The Tai Languages of Assam
a grammar and texts

Stephen MOREY, B.A. (Hons)

A thesis submitted in fulfilment
of the requirements for the degree of
Doctor of Philosophy

MONASH UNIVERSITY

LINGUISTICS PROGRAM,
SCHOOL OF LANGUAGES, CULTURES AND LINGUISTICS

August 2002

ကျမ္မာခြေကျမ္မာခြေကျမ္မာတ်မင်္ဂရ ။ ကျမ္မာ
ခြေဝိုခမ္ဘရ ။ ကျမ္မာတ်ရေခတ်တဖြမ်းမးနရခရ်ခံ ။
ကျမ္မာတ်ရေခ်မခ် လ၊ေ ကျမ္မာတ်ရေခ်မခ်ကံတွခ်
ကျမ္မာတ်ရေခ်မခ်ဝံပခ် ။ ခ်တ္တေရုတ်ဆံကျ်ုမုခ်ေ
တ်ဆခ်ကခ်ဲခံခ်တ၊ုုေခ ။

ဆ်ုဆ်ကခ်တ်ဆ၊တ်ဆ၊ကျ်ေုုရခ်ုံတ်ေ ။
ဆ်ုလ၊ခ်တ်ဝ၊မ္မတ်ဆ၊ကျ်ေုုရုခ်ေ ။

Table of Contents

Abstract

This thesis presents a comprehensive analysis of the Tai languages spoken in Assam, Northeast India, together with a substantial corpus of texts in the Tai languages. The thesis is presented in two versions; as a book in the traditional format, and as an electronic document, with links to sound and picture files as appropriate.

Of the several Tai varieties spoken in India, this study concentrates on the Tai Aiton and the Tai Phake. Both communities have strongly supported the writer's intention to not only describe their languages but to document and make available as much information about the language and literature of the Tai as possible.

The electronic format in which this thesis was written has allowed all of the language examples to be linked to sound files. The reader of the thesis is therefore able to check any of the claims made by reference to the original data. In the case of language examples that come from texts, the reader will also be able to check the context.

In order to present the language examples in the scripts used by the Tais of Assam, several special fonts were produced. The use of the script not only makes the thesis more available to the native speakers of the languages, it also conveys to the general reader the importance of the native writing system.

In addition to a description of the current linguistic situation of the Tai peoples in Assam, the thesis also includes a comprehensive study of previous work on these languages. Following that, a detailed analysis of the sound systems, writing systems and syntax of the languages is presented. Finally, two lexicons have been complied, an Aiton English lexicon and a Phake English lexicon.

This thesis is both a description and a documentation of the Tai languages of Assam. Over seven hours of text has been transcribed, translated, analysed and annotated and

these texts are all presented, linked to sound files, in the electronic appendix to the thesis.

It is argued in this thesis that the documentation of endangered languages is not only an urgent task, but that it is the documentation, the rich corpus of texts presented, that will be perhaps the greatest contribution to scholarship.

Statement of Originality

This thesis contains no material which has been accepted for the award of any other degree or diploma in any University, or which has been previously submitted for any degree or diploma. It contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Stephen Morey

Acknowledgments

An enormous number of people have made this thesis possible. I owe a great debt to all my teachers, to all my informants and to all those who helped me in diverse othr ways.

This thesis would not even have been conceived but for the great opportunities that I have had in studying linguistics. I am grateful to all my teachers from the Department of Linguistics at Monash University: Dr. Peter Kipka, Dr. Heather Bowe, Dr. Mark Newbrook, Dr. Peter Paul, Dr. Laura Tolfree, Dr. Joanne Winter, Dr. Keith Allan and Professor Michael Clyne. Dr. Anna Margetts, who has newly joined the Monash Department, has also been of great help.

Many people in the Linguistics Department at Monash University have been a great help. I wish to thank Paul Kube and Peter Kipp for copying tapes; Noel Williams, Ben Webb and John Tan for assistance with computer matters, and Lona Gottschalk and Glenda Benness for diverse support over the years.

One of the blessings of undertaking a PhD at Monash University is the high quality of fellow students that one encounters; I cannot name all, but I do want to mention the assistance I have received from Dr. Georgina Heydon, Humphrey van Polanen, Airdrie Farrar, Brigitte Lambert and Kylie Martin.

I am also indebted to several linguists from Latrobe University: It was Prof. Kate Burridge who first encouraged me to study linguistics; and Prof. Barry Blake and Dr. David Bradley both helped me to gain access to important texts on the Tai languages that were otherwise unavailable. Prof. R.M.W. Dixon and Prof Alexandra Aikhenwald from the Research Centre for Linguistic Typology have also been very supportive when I presented some of my ideas at seminars organised at the RCLT. Several Latrobe students have assisted this project: I am grateful to Alec Coupe who

Several people who keenly participated in this project have not lived to see it complete. The late Aimya Khang Gohain was a great scholar of the Tai language; he would have valued this thesis and it would have been a better document if he had been able to comment on the final draft. The late P.K. Buragohain, from an Ahom family, gave considerable support to Tai studies, and I have benefited greatly from the infrastructure that he created to allow scholars to work on Tai studies. The late Dimba Rajkonwar, a descendant of the Ahom Kings, would also have valued this work.

At the very end of my last field work trip for this thesis, I met and worked with Chaichuen Khamdaengyodthai of Chiang Mai, whose very considerable knowledge of all the Tai languages was a great assistance.

** **

Whatever work has been possible in this thesis is due in no small measure to the scholars who have gone before. In the field of the Tai languages of India, the giants of scholarship are the Rev. Brown, who first heard the tones, Sir George Grierson who made a comprehensive study of the languages of India and Dr. Banchob Bandhumedha whose work on Phake phonology and lexicography made this thesis possible.

Doug Cooper of the Southeast Asian Software Research Centre, Bangkok, has provided me with a great deal of help on all matters to do with computing. He set up the dictionary data base, and has explained to me how to make best use of computers. Without his assistance, the electronic version of this thesis would not have been possible.

A thesis cannot proceed without a good supervisor. I have been as lucky as one can be in having as my supervisor Dr. Heather Bowe. She is not only a fine scholar, but is solicitous for my welfare in every way. Whenever I needed her, she was available. Whenever she judged it necessary she challenged the conceptual basis of my thinking and I was thereby able to gain a deeper understanding of the languages studied here.

She encouraged the production of the electronic version of this thesis, and gave very detailed comments on earlier drafts of the text. The final versions of both the text and the electronic thesis are much the better for her help.

The support of one's family is essential. My parents Bruce and Rose Morey have encouraged me throughout. Their support also made it possible to purchase some of the technology that was necessary for this thesis.

The greatest thanks of all are to my wife Kimsiew Chuah; she has put up with frequent absences, considerable financial expense and a husband who has been at times quite overwhelmed by the task. A better support cannot be imagined.

All of the people mentioned above have helped to make this thesis what it is; for any shortcomings I alone am responsible.

Abbreviations used in this thesis

Terms used in sentence glossing

1	First Person
2	Second Person
3	Third Person
CL	Noun class marker
CLF	Classifier
COMP	Comparative
DEF	Definite
DI	Dual
EUPH	Euphonic particle
EXCL	Exclamation
HESIT	Hesitation
HORT	Hortative
I	Indirect Object
INTENS	Intensifier
NEG	Negative
NP	Noun Phrase
O	(Direct) Object
ONOM	Onomatopoeia
Pl	Plural
PP	Prepositional Phrase
PRT	Particle
QN	Question
RECIP	Reciprocal
RESP	Respectful
S	Subject
Sg	Singular
TAM	Tense, Aspect and Modality
V	Verb
VOC	Vocative
VP	Verb Phrase

ถ้าศึกษาภาษา ก็ศึกษามรดกของความคิด
When we study language, we study the heritage of human thought

– พระเปร่ง ปภสุสโร [1]

1. Introduction

Diversity of human language, and arguably therefore of human experience, is under serious threat. The small languages of the world – and most of the languages of the world are small languages – face threats that it is believed many of them will not surmount. The work of documenting the undocumented languages of the world is as urgent a challenge as any in the whole range of scholarship.

The reader will already have realised my ideological commitment to the documentation and description of endangered languages. It is not uncontroversial that the study of endangered languages should be a high priority for linguistic endeavour. However in this work, it will be taken as axiomatic that the diversity of human language should be recorded.

Many scholars have written of the importance and urgency of this work. Mithun (1998:163) wrote that the diversity of human language as "one of our most valuable human intellectual resources". She makes the strong claim that "The loss of language diversity will mean that we will never even have the opportunity to appreciate the full creative capacities of the human mind." (1998:189)

Dixon (1997:5) has written about what he considers to be the major priority for linguists, which is:

> "to get out into the field, and to provide description of some part of the wealth of human language, documenting the diversity before it is – as it will be – lost."

---

[1] In March 2002, when returning from Chiang Mai, having finished all the field work for this thesis, I met the Venerable Phra Preng Pathassaro, who asked me what I did. On being told, he uttered these words, which seem a fitting heading.

Documentation, surely, means more than merely listing a series of sounds and a series of strings of words which tell us what "the language" is or was made up of. Documentation must mean the making available to scholars, and to the future generations of the speech community, as much of the speech and literature of that community as possible, and presenting it in a way which advances the knowledge of human language.

Mithun (1998) encapsulated this point in her conclusion:

"Simply eliciting vocabulary and basic paradigms will not be enough; speakers must be allowed to speak for themselves. It is crucial to record how good speakers use their language, what they choose to say in the multitude of settings that constitute their daily lives, how they describe their own experiences, how they provide explanations, and especially how they interact with each other. Such a record will lay a foundation for their descendants to discover the intricate beauty of a system unlike any other, and a chance for us all to appreciate some of the capacities of the human spirit." (1998:191)

For most of the twentieth century, printing was the mode of presenting "the grammar of a language". Printing imposed limits on the amount of language data that could be presented. Furthermore, the orthography of the community concerned was often unable to be printed and neither the sounds of the language nor any but a limited photographic record was possible.

These constraints do not apply at the beginning of the twenty-first century. With the use of widely available computer technology, in addition to the descriptive thesis itself, a large corpus of language data, both analysed in written form and as audio or video, can be presented.

This thesis consists of eleven chapters which present an analysis of the several Tai languages of Assam. The electronic version of this thesis is linked to a corpus of texts in the Tai languages (see 9.5 below), presented in both audio form and with

transcriptions, translations into English and annotated analyses. This corpus, it is hoped, not only presents some of the intricate beauty of the Tai languages, but allows the speakers to speak for themselves, as Mithun has urged (1998:191).

In the electronic version of this thesis, links will be provided to both sound and picture files. This will involve the presentation of:

- sound files for almost all language examples referred to in the thesis
- sound files for all of the spoken texts on which the thesis is based
- photographs linked to thesis.
- brief biographies of the informants

This, it is hoped, will provide a much richer level of documentation than was possible in the past, without compromising the linguistic analysis that will be presented in chapters 6 to 10.

1.1 Meeting the Tai of Assam

In late 1996, I undertook a semester of study at the Prince of Songkla University, Pattani, Thailand. When my Thai dialectology teacher, Dr. Thananan Trongdi, came to learn that I was planning a holiday to India, he suggested making a visit to Assam, because "there are Tai people there".

So it was that in October 1996 my wife and I first arrived in Assam[2]. Thanks to the kindness of Professor Chatthip Nartsupha of Chulalongkorn University, Bangkok, we had the names of three Tai people in Assam. Only one of these, Nabin Shyam Phalung, was in Assam when we arrived. We visited him at the Department of Historical and Antiquarian Studies, in Pan Bazaar, Guwahati, where he works as an Ahom Pandit.

---

[2] The story of our arrival in Assam formed the basis of the *Poem in the khc² khyāŋ² style: How Stephen Morey came to the Tai Phakes*, composed by Ai Che Let Hailung of Namphakey Village.

By very good chance, a few days after our arrival, Nabin Shyam was due to travel back to his own village for a meeting of the Tai Aitons. We travelled with him and thus came to stay in the Aiton village of Bargaon (�’ baan³ luŋ¹)[3].

From there, we were able to travel to Dibrugarh, and through the kindness of Sri Atul Borgohain, reached the Phake village of Namphakey (�’ mān³ phā⁴ ke⁵ taü³), where by further good chance we were able to witness the annual ceremony to mark the end of the rains retreat of the Buddhist monks (’ pɔi² auk¹ wā¹).[4]

During our time in Namphakey, I had noticed that the Tai villagers possessed many books in manuscript form, but there were no printed books in the Tai language. I presumed that this was because having a script of their own that is read by relatively few people, the Tais had no resources to print books.

I asked the villagers whether it would be useful for them to have a 'computer printing block' for their script. At that time most of them had very little idea what computers were; but the older men in the village gathered together and discussed the matter and agreed that it would be a good idea. Returning to Australia, the first draft of the font was prepared (see 7.8 below) and my involvement with the Tai languages of India had begun.

After that initial three weeks in Assam, I have undertaken five more field trips, as follows:

> January 1998
> January 1999
> November 1999 - February 2000
> March - April 2001
> February - March 2002

---

[1] The Tai scripts are explained in Chapter 7 and the phonemicisation of them in 6.2 (Phake) and 6.3 (Aiton). The name of this village can also be spelled ’.

[4] This festival is also known as ’ pɔi² kan¹ tɔ²'festival-beg pardon'.

Altogether, I have visited about 12 villages in Assam, but I have only stayed in 6 of them, namely the Aiton villages of Borgaon (’ baan³ luŋ¹) and Duburoni (’ baan³ nam³ thum³), the Phake village of Namphakey (’ mān³ phā⁴kɛ⁵ taü³), the Khamyang village of Pawaimukh, the Khamti village of Mounglang and the Turung village of Rengmai.

Field work inevitably involves both some difficulties and some moments of great joy. Sometimes even the difficulties can be turned to advantage. It will be useful to briefly touch on both the difficulties and the joys of this particular project.

1.1.1 Some difficulties in undertaking this thesis

The Northeast of India has been somewhat disturbed by political problems for many years. There are a large number of insurgencies, described by Hazarika (1995). Some of these have as their genesis the desire for independence or autonomy by groups of people who speak a common, often non-official, language.

Naturally these insurgencies are no business of the visiting scholar, but their existence cannot be ignored. For example, among the Tai Ahom community (see below 2.1), there are some who believe in an independent country for Upper Assam. As Nartsupha and Wichasin (1998:179) state:

> "If they can not free all of Assam, they hope to cut Assam into two parts.
> Upper Assam will become an independent country called Ahomia or Mioung
> Dun Sun Kham[5] the old name of the Ahom kingdom."

Sometimes one group in the community or another will call for a *bandh*. These may be called by groups of students or political parties, but they are also called by insurgents. The term *bandh* often translated unsatisfactorily as 'strike'. For a bandh is not a strike, it is something like a curfew. Those calling the bandh will order that all transportation, schools and shops in the affected area be closed, although farmers may

---

[5] In Aiton this would be ’ muŋ² dun² sun¹ kham², literally 'country-cotton-garden-gold.'

continue with their work in the fields. People breaking the bandh sometimes run the risk of physical harm.

I have experienced about ten bandhs in the time that I have been w[ ]ing in Assam. On one occasion, I had been invited to attend a wedding in Moran, Dibrugarh District, Upper Assam. This was to be a full Soklong wedding, my first opportunity to view this ceremony which is conducted in the Tai-Ahom language.

For the three weeks prior to the wedding, I was undertaking field work in the Tai Aiton villages in Karbi Anglong and Golaghat districts. On the day before the wedding, my host, Chaw En Lai Phalung came in and announced that he had learned that a two day bandh had been called in Karbi Anglong district. I would therefore not be able to catch the bus from the nearby town to make my way to Moran.

However, a plan was developed. We would pack my bags and carry them across the river into nearby Golaghat district, then proceed on foot to Duburoni village. Having spent the night there, in the morning a taxi would come and take me through Golaghat district to the Doiyang river. This would have to be crossed by boat. From there, after a short walk, we would reach another bus stop, and then go by bus to Golaghat where a bus to Moran would be available. Since it was regarded as unsafe for me to travel alone, I would be accompanied by two young men of Duburoni village as far as Golaghat.

So it was that my laptop computer, video camera, and all my tapes were carried through the river, which was fortunately only up to waist height at that time of year! With the help of Bidya Thoumoung and Pradip Thoumoung I reached Golaghat, and from there caught my bus to Moran and attended the wedding.

This inconvenience had a positive side. In Duburoni there is a very knowledgeable old man, Sa Cham Thoumoung. On what was supposed to be my last day in the village he began to tell a story for me to record. It was a Jataka, an account of a former life of the Buddha, and was to last for two hours. Before it was completed, I had used up all

my empty cassettes, so Pradip Thoumoung was despatched to the nearest town to buy some blank ones.

After a gap of several hours, I was able to resume recording. Sa Cham, who is nearly deaf, needed to hear the last minute or so of the tape to be sure he knew where he was up to. This was achieved with him wearing a stethoscope, the end of which was then placed on the tape recorder. He then resumed the story and completed it in about another hour, after which I left the village.

When he came to know that I had unexpectedly returned, two days later, he told me that there was still a little more to add. In a Jataka story, every character is a previous incarnation of a person from the life of the historical Buddha, and the story usually ends with an explanation of who the characters were. it was this that Sa Cham wished to add. Thus it was most fortunate that the bandh had caused me to return to the village[6].

1.1.1.1 Crossing a bamboo bridge

The importance of crossing rivers was first brought home to me when I wished to visit Duburoni village in January 1999. In order to do so I had to cross a rather rickety little bamboo bridge[7], with four parallel bamboo pieces on which to walk; although sometimes they were only as wide as an inch and sometimes some of them were missing. The reader may imagine a picture of an Australian, with his computer hanging over one shoulder and tape recorder on the other, crossing this bridge. sure that these machines and their owner would soon fall in. It was getting close to the end of the year, but I safely the crossed that bridge and returned by another, rather longer, route.

Only later did I come to read the book of Tai proverbs, *Grandfather teaches Grandchildren*. Proverb number (11) reads as (1) below:

---

[6] Unfortunately I have not yet had the time to transcribe and translate this story.
[7] Photographs of this bridge can be found in the electronic version of the thesis at \biographies\Bamboo Bridge.doc

1)

    လ�gen,   ပီ   ယ�gen,   တဲ   ၵၟ် ၊

lā³   pī²   yā¹   tai¹   kho⁶

late   year   forbid   move along   bridge

    ၵၟ်   ゞၵ်   ၸၣ်   တၵ်   ၶၟ် ၊

kho⁶   nauk¹   cāŋ⁵   tok¹   khɔi³

bridge   decayed   maybe   fall   die

    မျ   ထမ်   ၸၣ်   အိၵ်   ꩫ ၊

mau¹   thəm⁶   ɪan⁶   ik¹   nā³

NEG   think   back   ɔre   face

    ၸၣ်   ယဵပ်   ၶျ   မၢႆ   ゞၵ်   တၵ်   တဲ ။

cāŋ⁵   yēp¹   khā⁵   mai⁴   nauk¹   tok¹   tāi²

maybe   tread on   branch   rotten   fall   die

Audio Link for this sentence example.

Yehom Buragohain translated this as: "Do not walk over the bridge at the end of the year for you may fall down from the bridge and die." The late Aimya Khang Gohain translated the last two lines a little more explicitly as: "thinking back and fore, you may tread over the rotten branch and fall dead.".

I have since learned that this type of bridge is built after the floods in about September. By January, coming close to the end of the Tai year, the bridge is getting rather rotten.

1.1.2 Life in Tai villages

Generally speaking, Tai people get up with or before the sun. In Assam, this is very early indeed, because India is a single time zone, and Assam is so far to the east in that time zone that the middle of the day is only a little past 11.00 AM by the clock.

After getting up, tea is offered, consisting of a cup of Assam tea and some biscuits or bread. A full rice meal will be taken around 9.00 or 10.00 AM and another meal after dark just before sleeping. I have never gone hungry in a Tai village, however. Tea and fruit and cake are offered throughout the day.

The main work of most people is rice farming. The rice season lasts from about May until December, with planting in May and harvesting beginning about four months later. After harvest the rice has to be separated from the stalks; this is done by having cattle trample on the stalks. During November and December this is a daily activity and can sometimes be heard in the background on some of the recordings made for this project (such as the *Butterfly Dance*).

During the rice season, the villagers are very busy. In the winter months, January to April, there is less work to do, and villagers were able to spend time helping with this Tai languages project.

There are several important festivals throughout the year. Some of the most important festivals are listed in Table (1):

Name of festival			Time of year
ပွႆးသၢင်းၵ်ꩡ်ꩫ်	poi² sāŋ² kyɛn²	Tai New Year (Sangkyen)	April 13-15
ပွꩦ်ဝၣ်ꩦ	khau³ wā¹	Entering the rains retreat	July
ပွꩦ်ゞၵ်ဝၣ်ꩦ	pɔi² auk¹ wā¹	Leaving the rains retreat	October
ပွꩦ်ၵꩵ်ထိꩫ်	pɔi¹ ka¹ thiŋ¹	Presentation of robes to the monks	November
မၢꩦ်ၵꩵ်ꩡꩵ်ၽꩦ်	mai³ kɔ¹ sum⁶ phai²	Burning of wooden pyre	February

Table 1: Tai festivals

In addition to regular festivals, there are also special festivals which may arise from time to time. The most important of these is the festival of ပွꩦ်လꩥ်ꩦ *pɔi² laiŋ⁶*, which is a ritual tug-of-war held to honour the passing of a Buddhist monk. This type of festival is discussed by Terwiel (1995). I have been fortunate to attend and make video recordings of two such festivals, in January 2000 and in February 2002.

1.2 The Tai language family

The *Linguistic Survey of India* (Grierson 1904)[8] contains a very important study of the Tai languages. Grierson (1904:59) includes the Tai languages in a family which he called Siamese-Chinese. Nowadays the thesis that Tai belongs to the same family as Chinese is not accepted by all scholars, rather Tai is grouped together with the Kam[9] languages of China as the Kam-Tai family, and at a higher level is related to the Kadai languages ɛ ⁻ːms a family called Tai-Kadai.

Within the Tai proper, Li Fang Kuei (1977) identified 3 subgroups which he named Northern Tai, Central Tai and Southwestern Tai. The Tai languages of northeast India belong to this Southwestern group. Grierson (1904:59) divided these into two groups:

> Northern - Khamti, Chinese Shan, Burmese Shan & Ahom
>
> Southern - Lao and Siamese[10]

The best known language of Grierson's Northern Group is Shan, and the group as a whole is often now referred to simply as Shan.

1.2.1 The Shan group of languages

Any discussion of the classification of Tai dialects proceeds from an understanding of Tai tonal systems, because it is often the difference in tonal systems that most clearly marks the different dialects/varieties. The Tai tonal system is explained in section 6.1.5.3 below.

Edmondson and Solnit (1997b) discuss the Shan varieties of Burma and China in detail, and, on the basis of tonal distinctions, have proposed two groups: (1997:355)

> Southern group - including the varieties of Shan Panglong, Mae Hong Son and Cushing's Shan

---

[8] The *Linguistic Survey of India* was originally published in several volumes between 1903 and 1928, but the volume relating to the Tai languages was published in 1904, and this is the date by which it will be referred to in this thesis. There was also a reprint of the Survey in 1966-1968, with the volume relating to the Tai languages being reprinted in 1966, and a second reprint in 1977.
[9] For a description of Kam, see Long and Zheng 1998.
[10] Siamese is better known as Standard Thai

Northern group - included Tai Nua and transitional Shan at Mangshi, Namkham and Mu-se.

Edmondson and Solnit provide tone boxes, based on the principles outlined by Gedney 1972 (see 6.1.5.3 below) for each of these two groups. In these tone diagrams, an alphanumeric tonal notation is employed, using a scale of 1 to 5 where 1 is the lowest pitch of the voice and 5 the highest. This differs from the arbitrary notation of tones employed in this thesis (following Banchob 1987), which is discussed in 6.1.5.2.

Merged boxes have been shaded:

	A	B	C	DL	DS
1					
2	325	221	332	31	35
3					
4	44	332	41	33	53

Table 2: Southern Shan tone box (tone values for Panglong)

	A	B	C	DL	DS
1	25				
2	33	2'	31	31	13
3					
4	51	33	52	55	

Table 3: Northern Shan tone box (tone values for Mangshi)

Edmondson and Solnit place Khamti in Northern Shan group (1997:356).

Since Aiton, Khamyang and Phake are spoken still further north, we might have expected them to belong to Edmondson and Solnit's Northern group (Table 3), and indeed the Khamyang tone box as described in this thesis (see 6.4.3) is similar to Table (3) above.

However, the Aiton and Phake tonal systems, particularly the later, do not appear to be related to the system as shown in Table (3). In the Phake tonal system, there is no merge between the boxes A2-3 and B4 that is observed in Table (3). Moreover in both Aiton and Phake there is a merger between A2-3 and A4. The tonal patterns of Phake and Aiton, as reported in this thesis, would appear to require a re-examination of the classification within the Shan group of languages.

1.2.2 The similarity of the languages in the Shan group

Already in the time of Grierson (1904:66), it was clear that the various varieties of the Shan group were closely related. As he stated: "Northern Shān (i.e the language of the Northern Shan states)[11] is closely allied to Southern Shān, indeed they form one language, with only slight difference of dialect. When they differ, Northern Shān is often in agreement with Khāmtī."

Needham (1894:ii) opined that "Khâmtî undoubtedly comes from the same sources as the Shan language treated in the Rev. Dr. Cushing's work, though almost all the words found in use in Khâmtî are quite different from those in use among Dr. Cushing's Shans."

Grierson (1904:66) disputed Needham's opinion that the Khamti words are quite different from the Shan, stating that "To me it seems as if the two languages were almost the same." He states that of the first 20 words in Needham's Khamti vocabulary, "fourteen can at once be found in the same spellings and meanings in Dr. Cushing's Shān Dictionary, and probably more would be found there if allowance were made for difference of orthography."

I have compared the first 20 words in Needham's vocabulary with the *Phake English– Thai Dictionary* (Banchob 1987), and found all but three of the words in that dictionary, the exceptions being:

ꩬꩫ	'about (near in time, &c)'
ꩬ	'abuse'[12]
ꩬꩦꩫ	'accused', *n*.

Table 4: Words recorded by Needham not in Banchob 1987

Grierson further added (1904:78) that "these languages are all closely related to each other. Indeed, they should not be considered as languages, but as members of the group of Northern Shān dialects."

Grierson went on to attempt to classify the Tai varieties. After declaring that the Tairong was the closest to Khamti, he added that:

> "The next nearest is Norā. It uses the Khāmtī alphabet, but has one letter, *q*, which has been lost by Khāmtī, but which existed in Āhom, and still also survives in Shān and Aitoniā. Its vocabulary has more words which are peculiar to Shān than Tairong has, and its grammar often uses both Khāmtī and Shān forms (when they differ) indifferently."

Grierson's classification seems to be based on orthography rather than spoken language, since it is doubtful that Grierson ever heard spoken Tai. With the near extinction of Khamyang (apparently the same as Grierson's Nora) and the fact that all the Tairung or Turung now speak a Tibeto-Burman language, it is now no longer possible to check his claims.

1.2.3 The Tai varieties of Northeast India

Grierson (1904) lists 6 varieties of Tai which have been spoken in Northeast India in historical times, namely: Ahom, Aiton, Khamti, Nora, Phake and Turung. A seventh variety, Khamyang, which may be the same as Nora, is still spoken in a single village in Tinsukia District, Assam.

The living Tai varieties differ in terms of consonant inventories, and more particularly in terms of tone systems. Unfortunately, little or no information about the tonal system of the Ahom and Turung survives, nor about Nora if it were not the same as the Khamyang still spoken in Pawaimukh village.

This thesis concentrates on the Aiton and Phake varieties, but reference will be made to all the other varieties as appropriate and a brief discussion of Khamyang phonology is also presented (see below 6.4). It is common amongst the Aiton and Phake communities to talk of the Tai language, or Tai-Aiton language or Tai Phake language. Since these are named varieties, and there are some features which are accepted by all to be distinguishing features between them, they will be termed as Aiton language and Phake language. When referring to features common to all the Tai varieties, the terms Tai languages of Assam, or simply Tai languages, will be used.

The fact that Aiton and Phake are named varieties does not mean to suggest that they are not mutually intelligible to native speakers. Diller (1992:7 fn6) reported that there was full mutual intelligibility between Aiton and Phake, that they "find each other's speech totally unproblematic." It has however been suggested that some Phake informants sometimes have difficulty understanding the Aitons, although the converse is not the case. This may be due to the reduced number of tone and vowel contrasts in Aiton, discussed in 6.3.2 and 6.3.4 below.

1.3 The data which forms the basis of this thesis

The phonological and grammatical analysis of Tai Aiton and Tai Phake presented in this thesis is based on a corpus of texts, including a word list, collected by the writer during field trips to Assam between 1996 and 2002.

This corpus includes oral literature of a variety of genres, the translation of manuscripts, and some more informal texts, which are discussed in Chapter 9.

All of the texts on which this grammar is based have been translated, analysed and rechecked with native speakers, and these transcriptions, translations and analyses are presented as an electronic appendix to this thesis (see 9.5). The process of collecting and analysing this data is discussed in Chapter 5.

This presentation of the texts generally includes the full text in Tai script, and a translation into English. A line by line analysis with phonemic forms and interlinear gloss will then follow. These texts are important not just for the way that they say things, but also for what they say. Therefore text and translation have been presented to encourage readers to read the whole text. Whilst these texts are useful for mining linguistic data, it is also important to bear in mind that they have deeper meaning for the community from whom they have been collected.

In the body of the analysis itself, particularly chapters 6-8, examples will be presented as in (2), where the text is given in Tai script, followed by a phonemic gloss, an interlinear gloss in English and a translation into English, with notes where necessary. The source and informant for the text will also be given.

2)							
	lik⁴	hau²	lāi²	hau²	lɔi² lɔi²	süŋ⁶	hāi⁶
	book	1Pl	pattern	1Pl	slowly	if	disappear

mɔ⁵	nüŋ⁵	nan⁴	cam⁴	hau²	ta¹	nāi¹
time	one	that	PRT	1Pl	will	regret*

'If our books and writing slowly disappear, then we will regret it.'

(Tai Phake Song, *Song of teaching the Tai language*, Nos. (10)-(11), composed and sung by Ee Kya)

Notes: * synonymous with *yɔn³* 'to regret'.

In the electronic version of this thesis, hyperlinks have been inserted to take the reader to the text from which the example comes, or to the sound file for the example, or to some information about the informant.

Readers of the printed version of this thesis will be able to access this information in one of two methods. The first method, which is somewhat simpler, would be to open the electronic version of the thesis at the same point in the text and then use the electronic links.

Alternatively, the reader can find the files which have been hyperlinked to the example. All of the texts are stored according to the arbitrary number given to the text. This number can be found in section 9.5 below. In the case of example (2), the number of the *Song of teaching the Tai language* is 1-7-2-3. Being a Phake Text, it will be found under \texts\phake on the accompanying compact disc as 1-7-2-3.doc.

The sound file for the whole text is 1-7-2-3.mp3. Being a Phake text, it will be found under \wave files\phake\texts. Aiton texts would be found under \wave files\aiton\texts. The particular sentence example given in (2) would be found by going to \wave files\phake\extracts\1-7-2-3 and then opening 10-11.mp3, because it is lines (10) and (11) from the *Song of teaching the Tai language*.

Finally, it is possible to find some brief biographical information and a photograph of the informant. These are found in the \biographies folder. In the case of example (2), open that folder and open Ee Kya.doc.

As discussed above, although in the past it was sometimes difficult to present texts in languages using scripts other than the Roman, now it is both possible and practical to do so. Several fonts have been produced as part of the present project. The Tai scripts are discussed in Chapter 7 below, and the creation of the fonts is discussed in 7.8 and Chapter 11.

