistics.
- Dahl, Otto Christian, 1973, *Proto-Austronesian*. Lund: Studentlitteratur
- Egli, Hans, 1990, *Paiwangrammatik*. Wiesbaden: Otto Harrassowitz.
- Ferrell, Raleigh, 1969, *Taiwan aboriginal groups: problems in cultural and linguistic classification*. Monograph No.17. Taipei: Institute of Ethnology, Academia Sinica.

- Halim, Amran, Lois Carrington and Stephen A. Wurm, eds, 1982, *Papers from the Third International Conference on Austronesian Linguistics*, Vol.2: *Tracking the travelers*. Canberra: Pacific Linguistics.
- Harvey, Mark, 1979, Subgroups in Austronesian. BA Honours thesis. Canberra: The Australian National University
- 1982, Subgroups in Austronesian. In Amran Halim, Stephen A. Wurm and Lois Carrington, eds 1982:47–99.
- Huang, Lillian M., Elizabeth Zeitoun, Marie M. Yeh, Ann H. Chang and Joy J. Wu, 1996, A typological overview of pronominal systems in some Formosan languages. Paper presented at *Fifth International Conference on Chinese Languages and Linguistics*. Hsinchu: National Tsing Hua University. Abstract, Meeting Handbook, 64.
- Li, Paul Jen-kuei, 1996, The pronominal systems in Rukai. Paper presented at *Fifth International Conference on Chinese Linguistics*. Hsinchu: National Tsing Hua University. Abstract, Meeting Handbook, 63.
- 1997, *Taiwan nandao minzude zuqun yu qianxi* [Ethnic groups and migrations of the Taiwan Austronesian people]. Taipei: Formosa Folkways.
- Li, Paul Jen-kuei, Dah-an Ho, Ying-kuei Huang, Cheng-hwa Tsang and Chiu-yu Tseng, eds, 1995, *Austronesian studies relating to Taiwan*, Symposium series of the Institute of History and Philology, Academia Sinica, No.3. Taipei: Academia Sinica.
- Lynch, John, 1993, *Pacific languages: an introduction*. Suva: University of the South Pacific Press.
- Pawley, Andrew K. and Lois Carrington, eds, 1985, *Austronesian linguistics at the 15th Pacific Science Congress*. Canberra: Pacific Linguistics.
- Pawley, Andrew K. and Malcolm D. Ross, 1993, Austronesian historical linguistics and culture history. *Annual Review of Anthropology* 22:425–459.
- Pawley, Andrew K. and Malcolm D. Ross, eds, 1994, *Austronesian terminologies, continuity and change*. Canberra: Pacific Linguistics.
- Reid, Lawrence A., 1982, The demise of Proto-Philippines. In Amran Halim, Lois Carrington, S.A. Wurm, eds 1982:210–216.
- 1997, New evidence for the Austric hypothesis. Paper presented at *Eighth International Conference on Austronesian Linguistics*. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Ross, Malcolm D., 1992, The sound of Proto Austronesian: an outsider's view of the Formosan evidence. *Oceanic Linguistics* 31/1:23–64.
- 1995a, Reconstructing Proto Austronesian verbal morphology: evidence from Taiwan. In Paul Jen-kuei Li, Dah-an Ho, Ying-kuei Huang, Cheng-hwa Tsang and Chiu-yu Tseng, eds 1995:727–791.
- 1995b, Some current issues in Austronesian linguistics. In Darrell T. Tryon, ed. 1995b:44–120.
- Starosta, Stanley, 1985, Verbal inflection versus deverbal nominalization in PAN: the evidence from Tsou. In Andrew Pawley and Lois Carrington, eds 1985:281–312.
- 1994, Rukai-Tsouic: subgroup or treetop? Paper presented at *Seventh International Conference on Austronesian Linguistics*. Leiden: Leiden University.
- 1995, A grammatical subgrouping of Formosan languages. In Paul Jen-kuei Li, Dah-an Ho, Ying-kuei Huang, Cheng-hwa Tsang and Chiu-yu Tseng, eds 1995:683–726.

- 1996, The position of Saaroa in the grammatical subgrouping of Formosan languages. In Suwilai Premrsirat, ed. 1996:944–966.
- Starosta, Stanley, Andrew K. Pawley and Lawrence A. Reid, 1982, The evolution of focus in Austronesian. In Amran Halim, Stephen A. Wurm and Lois Carrington, eds 1982:145–170. ['SPQR'].
- Suwilai Premrsirat, ed., 1996, *Pan-Asiatic linguistics: proceedings of the Fourth International Symposium on Languages and Linguistics*, Vol. III. Salaya, Thailand: Institute of Language and Culture for Rural Development, Mahidol University at Salaya.
- Tryon, Darrell, 1995, Proto-Austronesian and the major Austronesian subgroups. In Peter Bellwood, James J. Fox and Darrell Tryon, eds 1995:17–38.
- Tryon, Darrell T., ed., 1995, *Comparative Austronesian dictionary*, Part 1, Fascicle 1. Berlin, New York: Mouton de Gruyter.
- Tsuchida Shigeru, 1976, *Reconstruction of Proto-Tsouic phonology*. Monograph series No.5, Study of Languages and Cultures of Asia and Africa. Institute for Study of Languages and Cultures of Asia and Africa. Tokyo: Tokyo University of Foreign Studies.
- Wolff, John U., 1994, The place of plant names in reconstructing Proto Austronesian. In Andrew K. Pawley and Malcolm D. Ross, eds 1994:511–540.
- Wu, Joy J., 1995, Complex sentences in Amis. MA thesis, Taipei: National Taiwan Normal University
- Zeitoun, Elizabeth, 1995, Quelques problèmes de linguistique dans les langues aborigènes de Taiwan [Issues in Formosan linguistics]. PhD dissertation. Paris: University Paris 7.