These Tai fonts have become accepted by the Tai Aitons and Tai Phakes and in order to make at least parts of this thesis accessible to those communities, I regard it as absolutely essential that the Tai script is used in every example. (see 4.5 below).

In the electronic versions of this thesis, all sentence examples will be linked to a wave file of that sentence, as is shown in (2) above.

In the past it was not possible to do this, and linguists have grown to accept that when examining data from other languages it is not possible, and therefore not necessary, to actually listen to the linguistic data. Now it is possible, and indeed optimal, to be able listen to the data as it is being examined. This is particularly true for any examination of phonetic, phonological or prosodic detail.

By listening to the data as it presented, a deeper understanding of the richness of a language will be gained; linguistic features which cannot necessarily be put into words, or communicated by a transcription, will be grasped.

1.4 The basic thesis of this work

1.4.1 Key aims of the research

Much of the linguistics in the past century has concentrated on finding formal theories to explain the way that language works, and from that to tell us about human thought processes.

This work, on the other hand, aims to provide a documentation and grammatical description, allowing specialist scholars to use the data from these languages for more detailed analysis of any particular phenomena. Equally importantly this thesis hopes to help preserve the knowledge and literature of the Tai of Assam. The theoretical underpinning of this thesis is discussed in Chapter 4.

The aims of this thesis are summed up as:

- To write a descriptive grammar of the Tai languages of Assam, drawn both from manuscripts and tape recordings collected.
- To produce a lexical data base for these languages, drawing on previous research and my own data.
- To record stories, songs, and other traditional information of the Tais of Assam, to form a basis for linguistic analysis.
- To present the data and findings in an integrated electronic format.

The bulk of the analysis of the language is presented in Chapter 6 *Phonology*, Chapter 7 *Writing* and Chapter 8 *Syntax*. In each of these chapters there are sections relating to theoretical considerations.

2. The Tais of Assam & their languages

Tai people have been living in Northeast India for at least 700 years[13]. Today there remain at least 7 groups which identify themselves as Tai. Of these the Tai Aiton, Tai Khamti and Tai Phake still maintain their traditional languages. However the numbers of speakers are relatively small, possibly as few as 10,000 in India. Due to the small numbers of speakers, these languages can be regarded as endangered.

This thesis has been written not only because of the importance of documenting this variety of human linguistic experience for future generations, but also because of the enthusiasm of the Tai people that their language should be adequately documented, studied, analysed and published.

The Tai groups all use traditional writing systems, which they probably brought with them from what is now Burma. The ability to read and write the traditional Tai writing system, discussed in Chapter 7, is certainly more endangered than the spoken language.

The writing system makes the interpretation of written texts very difficult (see 9.1). It is possible to know the letters of the alphabet without being able to understand the old manuscripts. The task of recording these languages is thus more urgent than it may otherwise appear, because the number of Tai speakers who can interpret the old manuscripts is decreasing. There would be a loss of a considerable part of the Tai linguistic heritage if the growing generation were not able to interpret these old manuscripts.

Before turning to examine the Tai communities about whom this thesis is principally written, it will first be necessary to place these languages in the context of the linguistic mix which is Northeast India, and secondly to refer to the Tai Ahom, rulers

---

[13] If we accept the traditional date for the arrival of the Tai Ahom king Sukapha, Tai people have been in India since 1228.

of Assam for 600 years, whose hopes to recover their Tai heritage prompted my initial research interest in visiting Assam.

2.1 The Linguistic mix in Northeast India.

Four large language families are represented in India, the Indo-Iranian branch of Indo-European, the Dravidian, the Tibeto-Burman and the Austroasiatic. Of these, all but the Dravidian are present in the seven states of the Northeast, known as the *Seven Sisters*. Four of these states, Arunachal Pradesh, Manipur, Mizoram and Nagaland have large or majority populations of Tibeto-Burman speakers: while a fifth, Meghalaya, has a large population of people who speak Khasi, which is a Mon-Khmer language, one of the branches of the Austroasiatic family.

In Assam itself, there are a large number of indigenous Tibeto-Burman languages, some of which have very large populations. There are also a significant number of tribal people from the other states who have come into Assam in more recent times. In addition, there are a large number of Tea Garden labourers, descendents of mostly Munda (Austroasiatic) speakers who were kidnapped by the British in the 1860s. Taken from their homes in eastern India, they were forced to work on Tea Plantations. They now speak a patois of Santhali, Oraon, Munda and Assamese (Hazarika 1994:35).

Most of the people of Assam, however, speak Assamese, which is the most easterly of the Indo-Aryan languages, and was, until recent migrations by Russian and English speakers, the most easterly of any of the Indo-European languages. This part of India is thus the meeting place of these great language families, and it is into this mix that Tai people have arrived in recent times.

2.2 The Ahoms

Historical records show that the Ahom were the first group of Tai in Assam. They were led there by Sukapha (ఐత్రోత్తిం sɣ kaa phaa),[14] a Tai prince from Muang Mau near the region where it is believed the Tai originated, now Yunnan Province in China. Sukhapha lead a group of Tai into the Brahmaputra valley and set up a Kingdom in the country which he named ఐంత్రోంత్తింం *muɲ dun sun kham*, literally 'country-cotton-garden-gold'. The traditional year for this event was 1228. The Ahoms ruled the Brahmaputra valley for nearly 600 years until they were overthrown by the Burmese in 1819, before the Burmese were themselves conquered by the British in 1824 (Gait 1992:344).

From the surviving Ahom manuscripts, it can be seen that the Ahom language was closely related to the language of the Shan states, and even more closely related to the varieties of Tai still spoken in Northeast India, particularly Aiton (see below 6.5.4).

The Ahoms practised their own traditional religion, of which some traces remain in the present day wedding ceremony of the Ahoms (called Soklong), in an annual feast in honour of the ancestors (Me Dam Me Fi[15]) and in certain animal sacrifices. Almost alone amongst the Southwestern Tais, the Ahoms seem never to have followed Theravada Buddhism[16].

After the sixteenth century, it appears that the Ahoms gradually assimilated into the majority Hindu population of their kingdom, adopting Hinduism and many other cultural practises of Northern India. The language used by the Ahom population today is Assamese, an Indo-European language closely related to Bengali.

---

[14] Nothing is known of the Ahom tonal system, and therefore no tones are marked. Nabin Shyam Phalung states that the meaning of this King's name is "tiger-equal to-sky", and is pronounced *suɯ¹ kaa² phaa³* in Aiton.

[15] Me Dam Me Fi is now celebrated as a public holiday on 31st January in Assam. Me Dam Me Fi is Tai language, in Phake it is *me² nam³ me² phĩ⁶* "repair-ancestor spirit-repair-spirit".

[16] There are, however, a number of Buddhist manuscripts in Ahom script in the Department of Historical and Antiquarian Studies in Guwahati. One of these was examined and translated as *Advice to Women* (2-4-2-2). It is not clear what the presence of these manuscripts tells us about the practice of Buddhism under the Ahoms.

Despite the loss of the language in everyday communication, a large number of Ahom manuscripts survive, representing a substantial literature of histories, creation stories, astrologies and other religious texts.

Brown (1837b) was the first to attempt to translate an old Ahom manuscript, using one Juggoram Khargaria Phukan as his informant. These early attempts are discussed by Terwiel (1989). In more recent times, Terwiel and Ranoo (1992), and Ranoo (1996) are two excellent modern editions of very important Ahom texts. Ranoo (1996) is a very scholarly edition of the Ahom Buranji, the history of the Ahoms back to the very beginnings. These and other works on Ahom are briefly discussed in Chapter 3.

Diller (1992:6) has pointed out that "Ahom, as an ethnic term, has two associated but still somewhat distinct senses." The first sense refers to "the former Tai-speaking population who came to rule in the upper Brahmaputra valley", while the second sense refers to the "modern Assamese-speaking subgroup presuming themselves to be the descendents of the former."

In Assam, there is an increasing interest in the Ahom language by the descendents of the Ahom, Diller's second sense of the term. A language revival movement has been underway for several generations, involving the teaching of Ahom in primary schools in those areas where large numbers of Ahom people live, the setting up of a Tai language course at Dibrugarh University, and the publication of short texts including modern songs and poems in the Ahom language. Despite the genuine wide interest in the language, and the clear enthusiasm of many of the revivalists, the writer has rarely met any Ahom who can use the language. One of the leading revivalists, Chow Nogen Hazarika, confirmed to the writer that when Ahom revivalists meet, their conversation goes little beyond greetings and simple statements.

It has been claimed that some members of the priestly caste still have knowledge of the language. Terwiel thoroughly investigated this claim, and following a symposium where several priests were asked to assist with the translation of manuscripts, came to the conclusion that "whereas they could readily decipher the script, ... they did so

without assigning tones ... and without any idea of the meaning of the words, except for a few of the simplest expressions. I reluctantly drew the conclusion that ... Ahom really was a dead language." (1996: 284)

Terwiel went on to claim that

"There are marked differences between the Ahom of the old documents,... and what the revivalists call Ahom, which has totally abandoned the rules of Tai grammar and often uses Ahom words in a non-idiomatic sense." (1996: 284)

One of the examples of what Terwiel calls the Ahom pseudo-language is found in the Ahom wedding ceremony (Chaklong) described by Gogoi (1976). Terwiel says:

"... the sentence khung lu mao cao kao di di si hap ao jao, which is translated 'I respectfully accept your offer'. The most probable dictionary equivalents of these words are khrueang lu 'offering'; mo chao 'priest'; kao 'I', di di 'very good'; sia a suffix; hap 'to accept'; ao 'to take' and jao 'finished'. All these words can be found in Ahom Lexicons (Barua, B.K. and Phukan 1964). To any Tai speaker, however, it is gibberish." (Terwiel 1996: 284)

Terwiel's claims were further discussed by Morey (2002a:103). The Ahom priest, Biswa Sangbun Phukan, who had officiated at the wedding at Moran in December 1999 (see 1.1.1 above) translated the words as follows:

khrung lu "offering"	mao chao "you"
kao "I"	di di "cordially (=good)"
si "PARTICLE"	hap ao "to accept (lit:accept-take)"
yao "PAST TENSE"	

Several Phake speakers indicated that this sentence would be syntactically acceptable in Phake, although pragmatically unacceptable because in Phake *khauŋ⁶* 'thing' could never refer to a person. Furthermore, to be syntactically acceptable the sentence would be spoken as three intonation units, here separated by ().

1) khauŋ⁶ lū¹   maü² cau³ ()   kau¹ nī² nī² sī⁶   ()
    offering   you   I   good PRT

hap⁴ au² yau⁴

hap⁴	au²	yau⁴
accept	take	PAST

"I cordially accept your offering"

From the above, I conclude that at least some of the present day Ahom rituals are based on real Tai texts. The Ahom texts in the Soklong Wedding are divided into two types: responses made by the bride, her father and the bridegroom, of which (1) above is an example, and chanting which is performed by the priests. When I interviewed three senior priests in January 1999, they stated that these texts were sacred, and they were not ready to divulge their meaning.

This raises the issue of the sacredness of the language. Since many of the Ahom texts are religious in nature, and those who possess any vestigial knowledge of the language are mostly priests, the Ahom language itself has developed a sacred place in modern Ahom society. As with many communities that have lost most of their language and culture, the part which remains is held to be increasingly sacred and knowledge of it is often restricted.

It may even be that the Ahom manuscripts were traditionally not available to common people. At least one plausible etymology of the word Buranji, used to describe Ahom histories and extended to all Assamese histories, is that it should be read ᷒᷒᷒ (bao-ran-ci), literally 'never to be shown' (Aimya Khang Gohain 1991:54). In Phake language this would be *mau¹ han⁶ ci⁴* 'NEG-see-show'. Aimya Khang goes on to add "The Tai Royal Chronicles were never to be shown to any unauthorised person."

Yehom Buragohain stated that this etymology was unlikely, given the spelling of the word in the manuscripts was ᷒᷒᷒ *buu rɔn ʒii*, whereas in Ahom sources the word for 'not' was spelled ᷒᷒ *bau* and that for 'see' ᷒᷒ *han*. On the other hand, she had been told by her father, the late Cham Mow Gohain, that the historical books of the Phake (᷒᷒᷒᷒᷒ *lik⁴ khə² mɔŋ²*) should not be shown to any person unauthorised to see them, for fear that they might be misunderstood.

It is always possible that in the 19th century, and even up to the present day, there were people who knew more about the language than they felt able to reveal. Terwiel (1989), on the other hand, discusses the early attempts to translate the Ahom texts for the British scholars in the 19th century. It does appear unlikely that there were any fluent Ahom speakers at that time, otherwise better translations would have been produced.

It is not possible in this thesis to explore all of the issues involved with the Ahom language and its revival. It appears certain that Ahom is not spoken as a mother tongue by any speakers, but equally it appears that the Ahom priests do have some knowledge of their language and the religious texts which they are using do have some roots in the ancient Ahom culture.

It was my original intention and wish to study the revived Ahom. It quickly became clear to me that it would only be possible to interpret the old Ahom texts if I had a very thorough knowledge of the living Tai languages. It is significant that today the scholars in Assam most likely to be relied upon as Ahom authorities, are the two Tai pandits, Nabin Shyam Phalung (Aiton) and Yehom Buragohain (Phake) employed at the Department of Historical and Antiquarian studies. This would suggest that knowledge of the spoken Tai varieties is essential to be able to make sense of the Tai Ahom manuscripts.

2.3 The other Tai groups

Other smaller Tai groups have followed the Ahom into India, namely the Aiton, Khamti, Khamyang, Nora, Phake and Turung or Tairung. It is said that Nora and Khamyang are the same (see Diller 1992:12), although as far as I know the Khamyang speakers in Pawaimukh do not call themselves Nora. Diller (1992:5), also listed seven groups, but stated that the Khamyang and Nora "are perhaps by now one."

There may even have been more groups of Tai in the past than there are at present.
The Aiton manuscript, *The Treaty between the Aitons and the Turungs*, refers to the
Takum and Tarai, who came to live with the Aiton. These may not have originally
been Tai speaking groups. Nabin Shyam stated that he did not know what the origin
of these groups was.

Unlike the Ahom, who today are overwhelmingly Hindu and speak Assamese, the Tai
language survives amongst several of the other Tai groups (see 2.4 below). All of
these Tai speaking groups, together with those who have lost their language in the last
few generations, are now Theravada Buddhists, although some pre-Buddhist practices
still survive, such as calling the *khon* or spirit[17] (see *The book of calling back the
Khon*).

Terwiel (1981) records that in the early part of the twentieth century, animal sacrifice
was still practised by Phakes and Khamyangs. The stricter following of Theravada
Buddhist teaching precludes such ceremonies today. However the practise of calling
the *khon* still persists and has been witnessed by the writer in the Khamyang village of
Pawaimukh.

Other cultural practises also differ between the Ahom and the other Tai groups. For
example, whereas the Ahoms generally live in mud rendered houses built on the
ground, the other Tai generally live in houses built on stilts, often with bamboo walls.
These houses are called ꨥꨓꨲ *hən² hāŋ³* in Phake[18].

Whereas the Ahom dress in a manner similar to other Assamese, the dress of the other
Tai groups[19] is influenced by the Burmese, with men wearing a Burmese-like lungi
(ꨀꨵ *phā³* in Phake) and women distinctive long skirts (ꨕꨮ *sin³* in Phake).

---

[17] The word *khon* cannot be readily translated into English. It is akin to some aspects of 'soul' and
'spirit'.
[18] Photographs of several Tai houses can be found at \biographies\Tai Houses.doc in the electronic
version of this thesis.
[19] For some photographs, see \biographies\Tai Clothing.doc in the electronic version of this thesis.

Most of the Tai speakers live on the south side of the Brahmaputra river. The Aitons,
along with several several Turung and Nora/Khamyang villages are clustered around
the Dhonsiri River. The Phakes, the Khamyang village of Pawaimukh and the Khamti
village of Mounglang are clustered around the Dihing River.

In addition, there is a small cluster of Khamti speaking villages near Narayanpur in
North Lakhimpur District, and a larger group of Khamtis in Arunachal Pradesh.

A Map showing the middle Brahmaputra valley and the location of the various Tai
villages in which I have worked can be found in the electronic version of this thesis at
\photos\maps\Map of Tai Villages.jpg.

2.3.1 Some notes on the history of the Tai

The traditional view is that the Aiton, Khamti, Khamyang, Phake and Turung all
entered Assam between the middle of the 18th century and the early 19th century,
having migrated from Burma and bringing with them Theravada Buddhist religion
and scripts which are closely related to the Shan of Burma.

This view is not uncontroversial. Diller (1992:8-11) discusses the evidence including
the fact that some local accounts indicate that at least some of these groups arrived in
Assam with Sukapha, some five hundred years before the traditional date. Even today,
for example, Aiton villagers can recite by heart stories of the arrival of Sukapha in
Assam, and several ancient manuscripts of the Aiton tell this story. If it is not their
own history, why is it such an important part of their culture?

Some Aiton stories refer to contact with kings of the Tai Ahom. The *History of the
Aiton* (2-2-1-1) No. (51) refers to the Ahom King Suhungmung (ꩬꨮ *su hum
muŋ*) or Dehingia Raja (1497 - 1539). This king conquered the Dhonsiri valley in
which the Aitons now live. If the Aitons were already living in the Dhonsiri valley in
the time of Suhungmung, we would expect them to record this event.

However, as Diller (1992:11) points out, if we accept that the other Tai groups arrived at any date earlier than the 18th century, "their (Burmese-linked) Buddhism and relative lack of Indo-Aryan assimilation would then remain to be explained."

In the case of the Aiton, it may be possible to explain this. On the last morning of fieldwork for this thesis, Ong Cham, an Aiton elder from Bargaon (�’ဘောဂ္ *baan³ luŋ¹*) explained to the writer that until about five generations ago, the Aitons did not follow Theravada Buddhism, but rather worshipped spirits (ဘိဘ *suŋ² phii¹*) and sacrificed animals. He made it clear that the Aitons were only converted to Theravada Buddhism in the mid 19th century.

In the *Story of Barhula*, Ong Thun talks about the visit of the Burmese monk Pi En Duin Sayadaw (*pii¹ en¹ duŋ²*) who converted the Aitons to Buddhism (see *Story of Barhula*, No. (23)). Prior to this, according to Ong Thun, the Aitons had conducted animal sacrifice. (see *Story of Barhula*, No. (24)).

Nabin Shyam later told the writer that whilst it was true that the Aitons had not followed Theravada Buddhism before the visit of Pi En Duin Sayadaw, they were Mahayana Buddhists at that time.

It is also worthwhile to note Grierson's statement on the Aitons, namely that they were

"also called Shām Doāniyās, or Shān interpreters, are said to have been a section of the Shāns at Müng Kâng which supplied eunuchs to the royal seraglio, and to have emigrated to Assam to avoid the punishment to which, for some reason, they had been condemned." (1904:65)

The Aitons that I have interviewed all reject this suggestion by Grierson. However the statement about interpretation may be significant, given that the Aiton language is demonstrably much closer to what the original Ahom seems to have been than any of the other spoken Tai languages are (see below 6.5.4). Perhaps the Aiton speech of today does contain some echoes of the original Ahom speech. If indeed the Aitons have been in the Dhonsiri valley since the 15th century, and if they were not followers

of Theravada Buddhism at that time, then perhaps their language and cultural practises were closer to the Ahoms.

Less historical information has been collected for this thesis in the Phake villages. As mentioned above, Terwiel (1981) reports animal sacrifice having been performed in the past, with some elderly people alive when Terwiel undertook his field work still remembering those ceremonies. The Phakes also have a tradition of being converted to Theravada Buddhism. In 1992 they held a celebration at Borphakial village, to mark the centenary of their conversion.[20]

Of the origin of the Khamyangs, very little has been written. Muhi Chandra Shyam Panjok's pamplet[21] (1981) is the only reference that I have found. He does not cite any sources, but presumably his pamphlet is based on historical Tai manuscripts (*lik⁴ khy² mvŋ²*).

According to Muhi Chandra Shyam Panjok's account, a group of Tai, identified as Khamyangs, were sent into Assam by the Tai King Sukhanpha to search for his brother Sukapha who had founded the Ahom Kingdom in the Brahmaputra valley in 1228. After being welcomed by Sukapha, the Khamyangs set out to return to King Sukhanpha in what is now the Shan States. The Khamyangs then settled at Nang Yang Lake and lived there for about 500 years, before venturing back into Assam after 1780.

Caught up in the various troubles that accompanied the final years of the Ahom Kingdom, the Khamyangs fought both with and against the Ahoms and were divided. One group eventually settled in Dibrugarh district and from there migrated to Pawaimukh in 1922. Chaw Sa Myat Chowlik, born in about 1920, has confirmed to me that the village was founded when he was two, and that his parents brought him there from Dibrugarh district.

---

[20] I am very grateful to Dibya Dhar Shyam for pointing this out to me.
[21] Muhi Chandra Shyam Payok is from Chalapathar and is not a speaker of Tai. It must be presumed that that a Tai speaker, possibly not a Tai Khamyang speaker, conveyed the meaning of the old Tai MS to him.

A second group of Khamyangs settled in Dhali in 1798 and these were the ancestors of the Khamyangs who now live in Golaghat and Jorhat districts, mentioned by Borush (2001) below (see 2.3.4.2).

2.3.2 The Tai Aiton Villages

There are eight Aiton villages in Karbi Anglong and Golaghat Districts. These are listed in Table (1):

Tai name		Translation of Tai name	Assamese/ English name	District
ဗဘၟၬ	baan³ nam³ thum³	Flood village	Duburoni	Golaghat
ဗၑၟ	baan³ sum³	Sour village	Tengani	Golaghat
ဗၟၭၧၯ	baan³ hui¹ luŋ¹	Big fruit village	Eorhola[22]	Golaghat
ဗၟၭၬ	baan³ hin¹	Stone village	Ahomani	Karbi Anglong
ဗၟၬ	baan³ luŋ¹	Big village	Bargaon	Karbi Anglong
ဗၟၩ	baan³ nɔi²/dɔi²	Hill village	Sukhihola	Karbi Anglong
ဗၟၬ	baan³ luŋ¹	Big village[23]	Kalyoni	Karbi Anglong
ဗၟၬ	baan³ saai²	Sand Village	Balipathar	Karbi Anglong
			Jonapathar	Lohit

Table 1: List of Aiton Villages, Tai and Assamese names.

Buragohain (1998:63) listed the names, years of establishment, and population of the various Aiton villages as in Table (2).

[22] Dibya Dhar Shyam believes that Borhola might come from a Tai word meaning 'low land'.
[23] Most Aitons cannot give a Tai name for Kalyoni, a curious thing considering its importance in the history of the Aiton. The name baan³ luŋ¹ was given by Dibya Dhar Shyam in a letter he wrote to me after reading an earlier draft of the thesis.

Village	District	Year of establishment	No. of houses	Population
Ahomani	Karbi Anglong	1939	31	267
Baragaon	Karbi Anglong	1835	39	359
Balipathar	Karbi Anglong	1898	59	528
Chakihola	Karbi Anglong	not known	18	180
Kaliyani	Karbi Anglong	Man era 1239	15	154
Borhola	Golaghat	1836	26	235
Dubarani	Golaghat	not known	43	334
Tengani	Golaghat	not known	19	150
Jonapathar	Lohit, Arunachal Pradesh	1950s	15	148
			260	2155

Table 2: List of Aiton villages after Buragohain (1998)

The three Aiton villages which had already been founded by the 1870s are all found in the 1877 edition of the *Indian Atlas*.[24] However they do not appear with the names that they bear today. Bargaon appears in the Atlas as Mitgaon[25] and Kalyoni as Aitungaon, the latter meaning literally 'Aiton village' in Assamese.

Buragohain's survey does not state the number of people in each village who speak the Tai Aiton language. Although Nabin Shyam stated that all Aitons can speak the Tai language, it seems possible that in some villages there may be a proportion of Buragohain's total who do not know the language. For example, the population of Balipathar consists of both Tai Aiton and Tai Turung. Since the Turung no longer speak Tai, the lingua franca of the village is now Assamese and this is likely to speed the loss of Tai Aiton as a language in that village.

Of these villages, I have visited all but Jonapathar in Arunachal Pradesh. Jonapathar was set up by a group of Aitons who left the crowded Dhonsiri valley in the 1950s to live in an area of Arunachal Pradesh populated by Khamtis. Their children have grown up surrounded by Khamtis and it has been reported to me that they speak a

[24] To view an image of this, go to \photos\maps\Indian Atlas 1873.jpg in the electronic version of this thesis.
[25] Several Aiton informants have told me that Bargaon was once a Karbi Village and that this was its original name. The Knife Pond (nɔŋ¹ mit³) probably gets its name from the old Karbi name.

form of Aiton influenced by Khamti. Nabin Shyam Phalung reported that the language of the people of Jonapathar "is becoming like Khamti."

In February 2002, I was able to interview Sam Khang Chiring, an Aiton from Jonapathar. In an informal discussion, it appeared that his language was indeed influenced by the Khamti tones, but when a recording was taken, the tones and vowels were very similar to those recorded for other Aiton speakers.

2.3.3 The Tai Phake Villages:

Table (3) lists the Phake villages in Dibrugarh and Tinsukia Districts, Assam:

Tai name		Translation of Tai name	Assamese/ English name	District
ဗုၵ်ၚ္ခက္ခော	mān³ phā⁴ kē⁵ taü³	Lower Phake village	Namphakey	Dibrugarh
ဗုၵ်ၚ္ခက္ခော	mān³ phāk⁴ tā⁵	Other side of the river village	Tipam Phake	Dibrugarh
ဗုၵ်ၚ္ခက္ခော	mān³ phā⁴ kē⁵ nə⁶	Upper Phake village	Borphake	Tinsukia
ၚ္ခက္ခော	niŋ¹ kam⁴	Ning kam Nagas	Nigam Phake	Tinsukia
ဗုၵ်ၚ္ခက္	mān³ phā⁴ naiŋ²	Red sky village	Faneng	Tinsukia
ၚ္ခက္	məŋ² lāŋ²	Country of the Lang Nagas	Mounglang	Tinsukia
ၚ္ခက္	məŋ² mɔ¹	Mine village	Man Mau	Tinsukia
ဗုၵ်ၚ္	mān³ loŋ⁶	Big village	Man Long	Tinsukia
ၚ္ခ	nauŋ¹ lai⁶	Nong Lai Nagas	Nonglai	

Table 3: List of Phake Villages, Tai and Assamese names.

The Tai Phake names and their translations were supplied by Sam Thun Wingkyen of Namphakey village. It is interesting that so many of the villages are named after

---

different groups of Nagas (ဝၵ် *khāŋ⁶* in Phake). The term Naga seems to be used to refer to any group of non-Buddhist Tibeto-Burman speakers.

Ngowken Gohain (2001:31) listed the Phake villages, their foundation dates and the foundation date of their respective Buddhist Temples (here spelled Bihar following Assamese practice). This information is listed in Table (4):

	Name	Village	Buddhist Bihar
1.	Borphake	1830	1912[26]
2.	Man mow	1840	1942
3.	Nonglai	1936	1938
4.	Long gaon	1938	1940
5.	Nigam	1938	1940
6.	Faneng	1910	1921
7.	Mounloung	1900	1902
8.	Tipam phake	1850	1850
9.	Namphake	1850	1850

Table 4: List of Tai Phake Villages, after Ngowken Gohain (2001)

Yehom Buragohain (1998:127) also listed the Tai Phake villages, including two more in Arunachal Pradesh:

1.	Man Phake Tau	(Namphake village, Assam)
2.	Man Tipam	(Tipam Phake village, Assam)
3.	Man Phake Neu	(Bor Phake village, Assam)
4.	Man Mo	(Man Mo village, Assam)
5.	Man Phaneng	(Phaneng village, Assam)
6.	Man Long	(Long village, Assam)
7.	Man Nonglai	(Nonglaui village, Assam)
8.	Man Monglang	(Monglang village, Assam)
9.	Man Nigam	(Nigam village, Assam)
10.	Man Wagun	(Wagun village, Arunachal Pradesh)
11.	Man Lung Kung	(Lung Kun village, Arunachal Pradesh)

Table 5: List of Tai Phake Villages, after Yehom Buragohain (1998)

Mounglang is a village which is partly Phake and partly Khamti.

---

[26] Dibya Dhar Shyam of Bargaon, informed me that the temple was founded in 1883, according to a Souvenir published to celebrate the centenary of the Phakes conversion to Theravada Buddhism.

I have not been able to visit either of the Phake villages in Arunachal Pradesh. The names of both villages, Wagun and Lung Kung, are said to be either Naga or Singpho language.

Because I have not been able to undertake field work in most of the Phake villages, I cannot say whether the phonology described for Namphake village (see below 6.2) is also found in these other villages. I have interviewed one Phake from Ningkam village, Ngo Ong, and his language is very similar to that of Namphakey.

2.3.4 Other Tai Speaking Villages

2.3.4.1 Khamti villages

Boruah (2001:41) after listing 30 Khamti villages in Arunachal Pradesh, goes on to list some in Assam "In Assam seven Khamti villages are found in Lakhimpur District and these are - Bor-Khamti, Sribhuyan, Gosainbari, Borpather, Deotula, Tipling and Tunijan. More than one thousand souls are found in these seven villages.

Unfortunately I have not yet had the chance to visit these villages and record any linguistic data there.

As mentioned above in 2.3.2, there is one Khamti Village in Tinsukia District, Mounglang, which I have visited. It is a joint Khamti/Phake village and is listed in Table (3) above. In February 2000, I spent one night in that village and established that the phonology of Tai Khamti spoken there seems to be very close to that described for Khamti by Harris (1976).

2.3.4.2 Tai Khamyang village

There is said to be only one village in which the Khamyang language is still spoken, Pawaimukh in Tinsukia District, which is situated about 7 miles downstream of Margherita. Only older people in Pawaimukh can speak Tai Khamyang, although

some middle aged and younger adults can be described as semi-speakers. The speech of these younger people has yet to be investigated.

Tai name		Translation of Tai name	Assamese/ English name	District
ఐఠ	maan³ paa¹ waai⁶	Village of the Pawai²⁷ river	Pawoimukh	Tinsukia

Table 6: List of Khamyang Villages, (where Tai is spoken) with Tai and Assamese names.

Boruah (2001:41) lists in addition the following Khamyang villages. "In the Chalapathar area of the district of Sibsagar several Khamyang villages are found, Na-Shyam Gaon, Balijan Shyam Gaon and Betbari Shyam Gaon are in Jorhat Sub-division of Jorhat District and Rajapukhuri No. 1 Shyam Gaon is situated at the Golaghat Sub-division of Jorhat district."

2.3.4.3 Turung Villages

Boruah (2001:41) lists the Turung villages as "Pahukatia, Pather Shyam and Tipomiya are in Jorhat Sub-division and No. 1 Veleng Turung village, Naojan Balipather Turung village and Da-Basapather Turung village are situated at Golaghat Sub-division[28]."

In February 2002 I was able to spend a night in Rengmai Turung Village, Sarupathar. I am unsure whether this village is the same as any of those listed by Boruah for Golaghat district.

A number of Aitons have married into this village, and there is at least one Turung, Chaw Pong Chap, who has learned Tai language as a second language learner. The Turung language is briefly discussed below in 2.4.1.

---

[27] According to Chaw Sa Myat, waai⁶ means 'rattan'. This name was given because there are rattan plants growing there.
[28] Boruah 2001 seems to be a reprint of an earlier work; Golaghat sub-division is now know as Golaghat district.

2.3.5 Difficulties with identifying the various Tai groups

Two other names sometimes mentioned in connection with the Tai peoples of India are *Doania* and *Man*. As noted above, (see 2.3.1.1, Grierson (1904:65) commenting on the origin of Aitons, said that they are "also called Shām Doāniyās, or Shān interpreters." According to Nabin Shyam, the word Doania, adapted from the Assamese word meaning 'interpreter' is nowadays used to refer to the Singpho who live in the Margherita division of Tinsukia District. These people, although following Theravada Buddhism, do not speak Tai and are not regarded as Tai either by themselves, or by the other Tai groups.

As to *Man*, it appears that the word is applied to some Tai people because of their origin in Burma (*muuŋ² maan²* in Aiton and *məŋ² mān⁵* in Phake). The Government of Assam and the Indian Constitution both recognise the Tai speakers of Assam as "Man-Tai Speaking". There is a Man-Tai Speaking Association which includes many Tais both from communities that speak Tai and from those that no longer speak Tai. This organisation speaks on behalf of the Tais in the world of politics.

A number of Tai people in India were keen for me to use the term Man–Tai in this thesis, rather than the terms Aiton, Khamyang and Phake. However, the principal informants in each of the villages in which I work prefer the terms Aiton, Khamyang and Phake, and so these are the terms used in this thesis, with the term Tai referring to all these groups together.

The number of people claiming Tai ethnicity in Assam has been estimated as high as several million. By far the largest proportion of people claiming Tai ethnicity in Assam are Ahoms, whose population has been estimated as anywhere between one and four million.

As has been discussed above in 2.2, it seems certain that there have been no native speakers of Tai Ahom for many generations. Who, then, are the Ahoms? There is no

reason to doubt that they are descended from Tai speaking people who arrived in the Brahmaputra valley many hundreds of years ago. However, Nabin Shyam stated that, from his reading of the Ahom *Buranji* (historical chronicles), in all probability the number of Tai Ahoms who originally settled in Assam was quite small and that it was intermarriage with other groups like the Chutiyas and Kacharis that has produced the large population today.

Since very few of the Khamyang, very few of the Turung and almost none of the Ahom have any knowledge of Tai language, the vast majority – at least 99% – of the million or more ethnic Tais in Assam cannot speak Tai language at all.

2.3.6 A note on the names of the Tai groups

2.3.6.1 Ahom

The origin of the name *Ahom*, from which the word Assam comes, is debated. The most widely accepted etymology is that *Ahom* is cognate with the words *Shan* and *Siam*. The current pronunciation of Ahom arises from the fact that in Assamese, /s/ is lenited to /h/ and hence the name of the country and the Ahom people is pronounced [ahɔm]. We may presume that at an earlier stage of history, the name was pronounced [asam].

Grierson (1904:61) quoted at length from Sir George Gait's *Report on the Census of Assam for 1891*, in which Gait discussed the origin of the name *Ahom* as follows:
> "Many different derivations of the name of the province have been suggested, and some of these ignore the undoubted fact stated above, *viz.*, that the country derived its name from the Ahoms and not the Ahoms from the country."

The word *Ahom* is not found in the old Ahom books, where the Ahoms refer to themselves simply as တဲ *tai* or တဲရဝ် *tai rau* 'our Tai'.

Gait, quoted by Grierson (1904:61) also discussed the origin of the word *Ahom*, adding that prior to the advent of the Ahoms, the name was unknown, and that:

"How the name came to be applied to the tribe is still unknown. The explanation usually offered, that they are called 'A-sama' (the Sanskrit word meaning 'peerless') by the Morans and the Borahis, whom they conquered ... is based on the assumption that these tribes had abandoned their own Indo-Chinese dialects more than eight hundred years ago, an assumption which is clearly erroneous."

2.3.6.2 Aiton

Sa Cham discusses the origin of the word Aiton (*aai³ tɔn²*) in his *History of the Tai* No (10). The first element of the name is the word *aai³*'eldest son'.

There is less agreement as to the meaning of the second element of the name. Nabin Shyam stated that *tɔn²* means 'expert', and that in several Aiton books this meaning is to be found. He explained the meaning in (2):

2) ကိတွာ့်   ၊ာ   တိ   ၊လက်   ဖြ   ၊ကာ   ယင်

    aai³ tɔn²   nɛ³   tii²   lik³   phraa²   kɔ³   ʒaŋ²
    Aiton   DEF   at   book   Buddha   LINK   have

    တွာ့်   ၊ယ   ၊ယ   ပ္စ်   ॥

    tɔn²   sɛ¹   mɛ¹   pɛn²
    clever   PRT   COMPARATIVE   other
    '(The word) Aiton is also found in Buddhist books, (where) *tɔn²* means 'cleverer than others.'

In the meaning of 'expert', *tɔn²* is found in the Aiton dictionary, but a cognate word is not found in the *Phake–English–Thai Dictionary* (Banchob 1987). There is a Shan word, �typ်: *tɔn:* 'exceed, extend beyond.' which would be realised in Aiton as *tɔn²*.

As noted above in 2.3.1.1, Grierson (1904:65) mentioned that the Aitons "are said to have been a section of the Shāns at Müng Kâng which supplied eunuchs to the royal seraglio...". The word *tɔn²* 'castrate' is known in Aiton, as well as in both Phake (also

tɔn²) and Shan (ဟyp် *tɔn* 'geld, castrate'). This etymology of the origin of their name is rejected by the Aitons.

A further explanation of the name was offered by Ai Mya, of Bargaon (*baan³ luŋ¹*) which was that the original term was *aai³ thun²*, where *thun²* means 'bright, radiant'.

Sa Cham stated that the King of the Aitons was once called ဝ်ိတွၣတိဇ်ယ်ိ *cau³ phaa³ aai³ tun³*, where *tun³* means 'beginning, origin'. This name, he stated, came about when the royal ancestor of the Tais had five sons, and the first of these, *la¹ aai³* was the ancestor (origin) of the Aitons. Sa Cham went on to say that the present pronunciation, *aai³ tɔn²*, was a Kachari pronunciation of the original *aai³ tun³*.

One further alternative explanation was offered by Dibya Dhar Shyam. He stated that in ancient times large numbers of Tai, 'equal to nine baskets of grain on nine rafts' (see also *History of the Aiton* (2-2-1-1), No (14)), migrated into India. They shattered into many groups, and the group (*tɔn²*) of the elder son (*aai³*) became the Aiton.

2.3.6.3 Khamti

As pronounced by both Aitons (kham² tii²) and Phakes (kham² tĭ⁵), the meaning of Khamti appears to be 'golden place', although syntactically this would be *tii² kham²/tĭ⁵ kham²*. Sa Cham explained that are called Khamti because at some time in the past they came to live at a place called 'River of Gold', as in (3):

3.1) ဝ်   ၊ဒ   ၊ကာ   ၊ယ   တိ   ၊ဒ   ဟ်   ၊ဌ   ၊ယ   ॥

    khau¹   nai³   uu¹   maa²   tii²   nam³   kham²   waa²   saa¹
    3Pl   get   live   come   place   water   gold   say   PRT
    'They lived at the place called 'The river of gold.', it is said.'

3.2) ၊ဒ   ဝ်ိ   ၊ဌ   ဟ်ိတိ   ॥

    lai³   caŋ¹   waa²   kham² tii²
    so   so   say   Khamti
    'So they are called Khamti.'
    (Aiton text, *History of the Tai*, Nos. (47) & (48) told by Sa Cham Thoumoung)

Harris (1976:113) discusses the issue of the meaning of *Khamti* in some detail and offers an etymology similar to that proposed by Sa Cham Thoumoung. He also mentions "a second meaning in the literature for *Khamti* is 'the place where they stuck', which comes from a legend about a Khamti king in Northern Burma who stopped a Tibetan army in a mountain pass to keep them from entering his country."

Harris points out that the current Khamti pronunciation of their name would suggest the meaning 'the place (where) gold (is).'

2.3.6.4 Khamyang

According to Sa Cham, (see *History of the Tai* No. (31)), the name of the Khamyang comes from the fact that at some distant time in the past, they lived for a time at ꩠ꩷꩸ꩲ *nɔŋ¹ ʒaaŋ²* or Yang Pond. This account accords with that given by Muhi Chandra Shyam Panjok in 2.3.1 above.

2.3.6.5 Phake

Sa Cham offered the meaning of Phake as being *phaa¹ kɛ⁴*, 'wall-old', which would be realised in Phake as *phā⁶ kɛ¹* (see *History of the Tai* No. (40)). This is curious since Phake is today realised as *phā⁴ kɛ⁵* by the Phake community.

2.3.6.6 Turung

Grierson (1904:167) states that "The Tairongs (or great Tais) ... are also called Turung or Sham (i.e. Shan) Turung ...". The term 'great Tai' is first encountered in Buchanan (1799), who says of them that they are:

> "... a people called, by the *Burmas*, *Kathee Shawn*, to themselves they assume the name of *Tai-loong*, or great Tai. They are called *Moitay Kabo* by the *Kathee*, or people of *Cussay*. They inhabit the upper part of the *Kiaynduayn* river, and from that west to the *Erawade*."

This would appear to refer to the Khamtis who reside in what is now Burma.

It is possible that the name Turung is not derived from the Tai words ꩠ꩷꩸ *tai² luŋ¹* at all, but is in fact a Singpho word. There are several groups mentioned in *The Treaty between the Aitons and the Turungs*, that is the Takum, Tarung and Tarai. As with Turung, they seem to be pronounced with an initial syllable of [tə]. Perhaps these were names of smaller groups of Tibeto-Burman speakers, and the name has been re-interpreted as a Tai word, because of its chance similarity to the words for 'great Tai.'

2.4 Current linguistic situation

Of the various Tai languages which have been spoken in Assam, the situation of the Ahom is the hardest to categorise. There are no Ahoms who use Tai language as a mother tongue, as discussed above in 2.2.

Each of Aiton, Khamti and Phake is still spoken by several thousand people, and each is still taught to children. They are thus not under immediate threat of extinction, although those who have a deep understanding of the traditional literature of the community are often rather elderly and there is therefore a threat to the range of domains of use of the languages.

Despite being small and endangered languages, their use is expanding into at least one community of non-Tai speakers. Small numbers of ex-Tea Garden labourers now work as labourers for the Tai and live in their villages. Some of them have learned Tai, and some of their children are functionally multilingual in Tai and their own language, and no doubt in Assamese as well.

Khamyang is spoken only by around 50 older adults in the single village of Pawaimukh, near Margherita in Tinsukia district. As mentioned above in 2.3.4.2, some younger adults can be regarded as semi-speakers. Only two elderly men are able to read Tai script.

Both the Nora (if it is different from Khamyang) and Turung are extinct. All Nora

people now speak Assamese, whereas the Turung now speak a language which is

largely Singpho (Tibeto-Burman), and is discussed briefly below in 2.4.1.

All the Tais are multilingual, with Assamese as a second language in every

community, and as the language of Primary Education. There is almost universal

literacy in Assamese.

The following tables list each of these languages, with the most important references

to them and some information about the manuscript collections and speaker numbers

Language:	Ahom
References:	Brown (1837), Campbell (1874), Grierson (1903), Grierson (1904), Barua, G.C. (1920), Barua, G.C. (1930), Barua, G.K. (1936), Barua, B.K. & Phukan (1964), Phukan (1966), Terwiel and Ranoo (1992), Ranoo (1996), Terwiel (1996), Morey (2002a)
Manuscripts:	Large collections, many in private hands
Speaker Status:	No native speakers; some knowledge of manuscripts by Ahom priests.

Language:	Aiton
References:	Campbell (1874), Grierson (1904), Phukan (1966), Banchob (1977), Diller (1992), Morey (1999a), Morey (1999b)
Manuscripts:	There are collections of manuscripts in every village. Most are held in the Buddhist temple, but a few are in private hands.
Speaker Status:	Probably up to 2000 speakers. The language is still learned by children.

Language:	Khamti
References.	Brown (1837), Robinson (1849), Campbell (1874), Needham (1894), Grierson (1904), Harris (1976), Weidert (1977), Wilaiwan (1986), Diller (1992), Chaw Khouk Manpoong (1993), Gogoi (1993)
Manuscripts:	There are known to be significant manuscript collections in the Khamti villages in North Lakhimpur (Gogoi 1993). The Khamti village of Mounglang possesses a large collection, and there are no doubt many more in Arunachal Pradesh and in Burma.
Speaker Status:	10,000 +; Perhaps 500-1000 speakers in Assam; there are no doubt a larger number in Arunachal Pradesh and many more again in Burma. Khamti is not an endangered language, although it may be endangered in North Lakhimpur.

Language:	Khamyang
References:	Diller (1992), Morey (2001a)
Manuscripts:	There is a collection of at least a hundred manuscripts in Pawaimukh village temple, and several more in private hands.
Speaker Status:	Less than 50 speakers over the age of 40 in the one village of Pawaimukh.

Language:	Nora
References:	Grierson (1904)
Manuscripts:	Unknown
Speaker Status:	Dead language; Grierson stated that there were about 300 speakers in his time.

Language:	Phake
References:	Wilaiwan (1983), Wilaiwan (1986), Banchob (1987), Gogoi (1987), Diller (1992), Aimya Khang Gohain (1991), Gogoi (1994), Aimya Khang Gohain (1997), Morey (1999c), Aimya Khang Gohain (1999), Morey (2001b) Ngi Khang Gohain & Ai Che Lei Hailung (2001)
Manuscr pts:	Large collections of manuscripts in every village. Those in Namphakey village have been catalogued and the list is published in Gogoi (1994).
Speaker Status:	Perhaps 2000+ speakers. The language is still learned by children.

Language:	Turung or Tairong
References:	Buchanan (1799)?, Grierson (1904), Turung (1986)
Manuscripts:	Unknown - there are said to be manuscripts collections in some Turung villages.
Speaker Status:	As a Tai language, Turung is extinct; even in Grierson's day there were very few speakers. Turungs today speak a Tibeto-Burman language.

2.4.1 A note on the Turung

Very little work has been done on the Turung language. Aroonrat Wichienkeo (1985),

cited in Diller (1992), reported that "Tai names for certain foods, ornaments, items of

apparel, and so forth, were still used or recognized, although other communication

had to proceed 'through the language of gesture'" (Diller 1992:13).

On the basis of my own observations, it appears that Turung, at least that spoken in

Rengmai village, Golaghat district, is closely related to the Singpho and is clearly

Tibeto-Burman. There are a number of Tai words used by the Turung, particularly

religious and kinship terms, mixed in with the Tibeto-Burman base.

Singpho speakers from Margherita have reported to me that the Turung language of Golaghat has a different tonal system from their own variety. At the present time it is only possible to report that the Turung language at Rengmai certainly is tonal, having clearly two and probably three contrastive tones. To hear the Turung language being spoken, refer to the *Story of the old couple without any food*.

In the 1960s, Dr. Banchob Bandhumedha recorded some tapes which are said to be Turung. These recordings are clearly Tai language. Unfortunately it has not been possible to analyse these as part of this present project.

2.4.2 The Khamyang language revitalisation (2002)

In April 2001 I recorded Deben Chowlik discussing *The future of the Khamyang language*. In the course of the discussion, he mentioned that the Khamyang elders had met to discuss the fact that the language was no longer being spoken by children, and expressed the hope that something could be done about it.

By the time I returned to Pawaimukh in March 2002, a language revitalisation project had begun. Each day at around 4pm, the younger children in the village attend the village primary school for instruction in the Tai Khamyang language. I was able to video tape about 15 minutes of one class.

The teacher is Chaw Mihingta Chowlik, the younger brother of Sa Myat Chowlik. At the age of about 73, Mihingta Chowlik has commenced studying the Tai writing system so that he can pass it on to the children.

The lessons consist of practise in writing the characters, practising lists of words for everyday objects (foods, body parts, numbers etc.) and practising short dialogues which Mihingta Chowlik has written for the class.

It seems that all of the children in the village are now able to pronounce at least some words of Tai. Furthermore at least some of them can use simple expressions, like, for example, calling guests to take rice or tea.

3. Previous studies of the Tai languages

3.1 The need for a critical review of sources on the Tai languages of Assam

A significant number of authors have written on the Tai languages of Northeast India over the past three centuries. These previous studies of the Tai languages form three distinct groups:

- 19th century studies by British and other European scholars
- 20th century studies by linguists, either from Thailand or from western countries
- 20th century studies by scholars from Assam, many of whom are from the Ahom community.

These works consist of all types of linguistic analysis, from the highly theoretical (for example Weidert 1977), or purely lexicographical (for example Barua, G.C. 1920) to those that attempt to survey the Tai languages, such as Grierson (1904) and Diller (1992). All of these works are discussed in some detail in 3.2 below.

In addition to works which focus primarily on linguistics, there are a number of works that contain important linguistic information that is nevertheless peripheral to the main aim of the particular source. These are discussed briefly in 3.3 below. For example, Ranoo Wichasin (1996) is an excellent scholarly edition of the Ahom Buranji, carefully analysed and translated into Tai. It is not primarily a linguistic study, but in presenting a huge amount of reliable data on the Tai Ahom language, it would form an important basis of any future comprehensive study of the grammar of Ahom.

There are several reasons for critically reviewing these sources. Firstly many of the earlier sources contain valuable information about earlier phases of the Tai languages, and give a greater time depth to any study of them. Unfortunately, many of these earlier sources do not give reliable transcriptions of vowels and tones, and need to be read in the light of more recent research. Secondly, some of these sources are very difficult to obtain, particularly in India. It is important to present some information

about these sources for scholars who are unable to access the original materials. Thirdly, many of these sources, particularly Banchob (1987) and Diller (1992) have significantly informed this thesis.

Not all of the previous work, unfortunately, is of such a high standard. It is important not only to academic linguists, but also to the Tai community themselves, that the reliability of these sources be discussed in some detail.

3.2 Linguistic sources

3.2.1 Buchanan (1799)

In Grierson's bibliography of each of the Tai languages spoken in the then Indian empire (1904:76), the earliest reference is Buchanan (1799). Grierson noted that Buchanan gave three short vocabularies for "Tai-nay (i.e. Siamese); Tai-yay (apparently Burmese Shan) and Tai-loong (apparently Khamti or Tairong)."

Buchanan's vocabulary (1799:228-229) is the oldest Western source for the Tai languages that has been identified. It lists 50 words in each of these three languages. The list headed Tai-loong is a language closely related to the Tai languages now spoken in India. Several of these are words the initial consonant of which is proto Tai */d/, which is realised as /l/ in modern Shan, but usually realised as /n/ in Khamti and Phake.

As Table (1) indicates, the language could be Khamti, in which these words are usually realised with initial /n/. This is also true for Phake.

English	Buchanan 1799 Tai-loong	Khamti Robinson 1849	Khamti Needham 1894	Khamti Harris 1976	Phake Banchob 1987
moon	noon	Lün	nün or lün	nɤn[1]	nɔn[2]
stars	nau	Náu	nau	naaw[1]	nãu[2]
earth	need	Nin	nin	nin[1]	nin[2]

Table 1: Comparison of several words from Buchanan (1799) Tai-loong with other Khamti sources.

Buchanan's notation for the languages is poor, and the main reason for discussing his work here is that it is the first data about the Tai languages of India in any European source. There are a number of other references listed by Grierson containing word lists of the Tai languages, including Leyden (1808), Hodgson (1850), Hunter (1868), Dalton (1872), Damant (1880) and Gurdon (1895). These will not be discussed further in this thesis.

3.2.2 Brown (1837)

Brown (1837a 17-21) is the first study of a Tai language which attempts to go beyond simply listing words. He compares five branches of "the Tai language", namely Khamti, Shyán, Láos, Siamese and Ahom. For example he states (1837a:18)

"2. The Siamese d is changed by the Shyáns to l, and by the Khamtis to n, but the Ahoms give it its correct pronunciation."

By "correct pronunciation", Brown is presumably referring to the Standard Thai (Siamese) pronunciation, which, at least in the case of this particular phoneme set, is now believed the closest to the proto Southwestern Tai phoneme *ꞌd.

Brown goes on to give a comparative table of the five languages, part of which is presented here as Table (2):

Siamese	Láos	Shyán	Khamti	Ahom	
Dí	dí	lí	ní	dí	good
Deng	deng	leng	neng	deng	red
Doi	doi	loi	noi	doi	a mountain
Dáu	dáu	láu	náu	dáu	a star
Düan	dün	lün	nün	dün	the moon

Table 2: Extract of comparative vocabulary from Brown (1837a:19)

Brown (1837a:plate opp. p19) also gives a table headed "Alphabets of the Tai Languages", in which he compares the Burmese, Ahom, Khamti, Shyan (Shan state), and Laos scripts, the latter of which refers to the Lanna script of Chiang Mai.

Brown also provided William Robinson with the list of Khamti words which Robinson published (1849:342-349). Since these are marked for tone, it may be that it was Brown who did the research which Robinson published in 1849, and was thus the first researcher who was able to notate tones in the Tai language. There is no evidence of tonal marking in Brown's 1837 paper, however.

3.2.3 Robinson (1849)

Robinson's short paper (1849) is of great importance, containing as it does a very important insight: that the Tai languages are tonal and the tones need to be marked.

As Robinson says: (1849:312)

> "By its finely modulated intonations, sounds organically the same are often made to express different ideas. Thus má, for instance (with the rising tone) signifies a *dog*; *m*á (the Italic *m* denoting the falling tone) signifies *to come*; while the same syllable, with an abrupt termination, or a sudden cessation of the voice at the end of it, má, denotes a *horse*."

Elsewhere in the text Robinson notates another tone where italics are used for the vowel rather than consonant, as in P*o* 'father' (1849:344). Robinson does not describe any characteristics of this tone, but Grierson (1904:144) described it as a 'straightforward tone, of an even pitch'.

The importance of this notation of tone cannot be underestimated. Robinson's word list, which he acknowledges was provided by Rev. N. Brown, contains 282 words. The entire list has been examined, and following the methodology devised by Gedney (1972) (see 6.1.5.3) the following tone box has been derived:

	A	B	C	DL	DS
1	1 (rising)	4 (level??)	3 (glottal)		
2	2 (falling)				
3					
4			?? falling, glottalised		

Table 3: Tone Box for Khamti, derived from data in Robinson (1849)

The tone box presented in Table (3) differs quite significantly from that of present day Khamti as reported by Harris (1976) and Chaw Khouk Manpoong (1993), and presented in Table (4):

	A	B	C	DL	DS
1	3		6	3	
2	4				
3					
4	7	4	5	4	

Description of Tones:

Tone number (Harris 1976)	Description (Harris 1976)	Tone number (Chau Khouk Manpoong (1993)	Tone mark (Chau Khouk Manpoong (1993)
4	high level	3	--
1	mid falling	4	o --
2	low falling with glottal constriction	5	ʌ --
5	mid rising with glottal constriction	6	c --
3	high falling	7	ʃ --

Table 4: Tone box for Khamti (Arunachal Pradesh and Assam) after Harris (1976) and Chaw Khouk Manpoong (1993)

It is difficult to see how the present day Khamti tonal system could have developed from that described in Table (3), because, for example, it would require a split in the falling tone of Table (3), (boxes A2-4).

Nevertheless, the importance of Robinson in notating tones cannot be underestimated. In the 150 years that have followed his publication, unfortunately, the notation of tones has not become an obligatory part of language description in this part of India. Tones are not marked by Needham (1894), Grierson (1904), Gogoi (1994) nor by most of the present generation of Assamese scholars writing on Tai languages.

In addition to his vocabulary and notes on the tones, Robinson also provided a brief grammatical sketch, and several sentence examples. These show, for example, that the particle 𝑗𝑥̃ haŋ² 'to' is used to mark an animate non-subject argument of the verb (see discussion at 8.3.3), although Robinson does not fully understand its use, writing simply: "The preposition *Hang* is most commonly used as the particle to denote the dative case." He then gives the following examples:

2	3	1		1	2 3
Hang man haü da.				*Give (it) to him.*	
3		1 2	1	2 3	
Hang man maü po. | | | *You beat him.* | | |

Table 5: Sentence examples from Robinson (1849:313) showing the use of *haŋ²* "to"

Furthermore, although a number of the sentences given by Robinson show the verb in final position, it is not always the case, as in:

1) Sang tüt háng *man*, ta khán.
    If (you) pull her tail (she) will scratch you.

Example (1) can be analysed as (2)[29]:

2) ဃ၀ၓ ၓၞၓၐ ၂ၓ ၓၓ ၒ၁ ၒၓၓ
    saŋ³   tuit⁴   haaŋ³ man⁷ ta³   khan³
    if   pull   tail   3Sg   will scratch
    'If you pull (her) tail, she will scratch you.'

_____

[29] Tones numbered following Chaw Khouk Manpoong (1993).

3.2.4 Campbell (1874)

Campbell (1874:168-181) gives a list of words and phrases in several languages, namely: "Assamese, Ahom, Khampti, Aiton, Cacharee, Hojai, Mikir." The list of English words is virtually the same as that in Grierson (1904:213-233), but the Tai forms given are not the same.

Campbell is the first source to mention the Aiton. However, as with Grierson, his data does not seem to accord with present knowledge of the Aiton variety. For example, in his word lists there is no sign of initial /d-/, which the Aitons pronounce today. It may be that Campbell relied entirely on the written form, which uses the graph for /n/ to write /d/ (see 7.2). Some words recorded by Campbell are presented in Table (6):

English	Ahom	Khampti	Aiton
Sun	Bán	...	Bhan
Moon	Den	...	Neun
Star	Dáu	...	Nao
Good	Di	Nioh	Ni uga

Table 6: Some vocabulary examples from Campbell (1876)

As for the sentences, the following from the Aiton are interesting. They are compared with Grierson's version from 1904:

English	Campbell 1876	Grierson 1904
What is your name?	Chiuma chiuchang	Maü chü säng
My father lives in that small house.	Pukai uhun annai	Pö kau ü ti hün iñ (en)
Give this rupee to him	Inguna ennai hau hang mun	Ngün trä a nän haü häng man

Table 7: Some sentence examples from Campbell (1876)

Campbell's phonetic notation is very poor indeed, especially when compared to the excellent work done by Robinson and Brown nearly fifty years earlier. For example, he could not write initial /ŋ/, so that /ŋun/ comes out as <Inguna>. Nevertheless, the

sentences remain useful and can still be analysed. For example, the first sentence listed in Table (7) can be analysed as (3)

3)   ကၢၟ်   မႂ်   ကၢၟ်   သင်

      cɯ²   maɯ²   cɯ²   saŋ¹
      name   2Sg   name   what?
      'What is your name?'

It is interesting that in Standard Thai this question is expressed as '2Sg-name-what?', which is similar to the form recorded by Grierson in Table 7, whereas among the Tai Aitons today, the form is as in (4)

4)   ကၢၟ်   မႂ်   ကႃသင်

      cɯ²   maɯ²   ka¹   saŋ¹
      name   2Sg   what
      'What is your name?'

The Aiton pandit, Nabin Shyam Phalung stated that the form in (3) would be acceptable to Aiton speakers, just as (4) is, although (4) is much more common. The form given by Grierson for this sentence (the form which has the same syntax as modern Standard Thai), which would be rendered in Aiton as *maɯ² cɯ² saŋ¹*, is not acceptable to the Aitons. Early in my fieldwork for this thesis, I often asked this question in this form, and the speakers did not understand what I was asking.

Even though his notation is poor, it does appear that a comprehensive and careful analysis of Campbell's sentences, and their comparison with modern Aiton practice, would add to our knowledge of these languages. Such a study has been beyond the scope of this present thesis.

3.2.5 Needham (1894)

Needham (1894), at 201 pages, is a very considerably longer and more exhaustive text than any of the earlier works on the Tai languages. The title (*Outline Grammar of the ကၢၟ် (Khâmtî) Language as spoken by the Khâmtîs residing in the neighbourhood of*

Sadiya) makes it clear that this text refers to the Khamti language as spoken in what is now India, although the book itself was published in Rangoon.

In his introduction Needham admits that :"The character is not difficult, but the various tones met with in the language are very puzzling." (1894:ii).

He goes on to explain what he understands by these tones, and to give examples for the graph ကၢၟ် (khai). This is reproduced in Table (8), with the addition of a phonemicisation of the modern Khamti forms, using the tonal number of Chaw Khouk Manpoong (1993).

"By finely modulated intonation sounds organically the same are often made to express totally different ideas; thus ... there are no less than six words written ကၢၟ် (khai), but each one expresses a different meaning according to the tone in which it is uttered, namely, –

      ကၢၟ် = ill   (modern Khamti *khai⁶*)

      ကၢၟ် = sell   (modern Khamti *khaai³*)

      ကၢၟ် = buffalo   (modern Khamti *khaai⁷*)

      ကၢၟ် = egg   (modern Khamti *khai³*)

      ကၢၟ် = go, depart   (modern Khamti *khaai⁵*)

      ကၢၟ် = tell, inform"   (modern Khamti *khai⁴*)

Table 8: Discussion of Khamti tones in Needham (1894)

Needham goes on to explain that there are three tones which are "very marked and not difficult to acquire"(1894:8). These three tones are the very same that Robinson had reported in 1849, and are described in almost identical terms, as for example the third tone:

      "(iii) an abrupt termination of the voice, or a sudden cessation of it at the end of a word, as မႃ (mâ) a horse."

It is also clear that Needham could not distinguish contrast of height in the central vowels, where a contrast between [ɯ] and [ɤ] or [ʉ] and [ɔ] could be expected. He also confuses mid-close and mid-open back vowels, such as his transcription of ကႃ as

[oi], when it clearly should be [ɔi]. See for example 1894:2 ᵛᵒᵈ = (noi) hill, mountain, which should be transcribed as *nɔi⁴*.

However Needham's work remains a very useful study. He gives a comprehensive analysis of the writing system, (1894:1-8), a section on nouns (9-14), Adjectives (14-22), Pronouns (23-30), Adjective Pronouns (30-35), Verbs (35-65), Adverbs (65-80) and ther ʳ ʳᵉ̃ction of Syntax (81-85), then a section giving further sentence examples of the various features.

Finally there is section of Miscellaneous Phrases (107-115), Texts (115-130) and a substantial Vocabulary (130-201).

Needham's rules of syntax (1894:81), here reprinted as Table (9), appear to have been used by scholars over generations to justify the claim that the Knamti language has changed to become an SOV language, probably due to Assamese influence. Khamti is listed as an SOV language in Greenberg (1966). This claim is discussed in detail in 8.4.

a)	The adjective follows the noun it qualifies
b)	Numerals sometimes precede, at others follow, the nouns they qualify. When a numeral particle is used, they invariably *follow* the noun they qualify.
c)	Demonstratice pronouns are generally placed immediately *after* the nouns they qualify, except in relative clauses when they may *either* be put at the *beginning* or at the *end* of the sentence.
d)	Relative clauses should stand *before* antecedent clauses
e)	Interrogative pronouns rarely ever stand *first* in a sentence.
f)	Adverbs are *generally* placed *after* a verb.
g)	Prepositions precede their nouns
h)	The order of words in a predicative sentence are (1) subject, (2) direct object (3) indirect object, and (4) verb.
	In an interrogative sentence the order is (1) subject, (2) indirect object (3) direct object, and (4) verb.

Table 9: Needham's *General Rules of Syntax* (1894:81)

Needham's work is referred to frequently in this thesis, especially in the sections of Constituent Order (8.4) and the marking of Tense and Aspect (8.5.7). The greatest value of Needham is not so much in his analysis, but in the mass of data which he

provides; it is only to be regretted that he could not hear the tones and did not name his sources. Nevertheless it remains the most comprehensive discussion in English of the Khamti language, a variety very closely related to the languages examined in this work.

3.2.6 Grierson (1904)

Grierson is the first scholar to discuss all of the Tai languages of India; and his gigantic seminal werk, *The Linguistic Survey of India*, has been the first port of call for any researcher in this field for a hundred years, as it was mine.

3.2.6.1 Contents of Grierson (1904)

Table (10) lists the full contents of Grierson

p56	Siamese Chinese Family
p58	Map
p59-79	Tai Group
p81-105	Ahom - Grammar
p106-119	Ahom - Texts
p120-140	Ahom - Word List
p141-149	Khamti - Grammar
p150-165	Khamti - Texts
p167-169	Tairong - Grammar
p170-177	Tairong - Texts
p179-181	Nora - Grammar
p182-191	Nora - Texts
p193-195	Aitonia - Grammar
p196-211	Aitonia - Texts
p213	Phakial
p214-233	Comparative Word List
p226-233	(Comparative Sentences)

Table 10: Contents of Grierson (1904)

3.2.6.2 General Comments on the Tai Group (1904:59-79)

Grierson's classification of the Tai languages has been mentioned in 1.2 above. In addition, he proposed a classification of the Tai writing systems (1904:59) which also accords with modern knowledge noting that they have "seven distinct forms of the written Character - the Āhom, the Cis-Salwin Shān, the Khāmti, and the Tai Mau (Chinese Shān), the Lü and Khün (trans-Salwin Shān), the Lao and the Siamese."

He also discusses the grammatical structure of the various Tai languages in general terms. Some of these comments will be revisited when examining these languages further in this thesis.

As Terwiel (1989:133) points out, Grierson's analysis for all of the many languages and varieties that he looked at, was based on three types of texts, firstly a translation of *The Parable of the Prodigal Son*, secondly another piece of text in the vernacular, such as a piece of folklore or narrative prose, and thirdly a word list based on the standard list drawn up by Campbell. Unfortunately, for the Tai languages, the second text was missing, except for the short Nora text, containing riddles which belongs to the second category.

Although we know, for example, that the Deputy Commissioner of Sibsagar provided the Nora texts (see below 3.2.6.6.1), in general the sources of Grierson's work are not known and the native speakers who provided information are never named.

Grierson appears to have analysed and glossed the texts and written his comments on the basis of those texts. He did not have access to any information about the tones of the languages, and, at least in the case of the Aiton text, his analysis seems to be based on transliteration rather than a phonetic realisation of the text.

In this brief survey, I will examine the first two lines of each of his translations of *The Parable of the Prodigal Son* (except for that in Ahom), and make any other pertinent observations.

3.2.6.3 Ahom (1904: 81-140)

Terwiel (1989:133-141) discusses the Ahom material in Grierson in considerable detail, listing the sources from which Grierson probably got his information. In this present study, only his brief discussion of the tones of Ahom will be mentioned.

3.2.6.3.1 Tones

Grierson makes one important point about the Ahom tones, namely that there should be tones in Ahom, but noted that tradition is silent on the matter. He states

> "Moreover, in the one word, the tones of which I have been able to ascertain, they differ from those of the Khāmtī and Shān. This is the word *mā̆*, which, when it means 'a horse', has in Āhom a long tone, and in Khāmtī and abrupt tone, while *mā̆*, 'a dog', has in Āhom an abrupt tone, but in Khāmtī and Shān a rising inflection." (1904:90)

It is not known where this information about the tones has come from, but if it is true, it would be very unusual. The abrupt tone referred to is clearly a glottal tone, but it is reflexes of the proto Tai C tone (see below 6.1.5.3) which are usually glottal, not the A tone. Thus we would expect the word for 'horse' (which is C4) to be glottal, not the word for 'dog' (which is A1).

3.2.6.4 Khamti (1902:140-165)

3.2.6.4.1 Tones

Grierson provides very little information about the tones of any of the Tai varieties, and almost everything that he writes about the Khamti tones is based on Robinson (see 3.2.3 above. He does, however, give information about the tone which Robinson did not describe, as follows:

> "(2) The straightforward tone, of an even pitch. Robinson does not mention this tone, but in a number of words, nearly all of which have this tone in Shān, he puts the *vowel* of the word into special type. This *po*, 'a father'. As Robinson makes no other provision for this tone, it appears that he intended to indicate it by this typographical device, but omitted to draw attention to it.

This tone corresponds to Dr. Cushing's third or 'straightforward' tone in Shān."

(Grierson 1904:144)

3.2.6.4.2 Texts

Unfortunately Grierson did not record a genuine Khamti text. Example (5) is his analysis of the first two lines of the *The Parable of the Prodigal Son*. The presentation of the texts here differs from Grierson's original, in which the Tai script was on one page, and the transliteration with gloss on the opposite page.

5.1)

Mü-nan	kōn	kō-lüng	yang	lūk-chai	shâng-kō
Formerly	*man's*	*a*	*were*	*sons*	*two*

5.2)

Naü	lūk	man	shâng-kō	nai
Amongst	*children*	*his*	*the-two*	*aforesaid*

lūk-chai	ân-nai	pō	man	mai	wā-kā
child (or son)	*younger-the*	*father*	*his*	*to*	*said*

khüng	chā	kau	chât-khă-laü-ū-ko
property	*share*	*my*	*whatever (there be)*

pan-haü-tā.'.
divide-give

Notes:

- The spelling of ꧋ 'CLF', suggesting [ko], might indicate an /o/~/ɔ/ merge. The word is *kɔ⁴* in Phake.

3.2.6.5 Tairong (1904:167-177)

Unlike the Khamti, Grierson's work on Tairong is the only description of this language which is based on information from apparently native speakers. There is however some evidence that the translator of the *The Parable of the Prodigal Son* was not a fluent speaker of Tai. The first sentence of the text is unusual in several aspects, in particular in that the human classifier *kɔ* is not found, and that the noun *luk* 'child' is represented as a count noun, in which form it is not found in any of the spoken Tai languages today.

3.2.6.5.1 Texts

6.1)

Kun	fū-lüng	yang	sâng	luk
Man	*one*	*had*	*two*	*sons*

6.2)

Luk-mwān	koi-nai	wā-kā
son-his	*younger*	*said*

'pō	ēi	khüng	maü	ok-chā	khā-laü
father	*O.*	*property*	*your*	*property*	*how much*

tī-fât-kwō	hāng-kau	haü-mā.'.
portion	*to-me*	*give*

Notes:

- *phuu* is known as a classifier from several Aiton manuscripts.
- The verb *yaŋ* 'have' is most often in utterance final position in the other Tai languages
- The word *mwan* '3Sg' has a very unusual spelling in line 2
- *kaa* 'TAM marker' appears to be attached to *waa* 'say' in line 2.

. 2.6.6 Nora (1904:179-191)

As discussed above in 2.3, the word Nora is sometimes understood to refer to the Tai Khamyang. In comparing Grierson's Nora with the Khamyang which I have been able to record, it is interesting to note that Grierson stated that "the letter ဝ is sometimes *ba*, but more usually *wa*." (1904:179). The same process, by which words spelled with initial <w> are pronounced with initial /b/ is found in the speech of some present day Khamyangs (see below 6.4.1).

On the other hand, Grierson stated that the particle *mai* is "regularly used to make a kind of oblique form when a noun is governed by a preposition. Thus ... *tīFrā-mai* to (i.e. against) God" (1904:180). No sign of this postpositional *mai* has been found in my research on the Khamyang.

3.2.6.6.1 Texts

3.2.6.6.2 From the Parable of the Prodigal Son:

7.1) ကၞ်   ကၞႝ   လုၵ်   ယၞင်ေကၟ   ဝၚ် ॥

Kōn	kō-lüng	luk	sâng-kō	jāng
Man	*person-one*	*sons*	*two-persons*	*had.*

7.2) လုၵ်   ပၞးကၞ်   ၒ   �064ၚ်   ေၮၟ

Luk	pạ-ân	wā	hāng	pō
son	*male-younger*	*said*	*to*	*father*

'pō	hai	ạ-muñ (muy)	maü	kā	kau
father	*O,*	*good*	*your*	*to*	*me*

tak	nai-nai	haü	hāng	kau.'
will	*be-given*	*give*	*to*	*me*

Notes:
- It is interesting that the pronunciation of ဝၚ် yang 'have' is (presumably) [ᵈʒaŋ], as it is among the Aitons but not in the Khamyang data that I have collected.

3.2.6.6.3 Riddles

The only native Tai texts recorded by Grierson were several riddles, one of which I have analysed below. Example (8) is the riddle as Grierson presented it (with the Tai script together with the transliteration) and (9) is my analysis of the same, using a transcription based on the brief study of Khamyang phonology as described in 6.4 below.

8).   တꞗ်မꞗ်ကꞧလꞧမꞧ

Ton	man	kā	lam	met
Tree	*its*	*is-equal*	*rod*	*fishing*

ꞧꞧမꞧꞧꞧꞧꞧꞧ

Nok	pit	nân	kū	khā
Bird	*Tuni*	*sleeps*	*every*	*branch*

Its tree is equal to a fishing rod and the Tuni-bird sleeps on every branch

Answer: Ton-mặk-khū, the Brinjal tree

9)   တꞧ်   မꞧ်   ကꞧ   လꞧ   မꞧ်

tun³	man²	kaa⁵	lam²	met¹
tree	it	equal	rod	fishing hook

ꞧꞧꞧ   မꞧ   ꞧꞧꞧ   ကꞧ   ꞧꞧ ॥

nok⁴	pit¹	nɔn²	kuu⁵	khaa⁶
bird	"Tuni"	sleep	every	branch

တꞧ်   မꞧ်ꞧꞧ ॥

tun³	maak¹ khʏ⁶
tree	brinjal / egg plant

Notes:
- nok⁴ pit¹ in Phake refers to "any kind of small birds"

3.2.6.7 Aitonia (1904:193-211)

Grierson noted that there were two groups of Aitons - in the Naga Hills and in Sibsagar District, and that there were 163 of them in the 1891 census. It is possible that this refers to the same groups and the same villages as are still found, with the

modern day Karbi-Anglong district forming part of the former Naga Hills district and Golaghat part of the then Sibsagar district (see above 2.3.2)

No information is given about the Aiton informants, except that "The specimens which I have received from the local authorities of Sibsagar are evidently carefully prepared, and it has been easy to make out the meaning of the greater part of them." (1904:193)

3.2.6.7.1 Orthography

Grierson's comments on the orthography are curious indeed, because they seem to refer to another language altogether:

> "As regards consonants, we have the Khãmtī ကa *ka*, instead of the Shãn ꩬ, and the Khãmtī ဖ *pha*, instead of the Shãn ꩬ. In every other case, when the Shãn form differs from the Khãmtī one, the former is used. Thus we have Shãn ꩬ instead of the Khãmtī ꩬ for *sa*, and the Shãn ꩬ instead of the Khãmtī ꩬ for *na*." (1904:193)

There is no sign in the early Aiton manuscripts, such as *The Creation of the World*, of the Shan graphs for /na/ or /sa/. This manuscript is at least a century older than Grierson, and it is not clear how Grierson came to this view. Nabin Shyam stated that the Shan form for /sa/ was in current Aiton usage, although this may be because Shan /sa/ is the same as Burmese /sa/ and Burmese script is used by the Aitons for Pali and Burmese texts. An examination of the script of the *Parable of the Prodigal Son* suggests that some of the forms are intermediate between Shan and Khamti forms.

3.2.6.7.2 The texts

The first two lines of the Prodigal son were given by Grierson as follows[30]:

---

[30] Here reprinted with modern Aiton orthography.

10.1) ကုꩬ ကော်လုꩬ လုꩬ ဆင်ကော ယင် ။

Kun	kō-lüng	luk	sâng-kō	yang
Man	*person-one*	*sons*	*two-persons*	*had.*

10.2) လုꩬ ဆဲ အꩬ လောဟင် ပောမꩬ

Luk-	chai	ân	lau hãng	pō-man
son-	*male*	*younger*	*said to*	*father*

ပော ခင် မꩬ ယင်ဆင် ဝင်ဟꩬ.

'pō	khâng	maü	yāng-sāng	weng-haü'.
father	*property*	*thy*	*whatever*	*divide-give*

Notes:
- Nabin Shyam Phalung suggested that the first sentence would be better written as :

  ကုꩬ ကော လုင် လုꩬ မꩬ ၂ ကော ဝင်

kun[2]	kɔ[3]	lɯŋ[2]	luk[3]	mɯm[2]	sɔŋ[1]	kɔ[3]	ʒaŋ[2]
person	CLF	one	child	3Sg	two	CLF	have

- Nabin Shyam commented that the form ဝင် is not found for ဝင် *bɛŋ[1]/mɛŋ[1]* 'divide'. It is found in Shan as ဝင်, *wɛŋ*,.

3.2.6.8 Comparative Word list and Sentences (1904:214-233)

In addition to the translation of *The Parable of the Prodigal Son*, Grierson also gives forms in each of his target languages for the word list first used by Campbell. This list includes some sentences which have already been discussed briefly above in 3.2.4.

3.2.6.8.1 Vocabulary items

Grierson's substantial comparative word list – none more substantial has been published in the century since – contains some interesting features. It appears that his transcriptions of the Aitonia forms are based on the written rather than the spoken form, because they show initial /n/ in words such as 'sun', 'star', following the written orthography, rather than the initial /d/ which is found in contemporary Aiton speech for these words. This is illustrated in Table (11):

No.	English	Āhom	Khāmtī	Tairong	Norā	Aitoniā
52	Woman	Nūng, kūn-mī	Pā-ying	Kun pā-ying	Pạ-jing	Pạ-ying
62	Sun	Bān	Wan	Wān	Ban, khun-ban	Wān
63	Moon	Dūn	Nün or lün	Nün	Nün	Nün
64	Star	Dau	Nau	Nau	Nau	Nau
132	Good	Dị	Nī	Nī	Nī	Nī
163	Be	Nāng or ū	Chaü	Chau	Pin (pen)	Chaü

Table 11: Extract from Grierson (1904) Comparative Vocabulary

3.2.6.8.2 Tense

Grierson attempted to provide a full tense paradigm for several verbs. In sentences No 180-204 he attempts to do this for *po* (Aiton *pɔ³*) "to hit"

No.	English	Āhom	Khāmtī
202	I am beaten	Kau-mai-po-ū (mai *is used in the passive voice when the agent itself is an object*)	*Cannot be given*
203	I was beaten	Kau-mai-po-jau	*Ditto*
204	I shall be beaten	Kau-mai tī-po- ū	*Ditto*

No.	Tairong	Norā	Aitoniā
202	Hāng kau pō-yo	Kau kin khân (*I eat stripes*)	Pō hāng-kau
203	Pō kau	Kau kin khân kā	Hāng-kau pō-kwā
204	Kau-mai tī-pō	Kau tak kin khân	Tạ-pō hāng-kau

Table 12: Extract of Grierson's elicitation of 'tense'

Nabin Shyam confirmed that each of the forms listed by Grierson was found in contemporary spoken Aiton. They would be realised in modern script as follows:

11.1) ᨣᩮᩢ ᨦ ᨣ (Grierson's No. 202)
pɔ³ haŋ² kau²
beat to 1Sg

11.1) ᨦ ᨣ ᨣᩮᩢ ᨣᩤ (Grierson's No. 203)
haŋ² kau² pɔ³ kaa¹
to 1Sg beat GO

11.1) ᨲ ᨣᩮᩢ ᨦ ᨣ (Grierson's No. 202)
ta¹ pɔ³ haŋ² kau²
WILL beat to 1Sg

3.2.7 Barua, G.C. (1920)

As discussed by Terwiel (1989), G.C. Barua was the person who provided Grierson with his data on Ahom. Terwiel notes that Barua was "an Assamese speaker who in the mid-1890s had been appointed to learn Ahom from a committee of five Deodhai priests and to translate Ahom documents." (1989:133)

Barua states his work is based on "the collection of words from Ahom manuscripts", a task performed by an unnamed clerk, and the "further collection of words in consulation with a few Deodhais who could read Ahom and perform *Deopujas*."

It appears that the dictionary, like the *Ahom Lexicons* (see below 3.2.9) contain words which cannot be confirmed in other sources.

3.2.8 Barua, G.K. (1936) (*Ahom Primer*)

Terwiel (1989:143) writes of the *Ahom Primer* (Barua, G.K. 1936) that:

"It represented a whole new departure, and, refreshingly, provided a critical reordering of some sentences, in which the influence of at least some authentic Ahom texts is apparent. However, since the main text of the *Ahom Primer* is in the Assamese language and script, it has not had any noticeable influence upon scholars outside Assam.'

The *Ahom Primer* contains information which suggests that speakers of one of the spoken Tai varieties were consulted in its preparation. The text contains a word list and some extracts from the Ahom Buranjis, but it also contains some short stories which are similar in form to some of the short stories gathered for this study.

It also contains some short sentences such as (12):

12)　ti　pai　le　ti　kat
WILL　go　roam　to　market
'Will you go to the market.'

Chaw En Lai Phalung has informed me that when his father, the late Ai Ne Phalung (Mohenda Phalung) was a student, he lived with G. K. Barua in Jorhat. Apparently G. K. Barua would ask Ai Ne Phalung about the Tai language, and used much of the knowledge he gained thereby to produce the *Ahom Primer*.

This claim is backed up by the fact that there is some evidence of peculiarly Aiton forms in the *Ahom Primer*, such as the use of ka in words like ka nü 'above', which in Aiton would be *ka¹ nɯ¹*, a shortened form for *kan² nɯ¹* (literally 'direction-above').

Significantly, the Ahom letter w is rendered in Assamese as জ which would suggest a pronunciation of [dʒ], which is found in Aiton (see below 6.3.1.1.6), rather than using য which would have suggested [y]. This is consistent with G.K. Barua's informant having been an Aiton.

Unfortunately Barua did not credit Ai Ne Phalung in the published book. It was normal practice at that time not to name native speaker informants in works of this type. We are fortunate that it has been possible to establish who provided at least some of the information in the *Ahom Primer*.

3.2.9 Barua, B.K. and Phukan (1964) (*Ahom Lexicons*)

Barua, B.K. and Phukan (1964) is an important text for the Tai languages of Northeast India, in that it claims to be an edition of the *Bar Amra*, an 18th century Ahom word list now in the possession of the Department of Historical and Antiquarian Studies, Guwahati (MS 31). However, there are key differences between Barua and Phukan and the *Bar Amra*.

For a start, Barua and Phukan begin their text with the letter (a), whereas the manuscript commences with the letter (ka). The first two lines of the manuscript, which give all the meanings of the Ahom word (kā) is reprinted below as (13):

13)

A translation of (13) is presented below in Table (13). From this it can be seen how the Bar Amra is arranged. Written entirely in Ahom script, the Ahom word under consideration is presented first, followed by the Ahom words *cam ba* "means" and the Assamese gloss written in Ahom script.

Bar Amra	Transcription	Translation
	ka cam ba kaurik	Ka means a crow.
	ka cam ba dʒabɔr	Ka means to go.
	ka cam ba bɛhabɔr	Ka means to trade.
	ka cam ba dʒukhak	Ka means measure
	ka cam ba kɔpalik	Ka means a forehead ornament.
	ka cam ba kɔstɔk	Ka means hardship.
	ka cam ba ɛral dakurɔk	Ka means a tethering block.
	ka cam ba bhar khuak	Ka means to pay a dowry
	ka cam ba nasibɔr	Ka means to dance.

মৰৈৎৱৰিৰূৎশিৰ্শ ॥	ka cam ba kɔthiak	Ka means paddy seedlings.

Table 13: Translation of the first two lines of the Bar Amra.

The entry for ᴎᴋ as given by Barua,B.K. and Phukan (1964) is reproduced in Table (14):

ᴎᴋ	কা	kā	যোৰা, বেহা কৰা, জোখা, ছোৱালী খোজোতে দিয়া ভাৱ, নচা	To go; to trade; to measure; to offer things to the parents of a girl when wooing; to dance
ᴎᴋ	কা	kā	কাউৰী, কষ্ট, ঘৰ, এৰান, ডাকুল, কঠীয়া, কপালী, বেপাৰ, যথেষ্ট,	A crow; hardship; pain; a tether blow; a piece of wood tied to the neck of an animal; paddy seedlings; a forehead ornament; trade; sufficient
ᴎᴋ	কা	kā	অতীতকাল বুজোৱা ক্রিয়াৰ পাচত বহা অৱ্যয়, কোনো শব্দৰ আগত বহি অন্য অর্থ কৰে, যেনে - প্রতি, পাচ,	The verbal suffix denoting past tense; used as a prefix to denote: at; after, behind.

Table 14: Entry for ᴎᴋ in Barua, B.K. and Phukan (1964)

Barua and Phukan translate the word ᴎᴋ into 17 separate meanings, whereas the manuscript gives only 10. Three of the meanings are given twice by Barua and Phukan, and a further four meanings given by Barua and Phukan are not found at all in the Bar Amra, namely: 'sufficient', 'the verbal suffix denoting past tense', 'used as a prefix to denote: at' and 'after, behind'.

In addition, Barua and Phukan have grouped the words into three groups. This grouping is not based on anything in the original Bar Amra.

It appears that Barua has included a large number of words that N.N. Deodhai Phukan added but which were not found in the *Bar Amra*. The provenance of these extra words is not known, and while they may be authentic Ahom words, the fact that so many words of uncertain origin are included in this and other Ahom words lists has caused scholars considerable confusion.

There is an urgent need for a new edition of the *Bar Amra*, containing only the words found in the manuscript.

3.2.10 Barua, B. (1966).

B. Barua discusses and exemplifies what he claims are loan translations / calques in Assamese from Ahom. One of his examples is

> "TU - meaning an animal in Ahom and also suffixed to indicate birds and animals etc, is used in the same way as to denote a definite article in Assamese, such as 'Charai-to', 'Ghora-to' and 'Hati-to', that is the bird, the horse and the elephant ,etc." (1966:66)

The Assamese language does indeed contain a suffix / -tʊ/ (টো), defined by Baruah (1980:51) as "a kind of definitive for certain nouns and pronouns referring to them." Dasgupta (1993:58) defines it as being "u⹁. in case of inferior object or to indicate contempt or disrespect."

Barz and Diller (1985:170) discuss the possibility of Ahom influence on the Assamese classifier system. They find that the contemporary Assamese classifiers, including *to* do not have direct Tai cognates. Nevertheless, they point out that in the last six centuries there has been an increase in both the number and function of classifiers in Assamese. This time frame closely parallels the period in which Assam was ruled by the Ahoms.

3.2.11 Biswas, S. (1966).

Biswas (1966) contains information on the Khamti language from a conference paper. He does not present any data not in Needham (1894), and typically does not mention tones. His phonological analysis seems to be entirely drawn from the written forms, as The transcriptions seem based only on script, as (14) will show:

14)   hit (do), hing (dry)
      (1966:175)

There are no Tai languages in Northeast India in which these two words share the
some vowel, although they are written with the same vowel symbol. It is clear that in
Khamti they are distinct, as is shown in (15), from Weidert (1977)

15):   ʰɛŋ "to be dry dried" suggesting /hɛŋ/
       ʰiˌt "to do", suggesting /het/

3.2.12 Phukan (1966)

Phukan discussed the classifications of the Tai languages, and then proceeded to
present some comparative vocabularies. These contain some important information
about Aiton phonology in the 1960s, namely that the palatal semivowel was
pronounced with fricativisation, as is found today. This phenomenon is discussed in
some detail at 6.3.1.1.6. Phukan was also the first to report that the Aitons pronounce
/d/ (although writing <n>) in those words which are reflexes of words that have initial
*ⁿd in proto Southwestern Tai. His examples are given in Table (15):

Tai Ahom	Aitonia	English meaning
dap	written *nap* but pronounced *dap*	knife
di	written *ni* but pronounced *di*	good
doi	written *noi* but pronounced *doi*	hill
ji	written *yi* but pronounced *ji*	store

Table 15: Comparative word list from Phukan (1966)

Phukan further notes, in relation to *di*, that "Grierson puts it as *ni*; it is likely that he
has given the written form of the word."

Unfortunately, like so many of the Assamese scholars who have worked on Tai
languages, he does not notate tones. He lists the number of tones in some other Tai
languages and references his sources, stating:

> "It is certain that Tai-Ahom also used tones. But it is now difficult, without a
> thorough investigation, to say the exact number of tones used in Tai Ahom ...

from a general observation of words and their meaning in the Tai-Ahom
dictionaries and other records, it may not, however, be inaccurate to say that it
contained five to six general tone variations." (1966:11)

There is no direct evidence for the number of tones that might have been present in
Tai-Ahom, and in any case the number and distribution of those tones would have
changed over time. Phukan's claim appears to be based on the fact that most Tai
languages in the Shan group have five or six tones. Significantly, however, Aiton, the
closest Tai variety to Ahom, has only three phonemic tones.

Phukan's main contribution has been in the documentation of what he calls "couplets",
some of which are reprinted as Table (16):

two word luk lan (lit. son grandsom) - descendents
three word kin mung ban (lit. eat country village) - ruling the kingdom
four word kin mung kin ban (lit. eat country eat village) - ruling the kingdom
five word thao mung khry mung nyeu (lit. official big country big) - high officials of state
six word phan kan nan kan taw kan (lit. cut mutual fight mutual strike mutual) - mutual fighting and cutting
eight word rang kan plong kan sheng kan bang kan (lit. consult mutual decide mutual oath mutual bind mutual) - oath taking after consulting together

Table 16: Couplets in Phukan (1966)

These are also known as elaborate expressions, and are discussed below in 8.3.6.

3.2.13 Harris (1976)

Harris collected a substantial Khamyang word list in the 1970s. His informant was a Buddhist monk who was at the time living in Delhi. Harris is the first publication to systematically list tones for any Tai language of India, using the principles set out by Gedney (see 6.1.5.3). The tone box based on his study is presented in Table (17):

	A	B	C	DL	DS
1	4		5	4	
2	1				
3					
4	3	1	2	1	

Description of Tones:

Tone number (Harris 1976)	Description (Harris 1976)
1	mid falling
2	low falling with glottal constriction
3	high falling
4	high level
5	mid rising with glottal constriction

Table 17: Khamti tones (after Harris 1976)

Harris also discusses the phonetic realisation of consonants (1976:116f), of which the following are noteworthy:

/ph/ "represents a voiceless aspirated bilabial affricate [pɸʰ]"

/w/ "represents a voiced labio velar approximant [w] or a voiced bilabial approximant [ʋ]"

/y/ "represents a voiced front-palatal approximant [j] or a voiced front-palatal fricative [ʝ] ...In syllable final position this symbol represents a close front unrounded vowel [i]"

Vowel phonemes are discussed at (1976:119). Some of the most interesting observations are:

/ɛ/ "represents a falling diphthong [iɛ]

/ɔ/ represents [ɔ] "that occurs in free variation with a more common diphthong [uɔ]"

Unfortunately the word list given by Harris is very difficult to use; it follows neither Khamti nor English alphabetical order, and appears to be in several sections. The first section (p122-129) seems to follow Roman alphabetical order of presumed proto Tai forms. In this section of the word list, for example, words which have initial *ʔb in the proto language come after those having initial *a, but are generally realised with initial /m/ in Khamti.

From page 129 to the end of the list, there seems to be less motivation for the ordering of the items. It is regrettable that this is so, because it is very difficult to find any particular item in the list.

3.2.14 Weidert (1977 & 1979)

3.2.14.1 Weidert (1977)

Weidert (1977) is essentially a treatise on phonological theory, mainly relating to vowel phonemes. Weidert claims that Khamti has the following vowel phonemes:

i)	lax:				tense:		
	ɪ	ɯ	ʊ	ii)	i	ɯ	u
	ʌ				ɛ	ɔ	
						a	

Table 18: Vowel Phonemes after Weidert (1977)

In the realisation of these phonemes, the following rules apply

6.3.4	ɔₐ	>	o	(regularisation) - for vowels in open syllables
	ɛₐ	>	e	
6.3.5	uλ	>	o	(lowering rule) - for vowels in closed syllables
	iλ	>	e	
	ʊσ	>	w-	(semiconsonant realisation)
	ɪσ	>	j (y)-	
	ɯᵥ	>	final ɯ (non-syllabification)	
	ʊᵥ	>	final w	
	ɪᵥ	>	final j (y)	

Table 19: Vowel Phonemes after Weidert (1977)

In his data, there is no sign of the /ɯ/~/ɤ/ distinction, found in both Harris (1976) and Chow Khouk Manpoong (1993). This is exemplified in Table (20):

Weidert:	Harris	Phake	English gloss
`mɯɯ	mɯɯ³	mɯ̄²	hand
`mɯɯ		mə̄²	go back home
¯sɯk	sɯk⁴	sük¹	war
¯sɯ	sɤ⁴	sɔ⁶	tiger

Table 20: ɯ~ɤ distinction in Khamti and Phake

A short sample from Weidert's list will demonstrate the difficulty which his phonemic notation presents:

¯AIᵥ	4th female child
`phAIᵥ ¯AUᵥ	To be smoky
_AUᵥ	To take, accept
`AUᵥ ¯pɔt	To shorten, to reduce in size
_Aŋ	Foul or brackish water
`Aŋ	To be spread

Table 21: Sample of Weidert (1977) word list.

Needless to say it is difficult for native speakers (and for academic linguists) to make use of this list. Since both Weidert and Harris (1976) present word lists that are so difficult to work with, it is not surprising that many people tend to go back to Needham (1894) for information on Khamti.

3.2.14.2 Weidert (1979)

Weidert (1979) was an attempt to rediscover the Ahom tones. He compared Shan, Khamti and Standard Thai Tonal systems and proposed the following tone box for Ahom:

Consonant Classes	A	B	C	D
(1)	*5	*4	*2	*high
(2)	*5	*4	*2	*high
(3)	*5	*4	*2	*high
(4)	*1	*2	*3	*low

Table 22: Reconstruction of Ahom tones, after Weidert (1979)

The reconstruction is based on the assumption that where there is a regular correspondence between Khamti, Shan and Standard Thai, a correspondence can be expected in Ahom as well. So, for example, the Khamti high falling tone always corresponds with the Shan high tone and the Thai mid tone (A4 tone). Weidert assumes that this will be a discrete tone, as it is shown in Table (22). This reconstruction is discussed further in 6.5.3 below.

3.2.15 Wilaiwan (1983, 1986)

3.2.15.1 Wilaiwan (1983)

Wilaiwan (1983) presents a number of Tai Phake sentences and other data about the language. Her findings on Tai Phake tone are discussed in detail below in 6.2.4.3.

On word order, Wilaiwan observed that Tai Phake was

"in a state of changing from being a language which has the word order SVO to being one with the order SOV. We can see that of the old languages, both Ahom and Sukothai Thai were of the order SVO and we can show that this ordering is the original ordering of languages of the Tai family. As for present-day Phake, the form of the sentences is both SVO and SOV, but if we study the old manuscripts, we find the order SVO only." (1983:229)

While it is certainly true that both SVO and SOV sentences are found in Phake, the claim that SVO is the only order found in manuscripts cannot be sustained, as shown in the following example:

16)   ုၐ   ဃၐ   ၐ   ဓၐ   ၐ ။
      khām[2] saŋ[1] pū[1]   tak[1] lau[5]
      word   advise grandfather will tell
      "'I will speak ...ese (words) of teaching".'

(Tai Phake Manuscript, *Grandfather teaches grandchildren*, Introduction No. (10))

3.2.15.2 Wilaiwan (1986)

Wilaiwan (1986) is a study of Khamti, although it is not clear where the data was collected or how, whether by elicitation or from texts. Her conclusion (1986:178) is that

> "SOV is the dominant word order in Kamti while in other known Tai dialects SOV is the dominant one. We may conclude ... that Kamti has developed from SVO to SOV. It has also been shown that Kamti contains a set of object marking postpositions."

This claim is discussed in detail in Section 8.4.

3.2.16 Banchob (1970, 1987)

Dr. Banchob Bandhumedha is one of the giants of this field. Without access to her masterly *Phake-English-Thai Dictionary* (Banchob 1987), this thesis could not have been written.

Dr. Banchob Bandhumedha first visited Assam in April 1955, hoping to find speakers of the Ahom language. She met a number of Tai Ahoms but never met any who could speak Ahom, and came to the conclusion that there were indeed no speakers of Ahom (Navavan Bandhumedha 2002:24).

However Dr. Banchob did meet speakers of other Tai varieties and as a result published the *Phake-Englis'-Thai Dictionary* (Banchob 1987), and prepared an *Aiton-English-Thai Dictionary* t.... was nev-r completed. Thanks to the generosity of Dr. Navavan Bandhumedha. I have been able to obtain a copy of the *Aiton-English-*

Thai Dictionary and this work, together with copies of tapes collected by Dr. Banchob which will be referred to in this thesis.

Two Aiton texts collected by Dr. Banchob have been translated and analysed and are presented here for the first time, these are the *Story of the forest spirits and the opium pipes*, told by Ai Ne Phalung (Mohendra Shyam) and *Various Cultural Items* as explained by Nang Am.

3.2.16.1 Banchob (1977)

In the 1970s Banchob published a number of articles in the Thai journal *Satri San* in which she listed words and discussed the phonemes and tones of the various Tai languages in India and in other parts of Asia. It is reasonably clear that Banchob was never satisfied with her analysis of the Aiton tones, and when Banchob (1977) is compared to her manuscript *Aiton-Thai-English Dictionary*, a number of inconsistencies appear. These are discussed in detail at 6.3.4.1.1.

3.2.16.2 Banchob (1987)

Banchob's *Phake-English-Thai Dictionary* (1987) was a cooperative project between Dr Banchob and the Phake informant Ai Ney Ken Gohain. Banchob describes the phonology and tonal system of Tai Phake (see 6.2). Her key findings are supported by the analysis undertaken in the present work, and her transcription has been adopted for Tai Phake, with the exception that her notation of long vowels, except for /a/, has not been followed. Her decision to notate the phoneme /ɛ/ as <ai> before a velar creates a difficulty for its phonological status, because /ai/ is a reflex of proto Tai *aj. This is further discussed at 6.2.2.

The *Phake-English-Thai Dictionary* is the basis for all the lexicography in this thesis. It will be referred to throughout the thesis.

3.2.17 Nomal Gogoi (1987 & 1994)

3.2.17.1 Gogoi (1987)

Nomal Gogoi's *Assamese English Tai Dictionary* (1987) is a large work, of about 9000 entries, which he gathered by eliciting information from various Tai groups, comparing this with the old Ahom dictionaries and using the Thai-English Dictionary of Dr. Wit Thiengburanathum.

Whilst the endeavour of producing a dictionary for a reviving language is undoubtedly a difficult undertaking, and whilst any such effort would inevitably involve the gathering of data from a variety of sources, it is most essential that those sources be identified. To the speakers of any Tai variety in Assam, Gogoi (1987) is a most unsatisfactory work, because all of the words are mixed up. Examples (17) and (18) will clearly demonstrate this:

17)

ঘৰ , house	ৰুঁহ, হুঁহ

In (17), the Assamese word [ghɔr] and its English gloss are given, followed by two words, the first is the Ahom pronunciation, which we might regularise as [rɯn], and the second is [hɯn], which in Phake is pronounced [hɤn²] (written hən² by Banchob). These two words are simply different reflexes of proto South Western Tai *rɯanA4, one from Ahom where *r > /r/ applies and the other from the other Tai groups where *r > /h/ applies. One form has come from old Ahom manuscripts, and the other from spoken Tai, without any attempt to indicate which is which.

Standard Thai words are also intermixed, as in (18):

18)

ইংৰাজ , the British	ꩭ, ꩭ

The Assamese word [iŋraʒ] and its English gloss are given, followed by two "Tai" definitions. The first is the Tai term used in Tai villages ꩭ [ku¹ lā¹ phək¹],

whereas the second is clearly the Thai word /fa raɪN/ 'western foreigner' which is given as ꩭ [fa raŋ]. So far as I know this word is never used in Assam.

Academic linguists wishing to make use of Ahom lexicon for any purpose would certainly be advised to avoid this text, and rather make use of B.K. Barua and Phukan (1964), despite its shortcomings.

3.2.17.2 Gogoi (1994)

Gogoi (1994) purports to be a comprehensive study of the Tai Phake language, but, like so many works published in India during the last century, it completely fails to take account of the fact that Tai languages are tonal, and the tones are a crucial part of the language. Gogoi makes the basic error of assuming that tones mark out different meanings on words which are otherwise the same, stating:

> "Because of mono-syllabism, the same word is to express various meanings according to tones. Tones are the primary characteristics to differentiate meanings between homonymous words." (1994:7)

The reality is that in tonal languages such as Tai, two words which differ only in their tone are not the same word with different tones; they are different words, as different to a native speaker as two words which differ minimally only in their initial consonant or their vowel.

Gogoi (1994:19) lists the tones that Banchob found, but nowhere uses them. One is forced to the conclusion that he could not hear the tonal distinctions. It is not just the tones that are not distinguished. On page 18, he lists the vowel symbols with their names and pronunciation, of which Table (23) is a short extract.

ꩭ	wa choi sat	au
ꩭ	ha tho sat	au

Table 23: Extract from Table of Vowel symbols in Phake (Gogoi 1994:18)

The first of these two examples is indeed /au/, but the second is /auu/ or /aü/ in Banchob's transcription. This mistake is repeated throughout his thesis, as in example (19):

19) Tai script (Gogoi)   သၥ   ကၠ   �650ၠ   ၼ   �"ၠ   ၼ
     Gogoi's notation   cha   kau   khau   hau   ma   ta

     Tai script (corrected)   သၥ   ကၠ   ၶ   ၼ   ၼ   ၼ
     Banchob   cā⁵   kau²   khau³   haü³   mā²   ta⁴
     gloss   to   1Sg   rice   give   COME   PRT-request
     translation (Gogoi)   "Give me food."

As well as the distinction between /au/ and /auu/, Gogoi never makes the /uɯ/~/ɤ/ distinction (given by Banchob as /ü/ and /ə/), nor in his word list (Gogoi 1994:80-97) does he distinguish /u/ and /o/. For example, he lists *kui* as meaning 'cotton, banana'. Banchob gives *kui²* cotton, *koi³* banana).

On grammar, his thesis does not advance the study of these languages much. For example, of past tense, he says (1994:53)

"To form past tense, the following help-words are suffixed to the Verb, as ꩧ (ka), to go, ꧦ (ma) to come, ꩦ (yau) to complete, ꩧꩦ (ka yau), ꧦꩦ (ma yau) and ꩫ (koi) to end."

Even if we accept that some of these are tense markers, no attempt has been made to give any information about when they are used, or how they differ. For a study which has as its title *Morphological Study of the Tai Language*, this is most unfortunate.

The work also contains examples of Phake Literature, which show little improvement on the earlier generations of transcribers, as shown in example (20) from his transcription of *Grandfather teaches grandchildren*. (1994:72)

(20)   Passage 1   Kho Lik Pu Son lan

kham wan naŋ nam oi. au kin khau nau pak. wan lak san taŋ kheŋ. lum pha mü chai yiŋ. heŋ pi hau toŋ mat. kham son lan pu lat. koi mun wai phau mun ta wan...

Translation
Grandfather teaches Children
Words are sweet as molasses. Food taken through the mouth nourishes the whole body. Young children should remember the words spoken by grandfathers for thousands of years. O my dearest grandchildren, do not forget the old customs and traditions.

This section is not the beginning of the text, although from Gogoi's presentation it appears that it might be. Example (21) gives my analysis of the first portion of (20):

21.1)   ၶ   ဝၠ   ၼ   ၼ   ၼ
     khām²   wān⁶   nüŋ¹   nam⁶   ɔi³
     word   sweet   like   water   sugarcane
     '... words are as sweet as molasses.'

21.2)   ၼ   ၼ   ၶ   ၼ   ၼ
     au²   kin²   khau³   naü²   pāk¹
     take   eat   enter   in   mouth

     ဝၠ   ၼ   ၼ   ၼ   ၼ ॥
     wān⁶   lāk¹   sān⁵   tüŋ²   khiŋ²
     sweet   different   diffuse   all   body
     'Take and eat them, and (let) the sweetness diffuse through your body.'
     (Tai Phake Manuscript, *Grandfather teaches grandchildren*, Introduction Nos. (32) & (33))

Gogoi (1994) is the one of very few large works on the Tai languages published since Grierson's day. It does contain some useful data, but none of that data can be interpreted by anybody who is not fluent in the Tai languages, for all of the reasons discussed above. It has been necessary to discuss Gogoi in some detail both to point out the ways in which it could have been improved, and to ensure that it is not held up as a comprehensive grammar of the Tai Phake language.

3.2.18 Aimya Khang Gohain (1991, 1997, 1999)

The late Aimya Khang Gohain was a great help to me in my study of the Tai language. He would have greatly enjoyed reading this thesis, and he would certainly have

supplied corrections and opinions on many matters. He is the only author from Assam who consistently marks tones in his publications on the Tai language.

3.2.18.1 Aimya Khang Gohain (1991)

Aimya Khang Gohain (1991:47) discusses the phonology of Tai Phake. His vowel charts do not differentiate /ɯ/ and /ɤ/ (given by Banchob as /ü/ and /ə/) and curiously /ɔ/ is consistently written as <ɔ>, so that he writes /kɔk/ 'cup' (Banchob kauk[1]) and /ɔi/ 'yes' (Banchob ɜ̄[2]).

Although I shall be arguing that for Aiton /ɯ/ and /ɤ/ have merged, it is very clear that for Phake this is not so (see 6.2.2).

Aimya Khang Gohain also gives some useful text examples, which he phonemicises, with tones, and translates. (22) is a an example:

22)   From Phake Chronicle
ᦲᦜᧅᦜᦲᦵᦵᦲ ॥ pí laklī háimīt cām In the year haimit,

I have analysed this as (23):

23)   ᦲ   ᦜᧅᦜ   ᦵᦵᦲ   ᦲ
       pī[2]   lak[1] lī[3]   hai[2] mit[4]   cam[4]
       year   era   Hai Mit   PRT
       "The year Haimit."

Aimya Khang Gohain worried about the potential loss of the Tai Phake language. I reprint here, with my analysis of his texts, his final statement on the matter (1991:58):

> "The one time mighty Ahoms have lost their language. So also the Khamjangs
> and the Turungs. They still have a chance for revival or at least regain to some
> extent their lost language through the Phakaes, and also the Khamtis and the
> Aitons. The Phakaes are also now in the threshold of being lost, unless the
> Ahoms and the Government come to help. The Tai-Phakaes have a saying:

24)   ᦲ   ᦁ   ᦲ
       mã[6]   pai[2]   hai[5]
       dog   go   dry rice field
       'Dog roams in jhumfield, i.e. a useless journey.'

A riddle:

25)   ᦲ   ᦲ   ᦲ   ᦲ   ᦲ   ᦲᦲ   ᦲᦲ   ᦲᦲᦲ
       tãi[2]   sãm[6]   pī[2]   ma[1]   neu[5]   ho[6] khau[1]   taŋ[3]   tuŋ[2] khuŋ[2]
       die   three   year   NEG   rotten   knee   erect   INTENS.

'Died three years ago, but not rotten. Its knees are still erect as they were. (dead spider)'

And a proverb

26.1)   ᦲ   ᦜᧅ   ᦵᦲ   ᦲᦲ   ᦲᦲ
         mai[4]   phãk[3]   kɔ[2]   pen[2]   tauk[1]
         tree   separate   clump   be   thin strip
         'Trees separated from grove become pieces.'
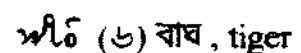
26.2)   ᦲᦲ   ᦜᧅ   ᦵᦲ   ᦲᦲ   ᦲᦲ
         mauk[1]   phãk[3]   phã[4]   pen[2]   phun[6]
         cloud   separate   sky   be   rain
         'Cloud separated from sky becomes rain drops.

26.3)   ᦲᦲ   ᦜᧅ   ᦲᦲ   ᦲᦲ   ᦲᦲ
         kon[2]   phãk[3]   məŋ[2]   pen[2]   pɔ̃n[5]
         person   separate   country   be   other
         'Man separated from country becomes alien.'

"Sometimes it occurs in their mind whether all their contributions will be "useless trip" or though appears to be dead but still not "rotten"; or have they become alien or are they soon going to be alien from their own people like the "rain-drops?" They have to wait and see."

3.2.18.2 Aimya Khang Gohain (1997)

Aimya Khang Gohain (1997) is a text book for teaching Ahom script but with Phake tones and grammar. Words are introduced in Ahom script with a tone mark in Assamese and then the meaning in Assamese and English, as in (27)

27) ꩵꩮꩥ (৬) বাঘ , tiger

In (27) an Assamese number 6 in parentheses indicates that the word has the sixth tone, with tone numbers following Banchob (1987).

It has not been possible to fully analyse this text, and in particular to identify which elements are purely Phake and which based on the old Ahom manuscripts.

3.2.18.3 Aimya Khang Gohain (1999)

Aimya Khang Gohain (1999) was the first translation of a complete Tai Phake text into English. It is a full and thoroughly annotated translation of the *Sucha Naraha*, an important Buddhist text. It will become a very valuable aid for scholars of the Tai languages into the future. (28) is the analysis of a small section from page 46 of the text (1999:66), using the manuscript text and Aimya Khang Gohain's translation, with phonemic transcription and inter-linear gloss added.

28.1) ꩰꨳꨵꨳꨱꨰꩮꨲꩳꨪꨵꨳꨯꩳꨱꩳ »

an²	nai⁴	phī⁶	kon²	sa¹ ta¹ wā¹	cam⁴
CLF	this	spirit	person creature		PRT

'Regarding gods, humans and other beings

28.2) ꨰꩥꨮꩬꩯꨰꩥꩮꨱꩳꨯꩳꨰꩥꩳꨮꩳꨱꨲꨵꩳꨱꩳ »

lɔ² phā⁴	tɔ² sā⁴	ik¹	mɔ² hā⁴	kī⁴ le¹ sā¹	khüm²	khiŋ²
greed	anger	with	delusion	passion	put a barrier	body

'the K<u>i</u>lesa (defilement) like <u>Lobba</u> (greed), <u>Dosa</u> (aversion) and <u>Moha</u> (delusion) that obstuct the Path of progress toward <u>Nibbana</u>.'

3.2.19 Kingcom (1992)

Kingcom (1992) is essentially a comparison of Tai Phake with Standard Thai, which is his own native variety. The comparison is based on the following key points:
- phonological comparison
- comparis. ⟩ of syllable structure
- comparison of certain morphemes - particularly classifiers
- classification and comparison of lexical items

There is a data set of 495 sentences which were presumably elicited and compared with Kingcom's own variety (Standard Thai). He undertook two field trips to Assam in April and October 1990

The phonology is basically that of Banchob (1987), but there are some curious findings. For example, on page 7 he gives /phi/ 'fat', whereas Banchob gives *pī²*. Kingcom also discusses the consonant correspondences in detail, which Banchob had previously done. It is likely that he did not have access to Banchob, which is not listed in his bibliography

There are some consistent misreadings. For example, Kingcom always notates ꩵ [auu] as /ɔ/. eg mɔ "you" (Banchob *maü²*) and he does not appear to hear the distinction of /ɯ/ and /ɤ/, although he gives it in his phonology.

Kingcom's treatement of syllable structure is quite thorough, and will be discussed in 6.2.3 below.

Unfortunately, it appears that Kingcom has not read Gedney (1972), which would have helped him to understand the comparisons between the Tai and Standard Thai tones. For example he speaks of the existence of one tone, the mid to rising tone, but is not consistent in marking it.

A large part of the text is spent in the analysis and comparison of classifiers. Kingcom divides the classifiers into 'semantically restricted classifiers' and 'lexically restricted

classifier'. The first is defined as "those where nouns that can occur with them share some semantic features[31]. Some features will be listed at the beginning of each classifier." (1992:68)

An example of such a classifier is (29), from page 69:

29)   /kŭ/
      <+pair> <+wearing>
            /kép tín/ 'shoes'

Kingcom categorises lexically restricted classifiers are those which "occur with a limited number of nouns in both the varieties. The nouns form closed systems." (1992:110). (30), from page 119, is one of these:

30)   /mák/   This classifier occurs only with the the noun /hôklăy/ 'staircase steps'

He then fully lists all classifiers, comparing Tai Phake and Standard Thai, grouping the classifiers into those which are (i) formally and functionally similar, (ii) formally different but functionally similar and (iii) additional.

Kingcom's lexical comparison is based on a basic word list added to words obtained in conversations. He found that Tai Phake and Standard Thai shared lexicon in varying proportions depending on the the semantics of the lexicon. The highest proportion of sharing was Animals 77.6% and Systems numeral 75.4 %, whereas the lowest proportions were Social Institutions 23.3% and Technology 0%. (1992:204)

Having noted a reduction in consonantal and vowel phonemes, number of classifiers and variety of lexicon in Tai Phake vis a vis Standard Thai, Kingcom makes the following extraordinary statement: "Tai Phake is a stunted language. There is no growth and it is not used in all fields of life as a living language." (1992:209). How exactly the national language of a large independent country and a minority language spoken in up to 11 villages are supposed to compare, is not clear to me. What is clear

---

[31] My copy of Kingcom (1992) is unreadable at this point.

to me after spending a considerable time among the Phakes is that their language can express everything that they wish to express, with the occasional borrowing of words from outside, although far less borrowing than is to be found in modern Standard Thai.

At the end of his thesis, Kingcom gives 495 sentences, of which No. 44 is here reprinted as (31). I have given Kingcom's example in Tai script - which he does not use, with phonemicisations after Banchob (1987).

31)   ကော် ဖော် ကျ ဖော် ဇော် ဇော် ဝဲ ကျ ဖော် ဂျ
      kau²   han⁶   kā¹   hən²   ɔn¹   cɔn¹   khau⁶   ŭ¹   ɛm³   nam⁴
      I   see   past house small John (of) is   near river
      "I saw John's small house near the river."

Some of Kingcom's key findings were summarised in Dhongde & Kingcom (1992).

3.2.20 Diller (1992)

Diller's invaluable sketch of the Tai languages of Assam (1992) includes the first publication of any text in any of the living Tai languages using the Tai script since Grierson published the Nora riddles in 1904. In addition to presenting this important text, Diller carefully discusses the historical background of the Tais in India, the state of the Ahom language, and gives some very useful information about the syntax of the living Tai languages.

Diller (1992) is an indispensable aid to any scholar working on the Tai languages of India and will be referred to throughout this thesis.

3.2.21 Boruah (2001)

Boruah performs a valuable service by comprehensively listing all the Tai villages in Assam and Arunachal Pradesh (2001:41-42). Unfortunately, like all other Assamese writers on Tai, he does not use tonal notation in this article, though he is clearly aware of the tones. For example, on page 45 he gives a table with a syllable *kā* and its five meanings without indicating which meaning is in which tone. He makes the statement

that: "five tones have been observed, three of which are more prominent. These are rising, falling and level. Sharp rising and sharp falling are also observed." (2001:46).

Boruah's analysis of the vowels is also at odds with other evidence. He states that

> "The phonetic system of all these six groups of Tai language is almost similar. Ahom, Khamti, Phake, Khamyang, Turung and Aiton have got eight vowel phonemes namely / i e u o a ɔ ā ü /. These vowel phonemes have their long and short forms." (2001:46)

Boruah gives the 8 vowel system, as follows:

	Front		Central		Back	
	Short	long	Short	long	Short	long
High	i	iː	ü	üː	u	uː
Higher-mid	e	eː			o	oː
High-mid low					ɔ	ɔː
Lower-mid					a	aː
Low			ā	āː		

Table 24: Vowel Phonemes in Tai languages, from Boruah (2001)

So far as I know, no Tai language exhibits such a system. It is true that in the Thai writing, there is more specification possible of back vowel distinctions than of front vowel distinctions, and I will be arguing in 6.3.2 that in Aiton the lack of vowel distinction in front vowels is clearer and more complete than it is for the back vowels. Nevertheless I know of no evidence for a four way distinction of back vowels. It may be that this has arisen because <a> is used in some Assamese transcriptions to indicate /ɔ/, and that as a consequence /ā/ is used to indicate the low vowel. It may be that <ɔ> and <a> in Boruah's transcription indicate the same thing.

Moreover, it is distinguishing feature of the Tai languages in Northeast India, and in neighbouring Burma, that length distinctions have been lost in all vowels, except for the low vowel. Boruah's chart is thus very misleading.

3.3 Other Sources

A number of non-linguistic publications contain important data on the Tai languages. The most significant of these are briefly reviewed below.

3.3.1 Barua, G.C. (1930)

The first attempt to publish the Ahom Buranji (historical chronicle), Barua, G.C. (1930) consequently contains a huge amount of data in the Tai Ahom language. His translations and text are still used as the basis for many other examinations of the Ahom language, and where the Ahom historical texts are published by other scholars in part, it is usually this text from which they have come.

The Ahom Buranjis have recently been comprehensively re-examined and republished by Ranoo 1996.

3.3.2 Yehom Buragohain (1981, 1998)

Yehom Buragohain, who occupies the position of Special Officer for Ahom Studies in the Department of Historical and Antiquarian studies in Guwahati, is one of the principal informants for this study, and has revised the translations of many of the Phake texts which are presented here. She has also published a small number of articles.

3.3.2.1 Yehom Buragohain (1981)

Yehom Buragohain (1981) lists the Tai divisions of time, as:

Po ting khun:	the period when the night is at the highest point, i.e., midnight
Po ting khun na kai kai khan:	the period of time after midnight but before the cocks crow.
Po na kai:	the period immediately before crowing of cocks
Po kai khan:	the period when the cocks crow
Po phong phow:	the twilight period in the morning

Table 25: Some Time Divisions among the Tai (after Yehom Buragohain 1980:7)

The second of these entries can be phonemicised (in Phake) as follows:

32) ဧꩻ   ဝၣၠ   ꩻꧠꩻ   ꩻꧠ   ꩻꧠ   ꩻꧠ   ꩻꧠ
     pɔ[1] teŋ[5] khün[2] nā[3] kai[1] kai[1] khan[6]
     time   middle night   face   chicken chicken crow

This very useful article lists the divisions of the day as they are found in the Ahom manuscripts. It is to be hoped that in the future, publication of such research will include tonal information.

3.3.2.2 Yehom Buragohain (1998)

Yehom Buragohain's list of Phake villages (1998:127) has already been referred to in 2.3.3 above. She also lists Phake texts with a non-phonemic "transliteration", and translation, as the following example from page 130 shows:

33)   Nuk yang lai sang mau kok kek ti aim nong
      (Oh crane why are you moving hither and thither on the lakeside?"

      To pa ma ok
      (Because fishes are not out)

This can be regularised as:

34.1) ꩻꧠꩻ   ꩻꧠꩻ   ꧠ   ꩻꧠꩻꧠ   ꩻꧠ   ꩻꧠ   ꩻꧠ
       nok[4] yāŋ[2] lai[4] saŋ[1] maü[2] kauk[4] kaik[4] tī[5] ɛm[3] nauŋ[6]
       crane why   2Sg   hither and thither at   edge   pond

34.2) ꩻꧠꩻ   ꩻ   ꩻ   ꩻꧠꩻ
       tɔ[5] pā[2] ma[1] auk[1]
       because fish   NEG   come out

3.3.3 Sharma Thakur (1982)

Sharma Thakur (1982) is essentially an anthropological study of the Tai Phake, and does not claim to be a linguistic study. It does, however, reproduce a number of texts, with translations. The translations appear to be good, but the transcription of the Tai Phake misses both vowel and tonal distinctions. Table (26) is an extract of Thakur's translation of *Grandfather teaches Grandchildren*.

3. Chom Che Hik Fu Khām	People condemn the lazy people.
4. Hāi Na Hāo Me Ka	When you plough your fields, bring your spade to every nook and corner
5. Khāt Chang Yā Khāng Non	Do not while away your time on the plea of tiredness

Table 26: Extract from the translation of *Grandfather teaches Grandchildren*, Sharma Thakur (1982:200)

Sharma Thakur's No 3, is the same as the second line proverb No (3) in my translation of *Grandfather teaches grandchildren*, here reprinted as (35):

35)   ꩻꧠ   ꩻꧠꩻ   ꧠꩻ   ꩻꧠꩻ
      hāi[4] yāk[4] pɔn[5] nü[2] yɔm[2]
      bad   poverty other people   look down upon

      ꩻꧠꩻ   ꩻꧠ   ꧠꩻꧠ   ꩻ   ꩻꧠꩻ
      sɔm[2] sɛ̄[6] hik[4] phü[3] khān[4]
      contempt/ridicule PRT call male (person) lazy
      'If you are poor, others will look down on you, they will ridicule you and call you lazy.'
      (Tai Phake Manuscript, *Grandfather teaches grandchildren*, No. (3))

As with so many of the works referred to in this chapter, it would have been much more useful if the transcription had marked tones and vowel distinctions. However Sharma Thakur wrote before the publication of Banchob (1987), and is writing in a different discipline (anthropology). Perhaps there needs to be more consideration given by those outside of linguistics to the problems of transcriptions such as Table (26), and some greater efforts made to produce transcriptions which will be useful in a multi-disciplinary sense.

3.3.4 Ranoo (1986, 1996)

3.3.4.1 Ranoo (1986)

Ranoo's 1986 thesis on the Ahom script remains the most comprehensive study of any of the scripts of any of the Tai communities of India. She has carefully compared the usage of the Ahom script at various periods, and her work can be reliably used for the dating of Ahom texts.

3.3.4.2 Ranoo (1996)

As referred to above in 3.3.1, Ranoo (1996) is a new and thoroughly researched edition of the Ahom Chronicles, or *Ahom Buranji*, translated into Standard Thai. The text is presented in four columns, as follows:

Column 1	Ahom text
Column 2	Transliteration of Ahom text into Standard Thai script
Column 3	Word by word translation into Standard Thai
Column 4	Translation into Modern Standard Thai.

Any study of the grammar of the Tai Ahom manuscripts would begin by thoroughly examining this work.

3.3.5 Terwiel and Ranoo (1992)

An earlier publication of an Ahom text, this time transliterated and translated into English, was Terwiel and Ranoo (1992). This also contains a useful summary of the Ahom script and some grammatical information.

3.3.6 Chaw Khouk Manpoong (1993)

Chaw Khouk Manpoong (1993) is a set of two primers to teach the revised Khamti alphabet (see 7.3 below). This was the first attempt to introduce tonal marking to the orthography of any of the Tai languages of India.

3.3.7 Jaya Buragohain (1998)

Jaya Buragohain worked in the Tai Aiton villages in the few years immediately before I first visited there, and she had the good fortune to have as her informant the late Sri Mohendra Shyam. Her information on the Tai villages has already been referred to in 2.3.2 above.

Jaya Buragohain gives a number of texts, which are prayers spoken at various important times of the year, such as when rice sowing begins. Unfortunately, as is typical of Assamese scholars, she does not provide tones. Example (36) is a text which is also translated into Standard Thai.

36)
"Nursery Bed (ta ka)
- the following prayer is offered on the day of sowing by the head man of the family, who "offers his prayer to the Lord Buddha at the Kheng-lik of his house. On reaching the nursery bed a prayer is made at first towards the east after praying to the Spirit of the earth thus:
    "ao mun Phra, ao mun Tra, ao mun Sanggha
    wan nai kao wan khao ka
    ao mun Phra, ao mun Tra, ao mun po mun me
    mun pu nai, ya nai, kao wan khao

translated as
    "O'Phra O'Tra O'Sanggha, today I have come carrying in my hands the seeds of paddy to throw in the name of Lord Buddha. O'Phra O'Tra let me have the strength to give alms to the Sangha ...
    (Buragohain 1998;72)

There are several errors in (36). Dibya Dhar Shyam pointed out that nursery bed should be ဝ်ကၢ (*tii² kaa³*) and what she has rendered as *Kheng-lik* is probably ကၢၢ်ၤဝ်ၵၢ် (*kyɔŋ² lik³*). Jaya Buragohain's transcription of the prayer is also confusing. The first two lines are re-analysed here as (37), employing linguistic methodology.

37)

ᵃᵘᵍ	ᵉᵘᶦ	ᵍᵖᵍ	ᵃᵘᵍ	ᵉᵘᶦ	ᵍᵖᵍ	ᵃᵘᵍ	ᵉᵘᶦ	ᵘᶜᵘᵍ
au²	mun²	phraa²	au²	mun²	traa²	au²	mun²	saŋ¹ khaa¹
take	glory	Buddha	take	glory	Dharma	take	glory	Sangha

ဝာ် ဈ် ကၟ် ဝါ် ၈် ကၟ

wan² nai³ kau² wăn¹ khau³ kaa³
day this 1Sg sow rice rice seedlings
'With gratitude to the glorious Buddha, Dharma and Sangha, today I (am going
to) sow the rice seedlings.'

Jaya Buragohain also lists the clans and subclans of the Aiton (1998:105).

Her work appears to be a valuable anthropological record of the Tai Aiton.
Unfortunately she never provided copies of her thesis to the Aiton community, and so
there has been no opportunity to properly assess her claims.

3.3.8 Hazarika, Chow Nogen. (1993, 1996).

3.3.8.1 Hazarika (1993)

This work is a collection of songs written by Hazarika in modern Ahom. There is no
English or Assamese language in the bulk of the text.

(38) is the title and first line of the first song, while (39) is my attempt to analyse it.

38) ၉ လစ္ခုဆ်ဆာ၀ာ၀<
လစ္ခုဆ်ဆာ၀ဆ်< ကရ၀၀ုဲ ၀ ၀ဟၖၖ၅<

39) ၀၀၆ ၀ဆ်ဆ်< ဆာ၀ဆ်< ကရ၀၀ုဲ ၀ ၀ဟၖၖ၅<
cau³ phraa² traa² a¹ loŋ² ɯi² sik¹ kyaa²
RESP Buddha Dharma Boddhisatva EXCL Sakka

3.3.8.2 Hazarika (1996)

Hazarika (1996) is a passionate essay in favour of Ahom revival and in favour of
Ahom culture. Referring to other communities, to philosophers and thinkers, the
author hopes to persuade the Ahoms to take their cultural revival seriously. The
following quote will illustrate the tenor of the work:

"'The color', says B. Russel, 'ceases to exist if I shut my eyes, the sensation of
hardness ceases to exist if I remove my hand from contact with the table...'
The sensation of the colourful past of the Tai Ahoms may not help if we
cannot assert , survive and REVIVE, we, after Descartes, shall have to
syllologise the first dictum of modern philosophy, viz. *cogito ergo sum* (I
think, therefore, I exist). Let us also like Sartre practise, 'If to die is to die in
order ... to bear ... witness ....etc., then anybody at all can die in my place.'"
(1996:2)

4. Theoretical considerations

4.1 Linguistic and Theoretical basis for this thesis

No piece of academic work exists without reference to other work; nor does there exist any piece of academic work which has been informed by all the other academic work which might be relevant to that endeavour. The exponential increase in academic publication over recent years, in many languages, renders it impossible to read all the literature from the present, let alone from the past, any of which might conceivably relate to and inform the matter being discussed. In this chapter I will touch on that part of the literature that I have found useful as general background to this thesis.

4.1.1 'Basic Linguistic Theory' and formal theories of syntax

Not only has the huge amount of literature made the task of writing a thesis like this more difficult, but a writer in my position needs to adopt a view on a basic controversy within linguistics. Is it the task of someone in my position to write this grammar within the framework of a particular theory of syntax, or outside of it?

In the opening to his work on the study of language universals, Comrie (1989) discusses what he terms "two major methodological approaches". Three important parameters mark out the two approaches, "the data base for research ..., the degree of abstractness that is required... and the kinds of explanations advanced ..." (1989:1)

Comrie's position is that to study language universals, data would be needed from a wide range of languages, and that:

> "linguists advocating this approach have tended to concentrate on universals statable in terms of relatively concrete rather than very abstract analyses, and have tended to be open, or at least eclectic, in the kinds of explanations that may be advanced ..." (1989:2)

According to Comrie, the alternative approach, associated with Noam Chomsky, argues that:

> "... the best way to learn about language universals is by the detailed study of a small number of languages; such linguists have also advocated stating language universals in terms of abstract structures and have tended to favour innateness as the explanation for language universals." (1989.2)

Since this thesis deals with a single language group, the first of Comrie's parameters is not relevant here. However the general direction of Comrie's thought, that concrete structures and open explanations are preferable, is indeed the direction of this thesis. In that sense, the underlying theoretical basis of this thesis may be termed 'typological'.

More recently, Dixon (1997) has proposed that linguistic endeavours of the kind which I am undertaking would be best done in the framework of 'Basic Linguistic Theory', a term which

> "has recently come into use for the fundamental theoretical concepts that underlie all work in language description and change, and the postulation of general properties of human language." (Dixon 1997:128)

Basic Linguistic Theory is contrasted with what Dixon refers to as "formalisms", a long list of which he names (1997:131), adding that

> "Each is useful for describing certain kinds of linguistic relationships, but it is put forward as if it were a complete theory of language. The word "theory" is being used in a novel way." (Dixon 1997:131).

Dixon outlines the program of Basic Linguistic Theory as follows:

> "When working in terms of Basic Linguistic Theory, justification must be given for each piece of analysis, with a full train of argumentation. Working within a non-basic theory there is little scope for argumentation ... Needless to say, such an approach tends to make all languages seem rather similar, and

ignores the really interesting features which do not conform to any
expectations." (Dixon 1997:132-133)

It is hoped that this thesis does provide full and comprehensive argumentation for the
claims contained herein. The model adopted for this thesis, then, is consistent with
Basic Linguistic Theory.

Nevertheless, a scholar in my position must take note of the endeavour of those who
are working within any of the various formal theories of syntax, perhaps best
represented by the figure of Noam Chomsky. Several features of Chomsky's program
need to be taken into account by linguists working on projects like this thesis:

- A key aim of Noam Chomsky and others would appear to be to fathom the
  mysteries of the human mind, using language as a window to the mind.
- Chomsky and others hold that non-native speakers can hardly be expected to
  know enough about the languages they are looking at to make really useful
  judgements.
- There must be a universal human language faculty; otherwise why do all humans
  develop language. There must therefore be language universals, which we will
  discover by introspecting.

It is obviously true that native speakers of a language have a deeper understanding of
their language than visiting linguists, and we may certainly hope that in the future
there will be trained linguists who are native speakers of the Tai languages and who
can advance their study far beyond what is possible in this thesis.

However, this thesis does contain considerable insight into the Tai languages.
Moreover, had I not proceeded with this thesis, I would not have collected any
recordings of the Tai languages of Assam. Much linguistic information would be lost
to future generations. For example, we can expect the best Khamyang speakers to
survive no more than another generation. The recordings that have been made can
always be re-examined in the future; but if this work is not done now, it never will be.

The aim of scholars like Chomsky and many others who are attempting to find and
refine the universal theories of language are worthy aims; however in the case of the
present study, it will be found more useful to describe the language in terms of pre-
theoretical terminology, such as 'subject' and 'object', 'noun' and 'verb' without
recourse to formal theories.

4.1.2 Text Based Grammar

Supposing we were to ponder what we would hope to leave behind of the English
language for future generations. Most would probably prefer to leave a work of great
literature rather than a grammar which explains how past and future time are marked
or how to express a sentence where the verb has three arguments.

There is a very real chance that the Tai languages will cease to be spoken in a few
generations. The Tai Khamyang certainly will be and even if Tai Aiton and Tai Phake
are spoken in fifty years time, the likelihood is that the domains of usage of these
languages will be even further reduced.

There is an equally very real chance that this present work will become and remain
the standard work on these languages until that time, and that by then the opportunity
to collect the rich variety of texts which is now available will be lost.

Therefore this thesis is in a great measure a collection of texts, a compendium of as
much of the knowledge and literature of the Tai that could be collected and analysed[32]
in the all too brief time available. Whatever analysis can be drawn is taken from those
texts, and every one of the texts is presented in transcription, analysis and with linked
sound files (see 9.5 below). The process of collection and analysis of these texts is
discussed in more detail in Chapter 5.

---

[32] Perhaps as little as 10% of the recordings that I made have been analysed. Great riches remain to be
explored.

Every single sentence example in this thesis will give its source; and where it is from a text, as the overwhelming majority will be, the example will be linked to that text, so that the reader can see and hear the context from which the sentence comes.

That context is not just a position in time within a series of utterances. Every text that has been collected has meaning, and that meaning is important. If I could, I would urge the reader to read the texts for their meaning, as well as for their structure. If we do not know the meaning, how are we different from the 90,000 ministers who could not interpret the Queen's dream in the story of *The Twelve Questions* (2-1-1-3), told by Sa Cham Thoumoung, of Duburoni village?

There is another reason for basing this study firmly on texts. The texts collected are of a wide variety, there are stories, songs, prayers, manuscripts, explanations of the manuscripts or of rituals, speeches and conversations, they represent a wide range of the speech act types in the Tai communities.

Himmelmann (1998:178) puts the question as to whether certain kinds of linguistic structures only occur in particular kinds of communicative events. If this is so, and intuitively it is (the type of linguistic structures exhibited in this thesis being very rarely found in chat), then we need to record as wide a range of communicative events as possible in order to even find all the linguistic structures.

Himmelmann (1998:180) lists events on a parameter of spontaneity, from the most spontaneous, events like exclamations, through to the most planned, such as ritual texts. His types of communicative events are listed in Table 1

Parameter	Major Types	Examples
:·-·nned		
	exclamative	'ouch'
		'fire!'
	directive	'scalpel!'
		greetings
		small talk
	conversational	chat
		discussion

	monological	interview
		narrative
		description
		speech
		formal address
	ritual	litany
planned		

Table 1: Types of communicative events according to the parameter of spontaneity, Himmelmann (1998:180)

Thus the writing of a text based grammar is underpinned by two theoretical principles. First, since this thesis is about the living language of a group of humans, for whom their language is of the greatest importance, we will pay tribute to that importance by recording as rich a variety of that community's heritage as we can. Second, by doing so we stand to gain a much greater insight into the language anyway.

By way of closing comments to this issue, let us compare the approach of Grierson (1904). Grierson, following the method of Wilhelm von Humboldt[33], collected examples of the *Parable of the Prodigal Son*, in every Tai language (and indeed in 364 languages of the Indian empire!) and used that as the basis for his study. From this text he drew conclusions about the Tai languages.

For one century ago, that approach can only draw praise; but for today it is no longer appropriate. If we try to ask the same questions of every language, we run the risk of receiving similar answers. It will surely seem incredible to the reader reading this thesis when it is as old as Grierson is today, that in the same opening years of the 21st century there are projects afoot to make a record of all of the languages of the world, using a translation of the *Book of Genesis* as the comparative text.

Surely, it is the linguistic and literary forms that differ from language to language and community to community that are most illuminating. That is why letting the grammar be based on the texts that emerge is likely to lead to the most interesting results. This is why, in Mithun's (1998:191) terms, we should let the speakers speak for themselves.

---

[33] I am grateful to Dr. Tamsin Donaldson for pointing out to me that Humboldt had pioneered this method.

4.1.2.1 Description and Documentation

Himmelmann (1998:165) proposes that the relationship between documentation and description of languages should no longer be seen as "one of unilateral dependency, with the documentary activity being ancillary to the descriptive activity (i.e. primary data collected IN ORDER TO make a descriptive statement of the language.)" Himmelmann goes on to add that the two activities should be seen as being in a relationship of "bilateral mutual dependency."

Himmelmann, therefore, proposes a new field of linguistic endeavour, which he calls 'Documentary Linguistics'. The basic aim of this discipline will be "to document the linguistic behavior and knowledge found in a given speech community," (1998:168). The data thus assembled should be "amenable to a broad variety of further analyses and uses."

A clear statement of the importance of this approach is that "there is no infallible way of drawing corpus from descriptive statements, that corpus is richer." (Himmelmann, pers.comm.). One can always write a grammar from the texts, but one could never expect to recover the texts from a grammar.

From the perspective of the current speakers of endangered languages, the documentation process is certainly more important than the descriptive. Sucn communities often hope for linguists to permanently record and make available large numbers of texts. This is one of the reasons why so many texts are presented here and why the sound recordings of them are also included. Moreover, for the speakers of these languages, the message conveyed in the texts is at least as important as the means of conveying it. If our only knowledge of these languages were based on a translation of Genesis 1, it would be very shallow knowledge indeed.

Let us suppose that in a hundred years our understanding of human grammar has advanced as much as it appears to have in the last hundred years, and let us further suppose that most of the languages now spoken have passed away. For some of those

languages a grammar will have been written, which, in traditional format, may have two or three short texts attached.

When the linguists of the year 2102 come to do their re-analysis of that language, and go back to those texts and see possibilities which we cannot see today, they may lament the lack of material with which to work. If, on the other hand, we provide them with a great deal of text, and a pointer as to what might be found where, will they not then be able to do much more with those texts?

Himmelmann notes that "language documentation is NOT some kind of "theory-free" enterprise. Instead, documentary linguistics is informed by a broad variety of theoretical frameworks and requires a theoretical discourse. ..." (1998:190)

This thesis will provide analysis drawn from the texts, which could be termed 'a grammar of the Tai languages', though the term implies the kind of deep knowledge of the language which I could not hope to acquire in the time allowed. It may be the texts themselves that will be the most important contribution of this work to future scholarship.

4.1.3 Language centred approaches

A very important issue in dealing with languages like the Tai languages of Assam, or indeed any language other than the researcher's native tongue, is that it will be difficult for the researcher not to be influenced by their native tongue in undertaking linguistic research.

This difficulty is multiplied if the researcher happens to have been born a monoglot English speaker, as I was. It is too easy for an English speaker at the beginning of the 21st century to assume that all knowledge is available in English and that the English language traditions will provide all the theoretical bases for the understanding of language that are needed.

Working in the field with non-Indo-European languages gradually makes one aware of other ways of describing things, other ways of thinking not just about the world, but about the language which it is our task to describe.

For example, readers of Noss (1964), a masterly survey of the Standard Thai language, will have noticed that many single Standard Thai lexemes are analysed as several different polysemous lexemes by Noss. Usually this is because this one Standard Thai lexeme cannot be translated into a single English lexeme. Yet for Thai people, it is a single word.

Long & Zheng (1998) show a fairly Chinese-centric approach in their grammar of the Kam language - most obviously in the title of the work, which uses the Chinese name for the people, Dong, rather than Kam, the named used by the people themselves.

But their Chinese centred approach extends somewhat further than that. For example, in their very useful brief description of the language (1998:17), they state: "The word order of Dong is essentially the same as that of Chinese, with one clear exception: in Dong, adjectives come after nouns while in Chinese they come before."

The really interesting observation is not that Kam (Dong) has noun-modifier word order, for that is typical of the Kam-Tai family from which it comes, rather, it is that in some cases Kam (Dong) differs from the patterns to be found in other languages to which it is genetically related and has, perhaps, converged to Chinese.

In Chapter (3), it has been seen that many of the writings about the Tai languages of Assam, both those from the early British and American researchers, and those more recently by Assamese researchers, have been seriously affected by the language bias of the authors. The extent to which this is also true of the present work will be left to the reader to judge.

4.1.4 Some particular difficulties in working with the Tai languages of Assam

When writing a description of a language, the tradition has long been to produce a synchronic description, in which everything is related to the stage at which the language is now.

Advances in the study of the history of the Tai languages cannot, however, be ignored. For example, the analysis of the phonology of the Tai languages may differ according to whether the history of the languages is taken into account or not. The Aiton consonant inventory is such a case in point. In section 6.3.1.2, a plausible analysis of the Aiton consonant inventory is presented which is based purely on synchronic observations and takes no account of historical factors or other related languages. A different inventory is presented in 6.3.1, having taken account of these factors. On balance, the latter approach is preferred, at least in this case.

Another issue which has affected this thesis is the question of grammaticality judgements, the intuition that native speakers have about whether a particular utterance is grammatical or not. Many times I have asked the Tai informants whether a particular sentence is grammatically acceptable or not; in many cases the answer may be yes, but only depending on circumstances. Since the circumstances are the linguistic context, it is much safer to rely on actual examples of spoken text where the context can be made explicit.

4.2 Approaches to the study of Tai languages

Several different theoretical approaches have been brought to bear in the description of Standard Thai, or Siamese. This language, the national language of Thailand, is often simply referred to as Thai, but that terminology may be confusing given that the languages being considered here are called Tai. In India it is referred to as Bangkok Thai.

In this section, the approaches of Noss (1964), Panupong (1970) and Lekawatana (1970) will be discussed, since they are referred to in the syntactic analysis of the Tai languages of India (see below Chapter 8).

4.2.1 Noss (1964)

Noss (1964:1) outlines his approach as "to outline the main structural features of standard spoken Thai, ... and also to elaborate by sub-classification and example those structural features which are least covered by existing grammars and dictionaries."

He gives considerable and detailed information about the phonology of the language, including a very detailed discussion of the prosody of Thai, defining prosodic phonemes (1.6), classed as stress phonemes – loud, normal and weak onset, in combination with sustained or diminishing stress; intonation phonemes, identifying eight possible intonation sequences for a single utterance; and rhythm phonemes, which apply at phrase level.

To give an example of the stress phonemes, he claims that the word /pay/ 'go', can have the following meanings which are differentiated only be a combination of stress phonemes:

1) loud diminishing stress   /!paj/   'Let's go'
    loud sustained stress   /!paj:/   'Sure (he) went)'
    normal diminishing stress   /'paj/   'Yes, (he) went)'
    (Noss 1964:21)

It can also be argued that the difference between these three utterances very much depends on the circumstances in which they are uttered, on the context of the utterance; but this is not a concern of Noss. In the present thesis this level of prosodic analysis has not been possible.

The largest portion of Noss (1964) deals with lexemes (words), and how they fit into the structure of the language. He divided lexemes into two basic classes, free lexemes

and bound lexemes. His bound lexemes are defined at as occurring "as constituents of syntactic units and higher-order constructions only." (1964:79). The term *bound lexeme* does not refer to the kind of bound morphology found in the various agglutinating or polysynthetic languages of the world, and indeed many of Noss's bound lexemes have the same form as a free lexeme, but are unstressed.

For example /pay/, which in (1) above was clearly a free lexeme, a full verb meaning 'go', also appears as a modal verb, a sub class of bound morphem, with the meaning 'to act away from the speaker, or so as to affect the speaker and his group'., as in (2)

2) yà-pay kʰian bon kradaan-dam
    Don't write on that (or their) blackboard
    (Noss 1964:135)

As will be seen in section 8.5.7 below, the Tai languages of Northeast India show analogous processes. The word *kaa¹/kā¹* is both a full verb meaning 'to go', and a word which has undergone grammaticalisation to be a TAM marker, expressing past time. It is a contention of this thesis that regardless of the level of grammaticalisation, the word is not fully bleached of its meaning, and that the two will not therefore be regarded as separate lexemes, in the way that Noss regards /pay/ and would regard *kaa¹*.

Diller (2001:146) assesses Noss' methodology as follows:

> "Studies in Thai in what is generally referred to as the structuralist tradition have emphasized synchronic form classes as determined by sentence test frames. In this tradition, frames are typically taken *a priori* as the main criteria in the setting up of a system of syntactic classification. This may lead to "structural homophones". If a form occurs in each of two test frames set up to distinguish form classes, then even if there is a strong semantic motivation for considering the form a single word, the guiding assumptions normally require it to be taken as two separate lexical items."

4.2.2 Vichin Panupong (1970)

Where Noss concentrated on words and how they fit into different constructions, Vichin Panupong's (1970) grammar of Thai is based around the concept of a constituent. She introduces her work by defining the basic sentence types as initiating sentences and non-initiating sentences – where the latter is "one which cannot start a conversation" (1970:1).

She exemplifies this as follows:

> "For example, **mây mi:** ไม่มี (No, there isn't) is to be regarded as a non-initiating sentence on the ground that it cannot be fully understood without some such preceding utterance as **mi: khon máy** มีคนไหม (Is there anybody there), which serves as a verbal clue."

This is an important insight into the languages of the Tai family. Tai languages permit the deletion of noun phrases such as *khon* in the non initiating sentence in Vichin's example (*mây mi:*).

An initiating sentence does not necessarily contain all of the noun phrases which the verb subcategorises for, as in the following example of an initiating sentence:

3)

ที่นี่	ทาน	ข้าว	กี่	โมง	อะ
thî nî:	**tha:n**	**khâ:w**	**kì:**	**mo:ŋ**	**há?**
here	eat	rice	how many	o'clock	PRT

"What time do you have dinner here?
(after Panupong 1970:2)

The subject is not stated; it is not required because pragmatically it is understood to be the interlocutor, even in the initiating sentence of a conversation.

Vichin Panupong goes on to discuss a range of sentence types and the constituents of those sentences. Her definitions of subject, object and indirect object are primarily structural, that is a subject is a "noun in its function as a sentence constituent of the sequence nv (noun-verb)" (1970:13), that is to say a noun or noun phrase which

precedes the verb, An object, on the other hand, is a noun phrase which follows the verb, and the indirect object is "the third noun of the sequence nvnn."

The canonical structure for a sentence, in Panupong's analysis, is therefore:

4)   (Subject) Verb (Direct Object) (Indirect Object)

The verb in a sentence which has no object is called an Intransitive Verbum (V_i), that which has a single object is a Transitive Verbum (V_t) and that with two objects a Double Transitive Verbum (V_{tt}).

Furthermore, where the constituents contain more than one word, one constituent may be anteposed. (1970:15)

This leads to the following list of possible structures:

5) SV_i, V_iS, SV_tO, V_tO, OSV_t, $V_{tt}OI$, $SV_{tt}OI$, $OSV_{tt}I$,

Two of these structures are similar, being both V N - namely V_iS and V_tO. They are different because the first type may be reversed to SV_i, whereas the second cannot be. Panupong gaves (6) as an example of V_iS and (7) as an example of V_tO:

6)

เจ็บไหม	คุณ
cèp máy	**khun**
pain-QN	you

"Are you in pain"

7)

ปวด	หัว
pùat	**hǔa**
ache	head

"(I've) got a headache."

Example (6) could also be expressed as *khun cèp máy* whereas * *hǔa pùat* is impossible.

Throughout this work Vichin Panupong gives lists of the possible arrangements of phrases, such as the possible combinations of items within a noun phrase, within a classifier phrase and so on. She does not discuss the context in which such structures might be used.

Another important insight of Vichin Panupong's is what she calls discontinuous sentence constituents. She suggested that:

"It is frequently necessary to postulate "discontinuous" constituents when the elements which make up the constituents are interrupted by another constituent in the sentence" (1970:20)

Both verb complexes and noun phrases can appear 'discontinuous' in surface realisation, as in (8):

8)　เพื่อน　　ยืม　　　หนังสือ　ไป　　เล่มหนึ่ง
　　friend　borrow　book　GO　CLF-one
　　S　　　V$_{t}$-　　O-　　(-V$_{t}$)　(-O)
　　"A friend's borrowed a book."

In the Tai languages of Assam, structures similar to (8) are not uncommon. They will be discussed in section 8.3.

4.2.3 Pongsri Lekawatana (1970)

Pongsri Lekawatana (1970) is a study of Thai based on the theories of Fillmore (1968). The following assumptions are basic to her analysis:

a) each case relationship occurs only once in a simple sentence (Fillmore 1968:21)

b) only noun phrases representing the same case may be conjoined (Fillmore 1968:22)

c) complex sentences involve recursion through the category Sentence under the case category Objective (Fillmore 1968:32)

d) every noun phrase begins with a preposition or a case marker(Fillmore 1968:33)

Fillmore proposes that a proposition is made up of a verb and cases, once a verb has been chosen, its case frame must be filled with noun phrases marked for those cases.

Pongsri Lekawatana has studied all of the possible cases in Thai according to this approach, discussing, for example, "verbs that take the agentive". Some of the features of such verbs are that they

• can occur in command imperative

• can occur as complements of verbs like สั่ง sàŋ "order"

• can co-occur with an objective or experiencer as in (9)[34]:

9)　แดง　　ฆ่า　　นก　　ตัว　　นั้น
　　dæɛŋ　khâa　nók　tua　nán
　　Dang(A)　kill　bird　CLF　that (E)
　　"Dang killed that bird"
　　(after Lekawatana 1970:48)

Lekawatana's analysis of the dative is also very useful for the analysis and is discussed below in 8.3.3.3.

Lekawatana also introduces another interesting proposal. Most of the oblique cases are marked by prepositions, but there is a subjectivalisation rule which moves one case to the front and triggers the deletion of the case marker. This has been summarised in (10), based on Lekawatana (1970:172-173), where the principles identified by Lekawatana are also exemplified with English glosses.

10)　if there is an Agentive, it will be in subject position:
　　　　as in　Suda (A) eat banana (O)
　　if there is no Agentive, the Instrument becomes subject
　　　　as in　Fire (I) burn factory
　　the experiencer becomes the subject if it co-occurs with objective
　　　　as in　Suda (E) see Dang

---

[34] The Thai script has been added by me.

the objective becomes subject if is by itself
　　as in　Window (O) open
or with a material instrumental
　　as in　This dress (O) is dirty with paint (Im)
the factitive becomes the subject if the agentive is unspecified at surface level
　　as in　house cl this (F) build when year that last (T)

The approach in (10) may prove useful as the Tai languages are studied in more detail, but a considerably deeper analysis than is possible in this thesis would be necessary.

4.2.4 Previous research on Shan

As discussed above in 1.2, the Tai languages of Assam are very closely related to Shan. The term Shan can be used to identify the subgroup of languages of which the Tai languages of Assam are part. Alternatively it is used to refer to the language of the majority of the population in Shan state, Myanmar. Young (1985:1-3) briefly discusses what is meant by the term Shan.

There is no modern linguistic grammar of Shan available in the English language, although the language was described in some detail by Cushing (1888) and Glick and Sao Tern Moeng (1991). The latter is a teaching grammar of the modern language of Shan state and the diaspora of Shan speakers who have fled the government of Myanmar.

The *Elementary Handbook of the Shan Language* by the Rev. J.N. Cushing (1888) contains a significant amount of data on the language. Cushing developed a system of marking the vowels and tones in his transcriptions, using numbers for the tones and the letters *c* 'close', *m* 'mediate' and *o* 'open' to identify vowel height, as in လုမ် lōm, 4o 'the wind'. He also presents several texts in Shan language and an English–Shan vocabulary.

Young (1985) presents a number of texts in Tai Mau, a variety closely related to Shan and the Assamese varieties. She introduces her book by noting that it was "the first major collection of materials made in recent decades within an area on which, in

general, very little has been written" (1985:1). Young presents a brief sketch grammar including a description of the writing system of Tai Mau. The bulk of her book is the presentation of a number of texts, in analysis and translation. Several of these have parallels with the texts presented in this thesis. Her first text, *Revenge* has a similar material to that found in *Story of a new king* (1-1-3-4).

Glick and Sao Tern Moeng (1991) is a teaching grammar of Shan. They give very comprehensive rules for the use of particular grammatical structures, as exemplified in Table (2), which explains the various permutations of TAM marking.

(i)	unmodified verb	may indicate that the action or description is either in present or past time
(ii)	V + yaw.	action has already been performed, or that a condition has been achieved. This pattern does not indicate whether the goal, if any, of the action has been achieved.
(iii)	V + yaw. + yaw.	action has definitely been performed, or a condition has definitely been achieved. The second yaw. emphasizes completion of the action of achievement of the condition
(iv)	V + yuu,	means that an action is still happening, or a condition still exists (this can refer to the present or the past)
(v)	te + V	means that an action will be taking place, or that a condition will be reached. te is also used to indicate potentiality
(vi)	te + V + yaw.	says that an action is about to take place at any moment, or just about to take place at the time of the speech, or that a condition is just about to come into being.
(vii)	te + V + yuu,	says that an action will definitely take place, or that a condition will definitely be reached.
(viii)	te + V + yuu, + yaw.	says that something will have already taken place, or that if the situation does not change, then the action or the description indicated by the verb will be taking place
(ix)	am, + V	the negative particle may be used before any verb or adjective
(x)	am, + V + yaw.	means that the action will not take place, or a condition will not be reached. It is similar in meaning to the English "not ... anymore", or "not ... after all."
(xi)	am, + V + yuu,	DOES NOT EXIST IN SHAN LANGUAGE
(xii)	te + am, + V	means that an action will not be taking place, or that a condition will not be achieved. It is like am, + V but refers to the future.
(xiii)	te + am, + V + yaw.	says that something will not be happening, or that a condition is not going to be reached
(xiv)	te + am, + V + yuu,	Similar to the previous. This must be either a phrase in a sentence beginning with the shan equivalent of "if" or "when", or else must be a response to a preceding sentence. te + am, + V + yuu, may be used where this is used, but the reverse is not always true

Table 2: TAM marking after Glick and Sao Tern Moeng (1991:41f)

As can be seen from the above, Glick and Sao Tern Moeng do not use the Shan script in their grammar, although it is used in the comprehensive index and some of the teaching texts appear in script at the end of the book. For the bulk of the book, they use the phonemicisation shown above, where the symbols <>, <ʼ>, <:> and <;> are used as tone marks just as they are in the Shan script.

As a teaching gra:::: :r, Glick and Sao Tern Moeng (1991) is not designed for use by linguists wishing to (conveniently) discover insight into the structure of the grammar. Fortunately every word used in the text appears in the index at the end which refers back to any mention of that word. A good deal of very valuable data is presented, in part from the author's own native speaker knowledge.

Finally, the *Shan–English Dictionary* (Sao Tern Moeng 1995) has been an invaluable aid in the preparation of this thesis. Every single word which does not appear in Banchob (1987) or in the Aiton word list elicited from Bidya Thoumoung (see below 10.1.1) has been checked in the *Shan–English Dictionary*.

4.3 Native Intuitions and Native Pedagogy

In addition to the work of previous scholars on the Tai languages (see above Chapter 3), and the work of scholars on related languages of the Tai family (see above 4.2), this thesis has also been informed by the knowledge and insight of the native speakers of Tai into their own language. Sometimes these insights are informed by Grierson (1904), the only reference widely available in Assam.

A very good example of the native speaker intuition as to the structure of the language was the late Aimya Khang Gohain's insistence that these languages were not verb final. In 8.4 below, the basic order of constituents (phrases) in the Tai languages is discussed, and the claims that they have become verb final, or specifically SOV, are explored.

It is appears to be true that a much larger proportion of utterances in the Tai languages of Northeast India exhibit verb final structures than is true for other languages of the family, such as, for example, Standard Thai. Yet Aimya Khang Gohain, and other Tai speakers, insisted that for this thesis I should make the clear statement that SVO is the basic word order of the Tai languages, as it is of the Tai family in general.

How are we to explain this insistence, if not in terms of some kind of understanding of the Tai language? Many previous researchers have concluded that Tai has become or is becoming SOV, but the people themselves do not feel that, even when they utter an SOV sentence.

Native speaker understanding of the language goes beyond intuition into native pedagogy. There are no grammars of the language written; there is, as yet, no monolingual dictionary, but there are pedagogical traditions.

For example, in the translation of Burmese loan words, devices such as (11) can be used:

11) ကျော် ဝ် ဝ§ ဟုင် ကျော် ပွန် ဆေ ပွန်
kyɔ¹ cam⁴ wā⁵ ṭ⁵ hoŋ⁶ kyɔ¹ pon⁴ se⁶ pən⁵
famous PRT PRT famous famous exceed COMP* other

နတ် လူ ဝ လဲ ကျွန် ။
nat¹ lu¹ wā¹ lāi⁶ kyun²
god village** many CLF of place
'The word *kyɔ¹* means famous, the fame of one exceeding all other gods and humans in every place.'
(Tai Phake Manuscript *Explaining the meaning of words to children*, No. (4), read by Ee Nyan Khet)

A monolingual dictionary is not necessary for most Tai people, but there is a need to have words such as *kyɔ¹* explained, because it is a Burmese loan word and its meaning is not known to all people.

4.4 Prescription and description.

A very basic principle of modern linguistics is sometimes called the "descriptive approach". The fact that linguistics is about describing the system of language as it is (descriptivism), rather than making any statements about what ought to be (prescriptivism), is emphasised to linguistics students[35].

Pure description, without any regard to community's perception of what is prestigious or correct, may not produce a grammar which is a true representation of the language. What happens when a sentence which is non-standard or even non-grammatical is nevertheless uttered in one of the texts? If there is no standard against which to measure the sentence, how are we to know?

One might collect a sentence like 'I want to talk about a guy that I knocked on his door of.' This sentence would not be accepted as grammatical by most English speakers in Australia[36]. It was spoken by a respected member of our community, an Australian Member of Parliament[37]. If a similarly ungrammatical but interesting sentence was spoken by a person of similar status in the Tai community, it is almost certainly included in this thesis.

It may not only be for the benefit of the linguist, but for the benefit of the speakers of the language that a prescriptive grammar should be done. The role of the linguist in this process is unclear; because of the small number of speakers of such languages, and because of the socio-economic conditions in which most of them find themselves – speakers of small languages are rarely amongst the wealthy classes of a society – the linguistic description which researchers like myself write may well become a de-facto prescriptive grammar for that community.

If, in one or two generations, these languages begin to slide into rapid decline, then the community will inevitably look back to this thesis, in the way that they now look

---

[35] In my case, it was emphasied in the very first linguistics lecture I attended at Monash University, given by Dr. Peter Kipka.
[36] It is acceptable according to my computer, however, unlike the sentence that this footnote is attached to.
[37] The practise adopted in this thesis of identifying all the informants has not been followed in this case.

to Grierson (1904), for information about their language; they may even begin to trust the written words more than their own native speaker intuitions.

Therefore, it is essential that scholars always bear the native speaker community in mind when writing works on small languages. The English speaking world as a whole barely notices the discussions which linguists indulge in about small pro or the syntax of relative clauses in English. For a small language, however, everything written acquires great importance.

In this thesis, where any prescriptive norms have been pointed out, they are noted. A good example is the possessive construction explained in example (207) in section 8.5.2. Aithown Che Chakap indicated that one structure is preferred for inalienable possession and another for alienable possession. Both structures are probably available for both types of possession, but for him at least, there is a clear preference.

4.5 The importance of the Tai Script

A clear decision has been made in this thesis to use the Tai script for language examples, both in order to make them more accessible to the Tai community and to stress the importance of the native orthography.

It is worthwhile digressing for a moment into history. When Sir William Jones' collected works were published (1799), scholars commonly wrote and understood several classical languages. Let us consider a footnote to his *Grammar of the Persian Language*:

> "and Apoftolius compiled an Ἰωνιά زار بنفشه *a garden of violets*, or a collection of proverbs and fentences." (1799, II:205)
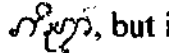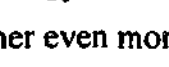
It was not unusual to switch into Greek or even Arabic, and the reader was apparently expected to be able to read the scripts and perhaps understand the languages. Even to this day, scholarly works which slip into French or Latin, untranslated, are still to be found.

However in the 20th century a new tradition of talking about languages without the use of native orthography arose. One can see that this might have arisen when the multitude of scripts in which the languages of Southeast Asia are written became apparent, and the costs of printing in these scripts became manifest.

However there are several serious objections to this. Firstly an ethical one: native speakers of these languages will scarcely be able to read anything about their language if their script is not included, they will have real problems reading any phonemicisation, whether it be the relatively user friendly Banchob (1987) orthography (see 6.2)or the somewhat less accessible realisation of Khamti in Weidert (1977).

Since these are living languages, the heritage of human beings, it is not acceptable to deny to the native speaker community access to the information collected about them.

The second objection relates to the semantic content of the symbol. There are times when the script conveys a message that no transcription can convey. The Tai languages of Assam have a vocative particle, /ɔi²/ in Phake, which also has a ritualistic function in some texts.

This particle can be written as ᩁᩮᩬᩲ, but in texts it often appears in very flourished form, as ᩅᩮ or ᩅᩮ or some other even more ornate form. This communicates the ritualistic nature of the particle, which the phonetic transcription cannot show.

Another example is the subscripted consonants /r/, /y/ and /w/. Each of these consonants can be the second member of an initial cluster in Aiton, with a stop as the first member (see below 6.3.3.1). But when they are written, they appear in a different form, as in Table (3):

phoneme	initial consonant form	subscripted consonant form	example		
r	ရ	ြ	ဖြာ	phraa²	Buddha
w	ဝ	ွ	ဖွာ	phwaa²	birth
y	ယ	ျ	မျတ်	myat¹	happiness

Table 3: Comparison of initial and subscripted forms

The approximants in word initial position and in clusters are pronounced slightly differently. Indeed it is possible to argue that Tai initial clusters should be treated as single phonemes, analogous to the affricates in English. The subscripting of the consonant, in the way that vowels are subscripted, shows this. The phonetic transcription, on the other hand, cannot.

A third objection is that the orthography sometimes affects the language. Among the Tai scripts we see that there is no initial <d> or <b>, yet /d/ and /b/ are phonemes in Aiton. However, they are in variation with /n/ and /m/, with which characters they are usually written. This variation may have arisen, at least in part, as a result of the writing system.

Becker (1993) has put the issue of the use of native orthography clearly before linguists: in a passage which he headed "Defamiliarizing Burmese", he wrote that:

> "transliteration of this passage into roman writing is not a 'meaning preserving' act, for the writing system which one uses shapes the way one imagines one's language and thinks about it. ... writing systems (and other systems of representation) are among the deepest metaphors in a language ... and so for us to substitute one technology of writing for another is not a neutral act ... It means to reimagine language itself." (1993:63)

Glossing a language is, to Becker, "clearly a political process", and one in which he argues languages do not meet as equals. He goes on to add that putting Burmese into

Roman tr.·· literation obliterates the very deep metaphor of centre and periphery in writing, adding that

> "Much traditional philology and modem linguistics depends upon this romanization as a first step (even before glossing) in analysis. The illusion is that nothing important is lost." (1993:64)

I find it greatly reassuring that there is knowledge of language which cannot be gained from the examination of a Roman, or IPA, transcription. Since we do not wish to lose any more information about the language than is absolutely necessary in the process of translation, the script must remain.

This is not to criticise those scholars of the past who were unable to use native orthographies in their work, because of the technical difficulties that then obtained. Those technical difficulties apply no more[38], and it is a contention of this work that native orthographies should be present in all linguistic work.

4.6 The presentation of text examples.

The linguistic method of text analysis followed in this thesis presents each example in four lines. The first is the script (see above 4.5). The second is a phonemic transcription, the third an interlinear morphemic gloss and the fourth a translation.

The phonemic transcriptions used in thesis are discussed at length in Chapter 6. They are, however, not ideal. Writing of his comparative study of several Southeast Asian languages, Enfield (2000:62) admitted that "Transcription systems used for these languages should not be regarded as final, but for purposes of this work they are satisfactory."

A transcription system regarded as final, in Enfield's terms, is not necessary if there is a practical orthography. In general a transcription is really only necessary for those

---

[38] This is not to say that there is not some extra effort involved. Considerable time was spent in this thesis on the production and refinement of the Tai fonts, (see 7.8). However fonts now exist for the practical orthographies of most of the languages of the world.

readers who do not know the practical orthography. However, it must be said that in the case of the Tai languages of India, the phonemic transcription is essential because of the problems of the orthography (such as underspecification of vowel contrasts, see 7.2, and lack of tonal marking, see 7.5).

The interlinear morphemic gloss presents further problems. Enfield raised some of the concerns; writing that there is

> "a perennial difficulty when dealing with morphemes which do not have ready one-word translations into English. Over-general glossing can obfuscate important and relevant information for the reader. An example is the extensive used of 'TAM' as a direct gloss, neutralising in a single term the substantial distinctions between the semantic fields of tense, aspect and modality, as well as many conceivable distinctions within each of those." (2000:9)

In an attempt to overcome some of the difficulties raised by Enfield, in this thesis those morphemes which are grammaticalisations of content words will be notated using upper case letters, but with their original meanings, for example $k\bar{a}^1$ 'GO', $m\bar{a}^2$ 'COME', both of which are TAM markers. When they are found as main verbs, they will be notated as 'go' and 'come' respectively.

Some important morphemes in this thesis will be glossed using neutralising terms like 'PRT', here used for 'particle'. This is not ideal; as Enfield points out, this class of words are "sentence-final illocutionary particles which are a semantically rich and yet poorly understood semantic field across these languages." (2000:9). In many cases the full semantics of these 'particles' is not yet understood and this is the reason for the glossing 'PRT'.

The final line of the examples is the translation. In many cases the need to maintain some of the syntax of the original leads to English which is not particularly idiomatic. Translating texts line by line is a different task from translating a whole; both have been attempted here, but the size of the task is such that in some cases the best possible result may not have been achieved.

Becker (1993) discusses in detail what he calls the the exuberancies (adding to what we hear and read) and deficiencies (taking away from what we hear and read) of translation, by referring to (12)

12)   ကား
       hta:
       put

He describes this as follows:

> "The point here is not to argue, for instance, *hta:* has a different range of
> meaning from English "put", although that is certainly true. It is rather that
> many words of the English translation have no counterpart in the Burmese text
> at all, words like: "be", "as", "we", "a", "is", "I" and past tense. I don't think it
> is possible to say that these words are somehow "understood" from the context
> in the Burmese text. They just aren't there." (1993:65)

4.7 Ethical considerations

As pointed out at the opening of Chapter 1, one of the main considerations in this thesis has been to offer some support to the community of Tai speakers in Assam. This community very much wants to conserve its language and linguistic heritage. Some important ethical issues have arisen in the process of undertaking this thesis.

This discussion does not propose to enter into the issues of consent, privacy and protection of anonymity that take up a good deal of discussion when dealing with the ethics of some fields of research, such as medical research. Rather, this discussion will concentrate on how this thesis relates to the Tai communites and their wishes for the survival of their language.

The presence of language researchers is not a neutral act. When we work in a community, we are not working in a vacuum. We are dealing with the very real languages of very real people, who are affected by the presence of the researcher, as

well as by many other pressures from outside. For many people in the Tai villages, my visit is an opportunity to practise English, and some ask for gifts or assistance to help them to better themselves in the society in which they live. I can hardly decline either request, yet in doing so, I am playing a part in changing their culture.

Mühlhäusler is of the opinion that

> "the decline of Pacific and Australian languages, ... is due primarily to the loss
> of their ecological support system (language ownership, cultural practices,
> speaker's lifestyles, settlement patterns, speaker's physical and spiritual well-
> being) and their functional relationship with other languages (language chains,
> bi-, dual- and multilingualism, sign languages, pidgins etc.)." (1996:323)

Many of the items listed by Mühlhäusler as 'ecological support system' are challenged by the presence of foreign researchers. For one, the whole concept of language ownership is challenged if most of the recordings and the copyright of the most comprehensive grammar of a language are held by a person who lives in far-away Australia.

Mühlhäusler further points out (1996:144) that the languages of the Pacific under the impact of contact with Christianity and European languages, have moved towards a

> "...Standard Average European typology, particularly in the semantic domains
> of time, being, causality, spatial organisation, nature and nurture, human
> relationships and emotions."

In my researches into the languages of the Tai of Assam, by the very questions that I ask, and the gifts that I bring, and certainly by the things that I write here, some effect on those small languages is certain. Mühlhäusler discussed the effects of missionary policies and quoted from Codrington and Palmer (1896:VIII), who said:

> "It is probable that some corruption of a native language is inevitable in
> Mission work."

Can this not also be said for linguistics? One hopes that it is less true of linguistics, perhaps because university linguists often do not spend as many years in the field as missionaries do. Nevertheless, it is surely the case that o⁻ presence and our work will have effects on the community, and therefore on the lang ge, that we may not intend. These effects are not necessarily good or bad, but they must be taken into account.

Rhydwen (1996) further reminds us of the human aspect of this type of work:

> "My original research aims were subverted in response to my interactions with Aboriginal people who, unlike the abstraction 'objects of research' are real people with whom relationships develop. Nevertheless, I was still constrained by my obligations to funding bodies and the formal requirements of writing a thesis." (1996:ix)

This present work has also undergone a similar 'subversion', in that the needs and wishes of the Tai community to have their culture and literature recorded has become a more pressing aim than I initially believed it would be.

Rhydwen also discusses the problem of literature and its use in academic work (1996:172), and the copyright thereof. Having read aloud a paper which included a story, she was asked by her informant (an Aboriginal Australian): "Who owns stories, you know that thing you got, copyright intit? Who got copyright of my story?". After answering that she would not in fact receive any money for publishing the story, Rhydwen admits that it is not the point, because

> "How can I explain, how could I say that her story is not of interest in itself, that it is my interpretation, my linguistic analysis, that is the point, the meaning of the text?"

The recording of traditional literature must be accompanied by a respect for that literature as literature; and a respect for the fact that it is important not just to the community but to the whole of human culture that these literatures be recorded for their own sake, not just for the linguistic information we can mine from them.

4.7.1 The approach in this thesis

Bearing in mind the issues raised by Mühlhäusler, Rhydwen and others, this thesis will take into consideration the effects which it might have on the languages and the informants.

I believe it is of the utmost importance that the sources of linguistic information be identified and named. There are communities and situations in which the identification of informants would be inappropriate or may create difficulty for those individuals. This does not apply among the Tai of Assam; there is no taboo on naming the dead and some of my elderly informants are enthusiastic about recording their language, in the knowledge that those recordings will survive them. Examples (13) to (15) are extracts from a conversation between Sa Cham Thoumoung, an Aiton in his eighties, his daughter Silawa, and myself:

Silawa

13.1)

khaam²	mau²	an²	caai²	sa¹ hap¹	hau²
word	2Sg	CLF	male	Sahib	1Pl

ta¹	tɛp¹	au²	pai²	tii²	ɔs¹ trɛ¹ lii¹ yaa¹
will	tape	take	go	to	Australia

'Our Sahib will tape your words and take them back to Australia.'

13.2)

ka¹ laŋ¹	mau²	pɔ²	taai²	kaa¹
after	2Sg	PRT	die	GO

khaam²	mau²	an²	ta¹	kut¹	wai³	kaa¹	nau²	tɛp¹
word	2Sg	CLF	will	remain	KEEP	GO	in	tape

'After you die, your words will remain on the tape.'

SaCham

13.3)

ur²	kau²	kɔ³	saa¹ thɔ²	naa¹	caŋ¹	...
yes	1Sg	LINK	grateful	PRT	then	...

'Yes, I am pleased!' ...

Audio Link for example (13) can be found at \wave files\aiton\extracts\2-9-2-6.

(14)

ဝ့်	ကၠ	ကၟာ်	ဆင်	ကိုတ်	ဒ	ကၠ
khaam²	kau²	an²	saŋ¹	kun¹	wai³	kaa¹
word	1Sg	CLF	if	remain	KEEP	GO

ကၠ	ဟ	ကၠ	ဝ့်	ကၟာ်	တၠ	ကိုတ်	ဒ	ကၟာ်	ဇ
kau²	taai²	kaa¹	khaam²	an²	ta¹	kun¹	wai³	an²	dii²
1Sg	die	GO	word	CLF	will	remain	KEEP	CLF	good

'If my words remain, though may I die, my words will remain. that will be
very good!'

Audio Link for example (14) can be found at \wave files\aiton\extracts\2-9-2-6.

15.1)

ကိုဒ်	ဝ့်	ကၟာ်	ကိုင်	တၠ
ɯ²	khaam²	an²	uŋ²	taai²
YES	word	CLF	NEG	die

'The language will not die.'

15.2)

ကၠာ်	ကၟာ်	တၠ	ကၠ	ဝ့်	ကၟာ်	တၠ	ကိုတ်
kun²	an²	taai²	ko³	khɹam²	an²	ta¹	kun¹
person	CLF	die	LINK	word	CLF	will	remain

'People will die but the language will remain.

SM:

15.3)

ဇ	ၣ
dii²	naa¹
good	PRT-QN

'Is it good?'

Sa Cham

15.4)

ကိုဒ်	ၣ	လၠင်	ၣ	လၠင်
ɯ²	dii²	luŋ¹	dii²	luŋ¹
YES	good	big	good	big

'Yes, very good.'

Audio Link for example (15) can be found at \wave files\aiton\extracts\2-9-2-6.

The full acknowledgement of the informants is both important and appropriate for the community, as well as for scholars. In this thesis, every sentence example will be sourced to the text from which it comes, and the author or teller of that text.

By acknowledging the informants in this way, we are recognising that their contribution to the body of knowledge presented in this thesis is on an equal footing with the many academic texts quoted or referred to here. We are treating the oral texts as we would a book, and quoting and acknowledging the informants in the same way as we would quote and acknowledge a professional linguist.

Furthermore, by presenting the texts in the electronic appendix (see 9.5 below), we are giving them the status of a published text. This is clearly an advantage from the point of view of scholarship, because it allows for the checking of any of the claims made in the thesis. In addition, it is a method by which the oral scholarship of the Tai community can be presented.

The Tai people themselves greatly value many of the texts that I have collected. When I recorded informants speaking a story, they not only knew that it would be used for the purpose of study (see examples (13)-(15) above), but they expected that I would make these texts available.

As I moved from village to village, people were often keen to hear some of the texts recorded in other villages. For example, a number Tai people, such as Sa Cham Thoumoung (Aiton), Sa Myat Chowlik (Khamyang) and Ee Nyan Khet (Phake) can tell Mahosatha stories (see below 9.2). Each of these was very interested in what the others were saying.

In previous work on the extinct languages   ctoria (Bowe and Morey 2000), it became clear how important the knowledge o₁ where information came from can be to a community that has lost its language. In the future, the Tai community will be able to use this thesis to look back on some aspect of their language, and to hear again the stories that people like Sa Cham Thoumoung told them.

In addition to the proper acknowledgement of the informants, there is an ethical requirement to return to the community whatever knowledge has been gathered from them. This, of course, includes the making available of the text analyses to the informants, and making copies of tapes and videos and printed articles which are written as a result of this study. It also includes the making available of this thesis.

Making this thesis available is not just sending a copy to India, it is also making this text as friendly to Tai readers as it can be.

4.7.2 A note on bad language and swearing

I have not been able to record anything by way of swearing, oaths or bad language, save only the single example given in Chapter 8.3, example (135).

While, no doubt, the recording of bad language would greatly add to the linguistic diversity presented here, and while it would be interesting, it is not possible in this thesis to present any such data. All my Tai informants were united in wishing that no 'bad language' be included in the thesis. If a community does not want to have any bad language recorded, then the same ethical reasons which protect the names of the dead, or some types of sacred texts in other communities, apply.

5. Methodology - collection, analysis and translation of the texts:

5.1 Theoretical considerations

The two main methods of obtaining the linguistic data on which a thesis such as this can be based are elicitation and the preparation of a corpus of text.

The very nature of this type of research almost always involves the significant use of elicitation, particularly in the initial stage of studying the language (Payne 1997:370). In this initial stage, the researcher will usually not yet have sufficient competence in the language to be able to make use of large amounts of textual material, and this was certainly the case with the writer of this thesis.

However, as the language competence increases, the researcher is likely to be able to make more use of texts of different types, and those texts are more likely to be given by the informants.

The example of my work with the Khamyang informant, Chaw Sa Myat Chowlik, will illustrate this point. On my first visit very brief visit to Pawaimukh (13th January 1998), I was not yet able to converse in Tai. Using the late Aimya Khang Gohain and Mr. Bibek Borgohain as my interpreters, I elicited a tone box (see 6.1.5.3) from Sa Myat, and some simple sentences.

Chaw Sa Myat was then asked if he could tell a short story perhaps a children's story. He replied that he did not really know any stories, although the text *Advice to Children* was collected. I left Pawaimukh that day with the impression that there was perhaps not very much information that could be collected there.

In the succeeding years, as my competence in Tai has reached the point where I can converse freely with Chaw Sa Myat, he has volunteered a large number of very long

texts, and has shown himself to be most knowledgeable in the traditions, literature and culture of the Tai[39].

Why then, did Sa Myat not volunteer a story in 1998? The answer may be that to him it would seem pointless to give such a text to a researcher who could not understand at least some part of it.

Thus it may be said that for this thesis, the texts have increased in importance viz a viz elicitation as the project has developed.

Payne (1997:366) opens his discussion on the issue of the relationship between elicited and text data with the following comments:

> "Both text and elicited data are essential to good descriptive linguistics. Each have advantages and disadvantages. The linguistic researcher needs to be aware of these in order to make the best use of all the data available."

In discussing the relative position of elicitation in a descriptive linguistics project, Payne concluded (1997:368) that

> "The controlled, systematic, and rule-dominated parts of language are best approached with an emphasis on elicited data. These would include:
>
> 1 phonology (excluding intonation);
> 2 morphophonemics;
> 3 inventory of derivational morphology (which derivational operations apply to which roots, etc.);
> 4 inflectional inventory (determining the ranger of inflectional possibilities for person and number "agreement" and case marking);
> 5 pronoun inventory (isolating the entire set of free pronouns);
> 6 lexical inventory (acquiring the words for a large number of culturally significant things and activities)."

---

[39] I have often referred to him as *Sa Myat Bor Pandit*, 'Sa Myat the great knowledgeable man'. I have not met any Tai person in Assam who knows more about a greater range of cultural practices and the texts, oral and written, that relate to them.

According to Payne, text data will be made use of more for analysing "the more pragmatic, semantic, and subtle parts of language" (1997:368f). These include:

> " 1 intonation;
> 2 constituent order;
> 3 inflectional morphology (determining the precise functions, including tense/aspect/mode);
> 4 voice (alignment of grammatical relations and semantic roles of verbal arguments);
> 5 sentence-level particles (evidentials, validationals and pragmatic highlighting particles);
> 6 clause combining (including relativization, complementation, adverbial clauses, and clause chaining);
> 7 lexical semantics (determining the nuances associated with various lexical choices, including derivational morphology and pronouns)."

For each of the processes for which Payne recommends the use of elicited data, it has been found that reference to the texts considerably deepens the understanding of the languages concerned. Two examples will suffice to illustrate this point.

The Aiton phonology discussed in 6.3 below is largely based on a list of 1900 words elicited from Bidya Thoumoung in December 1999 and January 2000. However a deeper understanding of the extent to which Aiton phonology follows this was only gained by a very close study of the texts that have been recorded. For example, the realisation of the proto Southwestern Tai initial /v/ as /ʒ/ apparent in Bidya's list is confirmed in most spoken texts recorded in the Aiton villages, although not in read texts. The decision to notate this as /ʒ/ consequently relied as much on the textual confirmation of what was collected in elicitation, as on the elicitation itself.

Even items such as the pronoun inventory and lexical inventory could only be done up to a point with elicitation. The word *hā²* '1st dual' has survived in some Phake texts (see *Words of gladness - in honour of Stephen Morey* (1-7-1-11), No. (8)), even though it is bleached of its dual meaning. No amount of elicitation would have led to

this word being volunteered as a pronoun. As to the general lexicon, a large number of words are literary and can only be found in the texts in which they occur. No amount of elicitation will produce them. I believe that a high proportion of the words which will eventually be listed in Tai dictionaries will be words that would not be established through elicitation.

While Payne makes the important point that "text data should be distinguished from elicited data in whatever cataloguing system is employed"(1997:370), he does not go on to suggest that examples in the written grammar should be marked according to their sources. Since elicited and text data are by their nature different, it is maintained here that it is most important to indicate what kind of data is being discussed. For this reason, every Tai language example cited in this thesis will be sourced as either an elicitation, a sentence spontaneously uttered or an extract from a text.

In many grammars that have been written in the past it is impossible to know whether the sentence examples presented as an illustration of some aspect of the syntax were elicited or not. It is the firm belief of the present writer that knowing the source of a particular sentence will help the reader to understand more about the language

5.2 The collection of data

5.2.1 Elicitation undertaken as part of this project

The principal elicitations undertaken for this thesis were:

i)   An Assamese–Aiton word list, elicited from Bidya Thoumoung in December 1999 and January 2000.

This was done by taking each initial letter of the Tai alphabet, and writing down each of the monosyllables which were possible in Tai orthography for that initial letter. Bidya Thoumoung then wrote every Assamese meaning he could think of for each written syllable. These were then recorded and translated into English. This list forms the basis of the Aiton Lexicon discussed in Chapter 10.

ii)   Sentence elicitation.

In the early stages of this thesis, a number of sentences were elicited from both Phake and Aiton speakers, particularly from Chaw Sangea Phalung (Aiton) and Aithown Che Chakap and Aimya Khang Gohain (Phake).

iii)   Discussions of grammatical features.

A number of particular grammatical features were discussed with a number of informants, especially with Nabin Shyam Phalung (Aiton) and Yehom Buragohain (Phake). One example of this is the TAM particle $uu^1/\bar{u}^1$, for which explanations and examples were given, see below 8.5.7.6.1.

The other major task which was necessary for the research on this thesis to proceed was the copying and typing into computer of the *Phake–English–Thai Dictionary* (Banchob 1987). I was very fortunate to obtain a copy of this dictionary in January 1998 in Namphakey village. A large proportion of the data contained in this dictionary was then entered into computer (see 10.3 below) and an English-Phake word finder produced from that data base. This allowed for much speedier learning of the language and analysis of the texts.

5.2.2 Text collection

The texts on which this thesis relies were collected in the period 1998 - 2002. An attempt has been made to collect a representative sample of the kind of texts that are regarded as important to the Tai communities, but inevitably the writer's own interests are reflected in the range of texts presented.

There are a number of reasons why some informants might not feel able to give a particular text at a particular time; some texts take considerable time to speak and that time may not be available. Some may not wish to take the time if they do not feel that the researcher is able to make use of the text.

Sometimes, on the other hand, people came to me wanting me to record a particular text for posterity. Often these were songs and as such not always the best texts for understanding the syntax of the language, since the songs almost invariably use either archaic syntax or syntax that allows variation not permitted in spontaneous speech.

Another important issue in field work is the need to balance the time taken in text collection with time spent analysing the texts collected. The latter is very time consuming and very much more demanding than the former.

However this research has been important as an archive of the Tai languages as well as an analysis of them. As many different speakers were recorded as possible. Priority was given to certain very knowledgeable elderly informants, such as Sa Cham Thoumoung in Duburoni or Sa Myat Chowlik in Pawaimukh. Whenever either of them had anything to say, it was recorded. Only a small proportion of the total recordings have been analysed for this thesis. The remaining texts still await transcription and analysis.

It may be that old people's language is over-represented in the data that forms the basis of this thesis. Certainly the language of the younger generation - those under 30 - is barely represented here. It will be the task of future research to overcome this shortcoming.

Another aspect was the need to try to achieve some kind of gender balance in data collection. Because the writer is male, and because males are much more likely to be able to read and write in Tai, there is a preponderance of data from male informants. This is particularly so for the Khamyang data. However it is not the case for the Phake, where my principal informant is a female, Ee Nyan Khet, who reads and writes Tai fluently.

Another issue with data collection is finding the right person to give a good example of a particular type of text. In March 2002, I was taken to Tipam Phake village to

meet Ai Than, who was said to be able to chant a particularly fine blessing, the *Prayer of Blessings*.

A number of practical issues also impinged on the data collection process. Minidiscs and video cassettes are not available in Upper Assam, so that sometimes I was running short of space and could not record all the texts that were offered. Furthermore, electricity supply was sometimes unreliable or unavailable, and the recharging of the minidisc player or video camera was sometimes a problem.

5.3 Data Analysis

5.3.1 Prerequisites for analysis

Before the data analysis could proceed, a number of tasks need to be undertaken. The first of these was an understanding of the phonology of the language. In the case of Tai Phake, my analysis of the phonology agrees with Banchob (1987) with minor variations (see below 6.2.4.1). Her phonemic notation was thus accepted and employed in the thesis.

The situation of Aiton was somewhat more complex. Prior to this thesis, no comprehensive study of the phonology of Aiton had been undertaken. It took some time to establish the Aiton phonemic system as discussed in 6.3.

In the case of Khamyang, the problem of establishing a phonemic notation is increased by the tendency of most Khamyang speakers to mix Phake and Khamyang when speaking to me.

A second prerequisite for the analysis was the development of the fonts with which the Tai script and Phonemic transcription could be rendered (see 7.8 below). These went through several versions before the versions employed here were finalised.

5.3.2 Difficulties encountered in analysis.

The greatest difficulty encountered in the analysis was the amount of time required. There are around eight hours of text presented here, and it is estimated that the analyses took around four hours for every minute of text recorded, exclusive of the time spent on the prerequisites mentioned above or the time spent collecting the texts.

Every text has been checked with at least one native speaker, and some of the more difficult texts have been discussed with several speakers. The phonemic transcription and glosses were typed in, rather than generated automatically by a program such as Shoebox. This was both necessary and desirable. It was necessary because the considerable amount of homography in the Tai languages meant that every single grapheme had several possible meanings and the Shoebox system would not automatically gloss any text. Furthermore, using Shoebox would have created a number of difficulties with the various fonts in use.

It was desirable to type the glosses and transcriptions in because this both helped in the process of learning the language, and ensured that every sentence was comprehensively studied. Although this thesis employs computers, at no stage is any computer program making any linguistic decisions on my behalf.

A number of texts present particular difficulties. Example (1) is Yehom Buragohain's translation of a line from Nang Pe's *Blessing*:

1) ⟨Tai script⟩ ‖
cau³ kham² haü³ nai³ lü¹ māk⁴ thāp⁴ tāŋ¹ saŋ¹ khā¹
RESP gold GIVE get offer plenty place place Bhikkhus
'My you get to offer many things to the venerable monks.'
(Phake Blessing *Blessing*, No (24), intoned by Nang Pe)

Yehom Buragohain assumed that the first two words referred to the person who was being blessed, and could perhaps be translated 'oh golden one.'

However this text was rechecked with Ee Nyan Khet, who pointed out that the translation should be as in (2):

2) ⟨Tai script⟩ ‖
cau⁴ kham⁵ haü³ nai³ lü¹ māk⁴ thāp⁴ tāŋ¹ saŋ¹ khā¹
morning evening GIVE get offer plenty place* place Bhikkhus
'Every morning and evening may you get to offer many things to the venerable monks.'

This text is an example of one which took a very great deal of time to translate, as was the case with all of the prayers and other liturgical texts.

Sometimes it was not possible to recheck everything, and there are places in several texts where the translation of a particular word remains unclear.

The writer's own lack of ability in Assamese and imperfect command of the Tai languages also made for difficulties in this regard. In the early field trips, I was reliant on multilingual speakers who knew English in order to translate the texts. Later it became possible to work with non-English speakers like Ee Nyan Khet and to work out the meaning of almost all the words, particularly with the aid of the Shan-English dictionary (Sao Tern Moeng 1995). Even so, some very difficult words may remain unclear.

Another difficulty in transcription was persuading the informants to notate items such as false starts and hesitations. On many occasions there would be such a false start but when listening back to it, the informant would say "Oh, he didn't say anything". For linguistic analysis, however, these false starts are very important. Wherever possible I have notated them when they occur in spoken texts.

I have been able to collect and analyse very little truly natural conversation between two or more Tai speakers. Usually when the tape is on the people speak as if speaking to me. To what extent, if any, the natural spoken form of the language really differs from the analysis in this thesis is not possible to say.

In order to write a truly useful description of these languages, it is essential for the writer to be able to speak the languages being studied fluently. In an Australian PhD, with its allotment of only three years, this is almost impossible. It is certain that some insights into the deeper structures of the language will have been missed. Clearly this thesis can only be regarded as a preliminary study; and the conclusions are only as good as the variety of text which I have been able to collect and analyse.

5.4 Data storage

The recordings collected for this thesis are a most valuable record of the Tai languages, a record which will survive the speakers and survive the writer of this thesis, just as Sa Cham Thoumoung observed in examples (13) to (15) of section 4.7.1. The recordings have been placed in the Department of Linguistics at Monash University. Copies have been made and placed in other safe locations, in case of accidental destruction of such a valuable resource.

A significant proportion of the video tapes collected for this thesis have already been copied onto VHS and sent to the Tai communities in India, as requested by them. It is hoped that in time all of the recordings will be copied and those copies placed in the communities in which the recordings were made. In many cases, printed copies of the texts and analyses were sent to the informant.

5.5 Grouping the texts

For the purpose of this thesis, a numbering system of texts was devised, which is presented in Table (1):

1st number	Language community	(1=Phake, 2=Aiton, 3=Khamyang, 4=Ahom, 5=Khamti, 6=Shan. 7=Turung)
2nd number	Type of text	(1=stories, 2=historical, 3=prayers, 4=Buddhist text, 5=ritual text, 6=other cultural texts, 7=songs, 8=explanations, 9=informal texts, 10=grammar, 11=speeches, 12=miscellaneous texts)
3rd number	Mode of recording	(1=video, 2=audio minidisc, 3=audio casette, 4=manuscript only, 5=other researchers text)
4th number	Number of text	

Table (1) System of numbering texts

The first number gives the particular language community from which the text comes, the second the type of text, the third the mode of recording and the fourth the number of the text.

For example, Text 1-1-1-1 is a Phake text, a story, recorded by video and the first of that type. The document containing its translation is named 1-1-1-1.doc and the wave file is 1-1-1-1.mp3.

There are a number of issues which arise in this process. First of all, all of these communities are multilingual and the first number will therefore not necessarily refer to the language in which the text is recorded. In all of the Tai groups of Assam some mixing with Assamese is to be found, and Pali language is certainly used in religious texts.

With the Turung the situation is even less clear, because the Turung are often trilingual with some speaking Tai fluently as well as Turung (which is Tibeto-Burman) and Assamese. As an example of this, the text 7-1-2-1 (*The old couple without any food*) is a story told by the Turung informant Pong Chap in Turung language, but he also switches frequently into Tai, of which he is a competent second language speaker.

Text 7-1-2-2 is the same story told by Pong Chap in Tai. The initial number 7, therefore, refers only to the community from which the speaker comes, and does not necessarily refer to the language in which the text is spoken.

There has been a greater difficulty in the arbitrary definition of the 'type of text' which is the second number of this system. From my earliest days of working with these communities, the most frequently recorded texts were children's stories and songs.

The other category of early recordings was word lists, short sentences and information about the tones which I have categorised as 'grammatical'.

However it is now clear that there need to be more categories, and the following system has been devised:

Number	Type of text	including:
1	Stories	Children's stories Buddhist Jataka stories (particularly Mahosatha)
2	Historical	Historical manuscripts Oral retelling of histories Oral histories Informant's reminiscences and stories of their own life
3	Prayers	Prayers used in Buddhist ceremonies Prayers used in other ceremonies
4	Buddhist Texts	Buddhist manuscripts, including Mahosatha Oral retelling of Buddhist texts Explanations of Buddhist practice Explanations of Buddhist ceremonies
5	Ritual Texts	Manuscripts relating to non-Buddhist rituals Explanations of non-Buddhist rituals Dances and other texts performed during non-Buddhist rituals
6	Other cultural texts	Manuscripts containing information such as traditional medicines, proverbs, astrology, Explanations of traditional practises such as the process of building a new house, Children's texts such as lullabies, Riddles
7	Songs	Love songs Other songs
8	Explanations	This is the class of text where an informant talks about and demonstrates a particular aspect of cultural practise, such as explaining the ကျောင်လိက် *kyɔŋ² lik³* (or spirit house) in each Tai house
9	Informal Texts	Conversations Personal Letters
10	Grammar	Word lists Discussion about the tones Sample sentences Discussion of grammatical issues in other texts
11	Speeches	Speeches in public places
12	Miscellaneous	Other texts, such as people's names and their meanings,

Table 2: Categories of texts presented in this thesis

These categories are purely arbitrary and are not intended for in depth study of the literature of these communities.

For example, oral Mahosatha stories (Jataka stories based on the Jataka of wisdom) are listed as stories, whereas the Mahosatha manuscript itself is listed as a Buddhist text. The Mahosatha manuscript is the Tai translation of the Pali text, Mahāummaggajātaka (Fausböll 1896), and is an accepted part of the Buddhist canon., However the stories based on it are sometimes stylistically related to the children's stories, and are therefore listed as stories.

The literature of the Tai is discussed in more detail below in Chapter 9.

6. Phonology

6.1 Some theoretical considerations

6.1.1 Proto Southwestern Tai phonemes

6.1.i.1 Proto Southwestern Tai consonants

Proto Southwestern Tai phonemes have been reconstructed by both Jonsson (1991) and Brown (1985). Both of them reconstruct a large series of initials including both single consonants and consonant clusters. Brown (1985:246) reconstructs the series listed in Table (1), presented using his notation:

h	ph	th	ch	kh	khw		hm	hn	hñ	hŋ
	phl			khl						
			s	x			hw	hr		
	f							hl		
ʔ	p	t	c	k	kw		ʔm	ʔn	ʔñ	
	pr	tr		kr						
	pl			kl			ʔml			
	b	d	j	g	gw		m	n	ñ	ŋ
	br	dr		gr						
	bl			gl			ml			
			z	ɣ			w	r		
	v							l		

h	refers to aspirate consonants
?	refers to pre-glottalised consonsants
no mark	refers to voiced consonants

Table 1: Proto Tai Initial Consonants, from Brown (1985)

Table (2) lists the initial consonants as reconstructed by Jonsson (1991:53), using her notation:

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiceless unaspirated stops	p	t	c	k	
Voiceless aspirated stops	ph	th		kh	
Voiced stops	b	d	j	g	
Preglottalised stops	ʔb	ʔd			ʔ
Nasals	m	n	ñ	ŋ	
Voiceless Nasals	hm	hn	hñ	hŋ	
Voiced fricative	v	z		ɣ	
Voiceless fricative	f	s		x	h
Semi vowel	w		j		
Preglottal semi-vowel			ʔj		
Voiceless Semi vowel	hw				
Rhotic Approximant		r			
Voiceless Rhotic Approximant		hr			
Lateral Approximant		l			
Voiceless Lateral Approximant		hl			

Table 2: Proto Southwestern Tai consonant phonemes, after Jonsson (1991)

Jonsson also adds the following clusters:

	Labial	Velar
-l	pl, ʔbl	kl
-l ~ -r	phr, ʔbr, mr~ml	khr~khl, gr~ɣr, xr
-w		kw, khw~xw, gw~ɣw

Table 3: Proto Southwestern Tai initial clusters, after Jonsson (1991)

Both of these reconstructions posit a phoneme inventory which is considerably larger than the present inventory of any the daughter languages from which the reconstruction has been done. Jonsson's inventory posits 37 proto initial consonants in Table (2) but none of the daughter languages has more than 22. None of the daughter languages, for example, includes any voiceless sonorants (here notated as *hm, *hn, *hŋ, *hl, *hr, *hw) in their consonant inventories.

In Jonsson's reconstruction, the glyph <j> is used, following the International Phonetic Alphabet, to refer to the palatal semi vowel. However, in this thesis the symbol <y> will be used instead, because it was the symbol employed for this sound by most researchers into these languages in the past, especially Panchob (1987). Furthermore, the use of <j> might imply to some readers, particularly the Tai speakers themselves, that the sound was actually [dʒ]. In Aiton, the reflex of the proto palatal semi vowel is usually realised as [dʒ] or [ʒ] (see below 6.3.1.1.6).

6.1.1.2 Consonant change in Southwestern Tai.

Previous work on the reconstruction of the Southwestern Tai languages, such as Gedney (1972), Li (1977), Brown (1985) and Jonsson (1991) posit that the voiced consonants as a distinct group have disappeared, as have the voiceless nasals. In the modern Tai languages the voiced and voiceless nasals have merged and are usually realised as voiced nasals.

The voiced stops, on the other hand, have either merged with the voiced aspirate stops (as in Standard Thai and Lao) or to voiceless aspirate stops (as in Shan and in the Tai languages of Assam). For this reason, Standard Thai and Lao are sometimes called 'ph' languages (for example Enfield 2000:50) and Shan and the Tai languages of India are called 'p' languages. This process is bound up with tonogenesis, which is discussed below in 6.1.5.3.

6.1.1 3 Proto Southwestern Tai vowels

The traditional reconstruction of proto Southwestern Tai vowels is that of Li who reconstructs nine vowels, as presented in Table (4)

i	ï	u
e	ə	o
ɛ		ɔ
	a	

Table 4: Proto Southwestern Tai vowel phonemes, after Li (1977:297)

Li posited that vocalic length was not phonemic in his Proto Tai and developed secondarily. Jonsson, on the other hand, claims that only seven primary vowels "are well-retained in all the daughter languages" (1991:104). These are presented in Table (5), using her notation:

i	ï	u
æ		ɔ
	a, aa	

Table 5: Proto Southwestern Tai vowel phonemes, after Jonsson (1991)

Jonsson explicitly states that "The mid vowels // e ə o // are not reconstructed. The evidence shows that they are a result of raising, lowering, shortening and reducing of other primary and complex vowels." (1991:104). It is worthwhile to note that this reconstruction is exactly the vowel inventory proposed in this thesis for Aiton (see 6.3.2) and for Ahom (see 6.5.2).

Jonsson also reconstructs three diphthongs, /ia/, /ïa/ and /ua/, although she puts the question as to whether there were complex vowels in proto Southwestern Tai or whether they arose as a result of diffusion (1991:122).

Since the various authors discussed here use different notations for the same vowels, a comparison of these various different notational systems is presented in Table (6)

Li (1977)	Banchob (1987)	Jonsson (1991)	Diller (1992)	This thesis
i	i	i	i	i
ï	ü	ï	ɯ	ɯ
u	u	u	u	u
e	e	e	e	e
ə	ə	ə	ɤ	ɤ
o	o	o	o	o
ɛ	ɛ	æ	ɛ	ɛ
ɔ	ɔ	ɔ	ɔ	ɔ
a	a	a	a	a

Table 6: Comparison of vowel notations - Li (1977), Banchob (1987),
Jonsson (1991), Diller (1992) and this thesis.

6.1.2 Diphthongs in Tai languages

As mentioned above in 6.1.1.3, three diphthongs (/ia/, /ɯa/ and /ua/) are reconstructed for proto Tai. In addition to these three diphthongs, there are a number of combinations which are usually analysed as being vowel + semi vowel combinations, in particular [aj] and [aw], and in the Shan languages [aɯ].

Edmondson and Solnit (1997a:15) discuss the issue of "syllable-final semivowels", pointing out that Gedney and others in what they term the American tradition

"treat them as consonants and write them with consonantal symbols such as [-y, -w, -ɰ]. They point out that in Tai there are no possible syllable shapes aiC/auC/aɯC (where C stands for an arbitrary consonant). By assuming that -Vi/-Vu/-Vɯ are in fact phonologically -Vy/-Vw/-Vɰ, then no further apparatus is necessary; double consonant codas are simply forbidden as syllable types in Tai."

Edmondson and Solnit also point out that in many Tai languages, vowel length is only contrastive in syllables that have final consonants. This vowel length contrast also applies to syllables with final semivowels.

Li (1977) and other Chinese scholars, on the other hand, notate these "syllable-final semivowels" with the vowels, (-i, -u and -ɯ). In this thesis, the approach of Li will be followed, rather than that of Gedney and the American tradition. There are several reasons why this is being done.

Firstly, phonetically, it is clear that in many cases the finals are vowels rather than semi vowels. Harris (1976:118), writing of Khamti, says of /y/ that "In syllable final position this symbol represents a close front unrounded vowel [i]". In the Tai languages of Assam, final /y/ and /w/ are often long after /a/, particularly if /a/ is short. This is clearly demonstrated in Table (7), which presents some Tai Aiton examples spoken by Bidya Thoumoung:

Assamese		Tai		gloss	length of /a/	length of /i/
কুকুৰা	kukura	ন3	kai[1]	chicken	0.1"	0.3"
শজুৰা	khɔdʒuwa	ন3	kaai[2]	irritated	0.32"	0.1"

Audio link to Pronunciation of these words (\wave files\aiton\sounds\kai.mp3)
Link to spectrogram analysis of kai[1] 'chicken' (\wave files\aiton\wave images\kai1.bmp)
Link to spectrogram analysis of kaai[2] 'irritated' (\wave files\aiton\wave images\kaai2.bmp)

<center>Table 7: Length of syllable final vowels/semivowels in Aiton.</center>

A second reason for the decision to use final vowel notation is that in Phake at least, the notation of these as <ay> (<aj>) and <aw> would create another difficulty. As discussed below in 6.2.2.2, synchronically Phake allows the diphthongisation of mid-open vowels in front of velars. This means that the vowel in ο𝔔ε laiŋ[2] 'light' (from proto *lɛŋA4)[40] is the same as that in ο3 lai[6] 'to flow' (from proto *laiA1). This can be heard in the words phā[4] laiŋ[2] yɛ[4] 'the light', as spoken by Ee Nyan Khet in the *Story of Deception* 1-1-3-6 No. (33). The wave form for this utterance also shows the transition from the low to the high front vowel[41].

If we notate the [ai] group as <ay>, then we have to admit the possibility for the final consonant to be a cluster, at least phonetically.

A third reason for writing these combinations as two vowels is that historically the Tai languages have written /ai/ as a single graph, which would suggest that the intuition of this is as a diphthong, rather than a vowel-semi vowel combination.

Finally, treating the finals in /ai/ and /au/ as semi vowels will necessitate treating the final in /au/ as a semi-vowel as well. This would involve the positing of a consonant phoneme which is restricted to a syllable final position, and which has exactly the same phonetic form as a vowel which has normal vowel distribution.

---

[40] The proto tone is symbolised by A4, which refers the tone box. See below 6.1.5.3.
[41] The sound file for this example is in \wave files\phake\sounds\laing2.mp3. The wave form can be viewed at \wave files\phake\wave images\laing2.bmp.

6.1.3 Syllable structure in Tai languages

Noss (1964:5) categorises the syllable structure of Standard Thai as follows:

"Within the syllable, there are four types of components: initial, nucleus, final, and tone. Consonant phonemes result from the analysis of initial and final components, vowel phonemes from nuclear components, and tone phonemes from tonal components. These are all syllabic phonemes."

Table (8) presents an analysis of the Tai syllable, where what Noss calls initial is described as onset (O), and the other three components, here called nucleus (N), coda (C) and tone (T), are part of the rime (R). The Phake word ο𝔔ν kin[2] "eat" will be exemplified
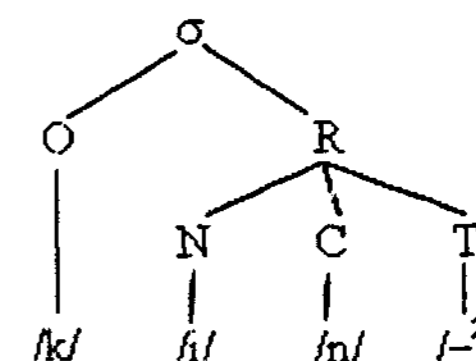
<center>Table 8: Syllable structure in Tai, as exemplified by the Phake word ο𝔔ν kin[2] "eat"</center>

Within this structure, the onset and rime are obligatory – with vowel initial words being regarded as having an initial glottal stop as the onset – and within the rime, the coda is optional. It is generally accepted by Tai scholars that a nucleus can be a diphthong, but only from the following restricted set of diphthongs /ia/, /ɯa/ and /ua/.

As discussed above in 6.1.2 there are problems with this model for Tai Phake, at least synchronically. In a word like Phake ο𝔔ε laiŋ[2] 'light' we either have to posit that /ai/ is a diphthong (the position adopted in this thesis) or that the word is /layŋ/ and two final consonants are allowed in the coda. It is beyond the scope of this thesis to explore in detail the theory of syllable structure in Tai.

6.1.4 The glottal stop

Most researchers into Tai languages mark a glottal stop in front of initial vowels. At least some slight consonantal onset is always necessary before vowels can be released in any language, although in many languages glottal initials are not written.

In some languages, glottal stop initial is contrastive with the lack of same, and minimal pairs, or near minimal pairs can be found. Table (9) illustrates such a near minimal pair for Maltese:

Written Maltese	Pronunciation	Gloss
arja	arja	"air"
qara	ʔara	"he read"

Table 9: Glottal contrast in modern Maltese

In Tai languages, the writing of glottal stop occurs for several reasons. Firstly, in Tai orthography – at least for the Indic based orthographies – vowels cannot stand alone (see Chapter 7 below), and so a dummy consonant (often called simply *a*) is written. This can then be marked in transcription using the glottal stop.

Secondly, the glottal stop is usually regarded as necessary to explain a tonal phenomenon in the Tai languages - the fact that vowel initial words can occur only with a restricted number of tones.

For example, in Phake, vowel initial words are mostly confined to tones 1, 2 and 3. (see 6.2.4 below). The lack of the rising tone for vowel initial words in Phake is explained by positing the glottal stop *ʔ/ as an initial phoneme in the proto form of such words. Since the tone of a Tai word is dependent on what its initial syllable was at the time of the tonal splits (see 6.1.5, Table 10), it is argued that glottal stop is still a phoneme, causing words with glottal initial to be able to take only certain tones, the same tones as the reflexes of the words which began with proto */b/ and */d/. Together these three proto phonemes make up the proto glottal series of initials in Gedney's (1972) tone chart (see 6.1.5.3).

A further argument for marking glottal stop is that it conforms to the theory of Tai syllable structure, namely that a syllable consists of an onset (the initial consonant or consonant cluster), a nucleus (the vowel), a tone and an optional coda (final consonant). The glottal stop is the onset in a word which has a vowel initial.

Finally, glottal stop is often present phonetically where vowel hiatus would otherwise occur, as in Phake ꩬꩧꧡꩫꩬ [má ʔúik] /ma¹ ük¹/ "tomato".

One example in which the vowel hiatus is overcome by a glottal stop is presented in (1), from Example (53) in *Story - Chaw Mahosatha meets Nang Pingya*, spoken by Ee Nyan Khet. Here there is a glottal stop, between the [ü²] of *kauŋ² mü²* 'pagoda' and the following word[42]. This would be expressed phonetically as [kauŋ² mü² ʔan²].

1)   ... ꩬꩫꧡꩬ   ꩫꩬꩫ   ꩫꩫꧡ   ꩬꩫꧡ ...
     ... kauŋ² mü² an²   nüŋ⁵   yaŋ² ...
     ... pagoda   CLF   one   have ...

All of the above reasons have lead many scholars to mark glottal stop in their transcriptions. However it is not universal. Some of the scholars that do not mark glottal stop include Nos. (1964), Weidert (1977) and Banchob (1987). Along with these scholars I will not be marking glottal stop in phonemic transcriptions.

In support of this decision, there are cases of vowel hiatus where not only was there no insertion of the glottal stop, but the two vowels are run together. The best example of this is (2), where the words *kaa²* 'crow' and *an²* 'general classifier' are spoken run together so that it is phonetically [kaan²].

2)   ꩬꩫꩬꩫꩫꩬ   ꩫꩬ   ꩫꩬꩫ   ꩫꩬ   ꩫꩬꩫ   ꩫꩬ   ꩫꩬꩫꧡ
     mü² ʒaam² nan³ kaa²   an²   uu¹   luk³   tii²   tun³ mai³
     at that time   crow CLF   live from   at   tree
     'At that time the crow was living in a tree.'
     (Aiton Story, *Story of the Crow and the Fox*, No. (5), told by Ong Cham).

---

[42] The audio link for example (1) is at \wave files\phake\sounds\hiatus1.mp3.

Had a glottal stop been present at the onset of *an²*, we would have expected this merge would not result. Although the recording quality is poor, this phenomenon is still audible.

The argument about the need to express a glottal stop to explain the tonal distribution of words with vowel initials, can easily be rephrased as vowel initial words behave like those which had proto-glottal initials and are basically confined to tones permitted for A3, B3, C3 and D3 (see below 6.1.5.3).

In Aiton, at least synchronically, it is no longer necessary to posit the glottal stop even for this reason. Because of the reduction to three tones, all three tones are available in Aiton for all initial consonants, as well as for all initial vowels. It is extremely unlikely that any scholar describing Aiton without the benefit of the various theoretical models discussed above would ever think of marking glottal stop. No minimal pair can be found to distinguish a glottal initial from a vowel initial, and it can just as easily be said that *an²* "CLF" and *pan²* "to give" are distinguished by the absence of a segment (as *Ann* and *pan* are in English) than that they are distinguished by the presence of an initial glottal in one, and an initial bilabial in the other.

6.1.5 Tones

6.1.5.1 Definition of Tone

In their discussion about Cantonese, Bauer and Benedict categorised tone as follows:

> "By systematically manipulating the pitch of the voice, speakers of a tone language have at their disposal a second phonetic dimension – in addition to the consonants and vowels – for contrasting words. By convention, such distinctive and indispensable pitch differences are called *tones* (or more precisely, *lexical tones*)." (1997:107)

The importance of tone to languages such as Tai (and Cantonese) cannot be overemphasised. Tone is a crucial and obligatory feature of every Tai word. As Bauer and Benedict explained it:

> "..the High level tone of *tãm* ... 'to carry on a shoulder pole' is an essential component to the complete form of the word as are the initial unaspirated alveolar stop consonant *t-*, the long nuclear vowel *-a:-*, and the final bilabial nasal consonant *-m*. A speaker's arbitrary change of the High Level tone ... to some other tone ... would prove quite puzzling to his/her addressee...; in a comparable situation an English speaker might pronounce a word with a different vowel, eg. *bed* [be:d] as *bead* [bi:d]. However in the context of a sentence, eg. *What time are you going to* [bi:d] *tonight* ..., addressees might be able to figure out what words the speakers had intended to say." (1997:107)

Tones take many forms in the languages of the world. In the Tai languages most of the tones are contour tones. Contour is the relationship between pitch height and pitch movement in time. A high level tone has a high pitch height, and little or no pitch movement. A high falling tone may have high pitch height and then lowering movement over time. It must be noted, however, that none of the pitches will be absolute, because of the different levels of voices, the most obvious of which is the different pitch levels of the voices of men and women.

There are other types of tones in Tai, which are sometimes categorised as register tones. In Tai the presence or absence of glottalisation or creakiness is a register characteristic. In both Phake and Khamyang, glottalisation is a distinctive feature between tones. In Khamyang, for example, the 2nd and 4th tones are distinguished only by the glottalisation of the 4th tone and its absence in the 2nd tone. The contours of these two tones are identical (see below 6.4.3).

6.1.5.2 Notation of tone

Several different systems have been devised for the notation of tone. These can be categorised as follows:

i) The use of pitch numbers, where 1 is low, 2 mid-low, 3 mid, 4 mid-high and 5 high. In such a system a high level tone on *pi* would be notated *pi55* or *pi⁵⁵*. This system is known as alphanumeric.

ii) The use of symbols, in which a high tone on *pi* would be written as *pi⌐*. In this system, the vertical line is the axis and the horizontal line categorises the shape of the tone.

iii) The use of IPA diacritics, where a high tone on the syllable *pi* might be written as *pí*.

iv) The use of arbitrary numbers to categorise the tone, but which refer to a particular type of tone. In Tai Phake, for example, the tone which is numbered (1) is categorised as being mid to high level with a possible final fall (see 6.2.4.4 below). By marking such a tone on the syllable *pi* as *pi¹*, no particular part of the tonal description is present in the notation, but all words with a similar tone are easily grouped together.

An arbitrary system is also used by the only Tai language to have developed an orthography based on the Roman alphabet, namely Zhuang, which uses syllable-final letters to mark the tones. For example the letter *z* marks the tone in *naz* /na³¹/ 'ricefield' (Edmondson and Solnit 1997a:16).

6.1.5.3 Tone in Tai Languages - Theoretical & historical background

Gedney (1972) recognised that the native vocabulary of the languages and dialects of the Tai family falls into 20 groups of words. At an earlier stage of the Tai language (described by Gedney as "at the time of tonal splits"), there were three tones on what Gedney calls 'smooth syllables', that is live or open syllables, and two more on what he calls 'checked syllables', that is dead or closed syllables[43]. The three live tones have

---

[43] The terms 'live' and 'dead' syllable are used by some scholars in the Tai field. Since words with final stops (called 'dead' or 'checked') have different tonal outcomes from words with final vowels or nasals (called 'live' or 'smooth'), it is this distinction that is crucial in the study of Tai tones.

become known as A, B and C; whilst the dead syllables, that is final /-k/, /-t/, /-p/ are known as D-long and D-short, depending on the length of the vowel, as in Table (10):

| | | Proto-Tai Tones | | | | |
		A	B	C	D-short	D-long
Initials, at the time of tonal splits	1 Voiceless friction sounds, *s, hm, ph, etc.	1	5	9	13	17
	2 Voiceless unaspirated, *p, etc.	2	6	10	14	18
	3 Glottal, *ʔ ʔb, etc.	3	7	11	15	19
	4 Voiced, *b, m, l, z, etc.	4	8	12	16	20
		Smooth Syllables			Checked Syllables	

Table 10: The Tai Tone Boxes (after Gedney 1972)

In this system, the tone of a word depends on one of five proto tones (A, B, C, D-short and D-long) in combination with one of four types of proto initials, namely voiceless aspirated, voiceless unaspirated, glottal and voiced. Following the orthography developed for Standard Thai, the voiceless aspirated is sometimes known as the "high series", the voiceless unaspirated and glottal the "mid series" and the voiced the "low series".

In many Tai varieties, particularly those spoken in China, only a bipartite division between the originally voiceless (the first 3 lines in Gedney) and originally voiced (the 4th line) is necessary. In some publications on Tai, therefore, the term A2 will refer to words which are reflexes of originally voiced initials with the A tone. In this thesis, following Gedney, these will be referred to as A4.

The original A, B and C tones may have been tones with registral qualities such as breathiness or creakiness, rather than of pitch and contour as they are today. Since words with initial voiced stops tend to have lower pitch than those with initial voiceless stops, it is theorised that this process, combined with a change of the originally voiced series (numbered 4 in Table 10) into voiceless stops, lead to the tonogenesis of the contour and pitch tones which are found in Tai languages today.

A change in the initial consonant from voiced to voiceless would mean that some words in line 4 would become homophonous with those in line 1. For example, the word which was *bii (A4) 'fat' would become *phii (A4) and thus identical to *phii (A1) 'ghost'. A tone split would then become necessary to maintain the contrast between these words.

Gedney gives example words for each of his 20 boxes. These words, with some additions, can be used to discover and test the tonal system of any Tai language or dialect. Generally Tai dialects have between 4 and 7 separate "tonemes".

I have found that Gedney's tone chart works extremely well with Phake informants, but less so with both Aiton and Khamyang informants. In part this appears to be because both Aiton and Khamyang are languages undergoing or coming to the end of a process of change - Khamyang because it is a language in its last phase and Aiton because it may have undergone toneme merge in recent times.

For Aiton and Khamyang, therefore, it was necessary to find minimal pairs in order to be able to work out the tonal system. The syllable *maa* was examined first, because in most varieties it is possible to find most of the tones with this syllable. The possible tones on other syllables such as *kaa* and *taa* were then explored.

When a list of words had been collected and the contours of the contrasting tones made clear, a Gedney style tone chart could be produced for each variety.

6.1.5.4 Changed Tone

Bauer and Benedict (1997:175), describe a process in Cantonese, known as *pinjam* 'changed tone'. There are a number of different processes subsumed under this title, but in one of them "in certain lexical contexts the tone of a syllable assimilates to the high endpoint of the following syllable to become High Rising and then the second syllable is lost."

One example of the process is the loss of the Cantonese perfective bound suffix *-tsɔ́:*, which can be lost as in 3b)

3a) *khéy tso'w -tsɔ́: la:*
3b) *khéy tso'w* la:*
'He has done it'

The word *tso'w* 'done' has a mid low level tone in citation[44], whereas *tsɔ́:* has high rising tone. In example (3b), the *pinjam* (changed tone) is marked with an asterisk. When example (3b) was subject to analysis, Bauer and Benedict concluded that the dip and peak values of the high rising *pinjam* contour on *tsow** 'done' were very similar to those for the regular High Rising tone contour on *-tsɔ́:* in sentence 3a). In effect, the tone of the perfective suffix *tsɔ́:* has been transferred onto the verb, and *tsɔ́:* is then dropped.

As will be discussed below in 6.2.4.5, some analogous processes have been observed in Phake. As far as I know, no other Tai languages of the Southwestern group within the Tai family exhibit such characteristics.

6.2 Phake

The following discussion of the phonology of the Phake language is largely based on the *Phake-English-Thai Dictionary* (Banchob 1987), together with some observations from the author's own experience. Banchob's notations have been retained for all consonants and tones, both in the *Tai Phake Dictionary* (see below 10.2), and in all transcriptions and in all analyses. Some minor alterations to her transcriptions of vowels are discussed below in 6.2.2.

6.2.1 Consonants

Table (11) lists the consonants observed by Banchob (1987:12). The symbols used for the consonants are those chosen by Banchob. This decision was made because of the

---

[44] This is marked in Bauer and Benedict with a vertical stroke above the vowel. Here, owing to the lack of a standard font which can do this, it has been marked with an apostrophe following the vowel.

availability of the *Phake–English–Thai Dictionary* in India, and the feeling that it would only add to confusion to alter them.

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiceless unaspirated stops	p	t	c	k	(ʔ)
Voiceless aspirated stops	ph	th		kh	
Nasals	m	n	ñ	ŋ	
Voiceless fricative		s			h
Semi vowel	w		y		
Rhotic Approximant	r				
Lateral Approximant	l				

Table 11: Consonant Phonemes in Phake

Since the status of the consonant phonemes in Phake is uncontroversial, and since they have been discussed by Banchob (1987), no fuller description of the consonants will be presented here. Readers who wish to see minimal pairs may wish to consult the *Tai Phake Dictionary*[45] where all the words recorded by Banchob can be found.

There is one allophonic process of Phake consonants which needs to be pointed out. Certain very common function words with initial /k/ show initial lenition when they are unstressed. This is particularly common with reciprocal marker *kan²* and with *kā¹* 'go', when it is not a main verb (see below 8.5.7.5.1). This is demonstrated in example (4):

4) ꩬ   ꩬ   ꩬꩬ   ꩬ   ꩬ   ꩬ   ꩬ   ꩬ   ꩬ
    nɛ⁴   cau³   ma¹ hɔ¹   nɛ⁴   hū⁴   kā¹   hū⁴   thom¹ kā¹
    DEF   RESP   Mahosatha   PRT   know   GO   know   listen   GO

    ꩬꩬ   ꩬꩬ   ꩬꩬ   ꩬ   ꩬ   ꩬ   ꩬ ॥
    ɔ¹   ū³   mün²   nɛ⁴   kā¹   thai⁶   nā²
    EXCL   father   3Sg   DEF   go   plough field
    'Chaw Mahosatha knew that her father had gone to plough the fields.'
    (Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (33), told by Ee Nyan Khet)

[45] See \lexicon\phake\phakdic.doc in the electronic version of this thesis.

As a full verb in the second line of (4), *kā¹* is realised as [kā¹]. In the first line, however, as a function morpheme it is first realised as [ɣā¹] after the vowel /u/ and then with nasalisation after /m/. These processes appear only to apply to common function morphemes with initial /k/.

6.2.2 Vowels

Based on Banchob (1987) and observations by the writer, it is believed that there are 10 contrastive vowel phonemes in Phake, as in Table (12): (The symbols are following Banchob).

i	ü	u
e	ə	o
ɛ		ɔ
	a, aa	

Table 12: Vowel phonemes in Phake

Table (13) lists minimal pairs distinguishing respectively the two back unrounded vowels, the three front vowels and the three back vowels.

Assamese		Gloss	Tai Phake
হাত	hat	'hand'	mü²
যোৱা	dʒuwa	'return home'	mə²
কিমান	kiman	'how many'	kī¹
অলপ	ɔlɔp	'a little'	ke¹
বুঢ়া	bura	'old'	kɛ¹
সাপ	hap	'snake'	ŋū²
গৰু	gɔru	'cow'	ŋo²
–	–	'crooked'	ŋɔ²

Audio Link to the Pronunciation of Phake Vowels (\wave files\phake\sounds\vowel2.mp3)

Table 13: Comparison of the Phake vowels.

6.2.2.1 Vowel length in Phake

Banchob (1987:12) actually lists 17 vowels, as in Table (14):

i,ī	ü,ǖ	u,ū
e,ē	ə,ə̄	o
ε,ε̄		ɔ,ɔ̄
	a,ā	

Table 14: Vowels listed by Banchob (1987)

Banchob does not give any examples of contrasts between short and long vowels in her phonological discussion. In the bulk of the text it is clear that there are many examples of length contrasts for the vowels /a/~/aa/, for example *kaŋ³* "to screen" and *kāŋ³* "fishbone".

In general, for vowels other than /a/, Banchob notates the vowels of open syllables as long and those in closed syllables as short. In part this follows the practice of the Phake script, where, for example, /u/ in an open syllable is marked with a symbol which is historically a long vowel <-ᷓ> and in a closed syllable with a symbol which is historically a short vowel <-ᷕ>. There is clearly a length difference between the open and closed syllables, and for that reason, I have followed Banchob in marking at least syllables which end in final /i/ and final /u/ with long vowels. This is not intended to suggest that there is a phonemic distinction.

Banchob (1987) generally only marks vowels in open syllables as being long (except for /a/ where length distinction clearly still applies). There are some exceptions, however. The vowel /e/ is sometimes marked by Banchob as long in closed syllables[46], as in (5.1), and sometimes not, as in (5.2)

5.1) ᷽ᷭᷟ   cēŋ¹   a corner, an angle

᷽ᷭᷟ   cēŋ²   the first month of the lunar year

᷽ᷭᷟ   cēŋ⁵   to learn how to stand, as the infants

᷽ᷭᷟ   cēŋ⁵   to take morning ... h, esp. for animals

---

[46] Note that 'closed' syllable here is used ... al consonant. This is not the same as 'dead' or 'checked' syllable as used above i... syllables with final stops. The term 'closed' syllable includes both final stops an...

5.2) ᷽ᷭᷟ   keŋ²   a kind of grass, softer than the others

᷽ᷭᷟ   keŋ²   an iron tripod to support cooking pots or pans over a fire

᷽ᷭᷟ   keŋ²   a period, time

᷽ᷭᷟ   keŋ³   decent or nice to look at (mostly the fruits)

The investigation for the present thesis has also revealed that at least in closed syllables, the mid vowels tend to be somewhat longer than the high vowels, as demonstrated for the back vowels in Table (15):

English	Tai Phake		Approximate length of vowel
dig	᷽ᷭᷟ	khut¹	0.13"
curvy	᷽ᷭᷟ	khot¹	0.2"
to tie	᷽ᷭᷟ	khɔt¹	0.2"

Audio Link to the pronunciation of these words (\wave files\phake\sounds\khut.mp3).

Table 15: Vowels listed by Banchob (1987)

There are however no minimal pairs to indicating that this length is ever contrastive. I have therefore adopted the practise of marking all non /a/ vowels as being short in closed syllables, even when Banchob marks them as long in her Dictionary.

6.2.2.2 Diphthongisation

Banchob reports a regular process of diphthongisation of the mid-low vowels before velar consonants.

6) /ε/ → [ai] / __ velar
    /ɔ/ → [au] / __ velar

This process has already been discussed and exemplified above in 6.1.2. It is frequently observed in Phake villages today, but is not observed in other Tai varieties in which neither the diphthong /ai/ nor /au/ can be followed by a final consonant.

It does appear that a mid-low simple vowel allophone is still present in the speech of some Phake informants. In example (7), the word /laiŋ⁶/ is clearly pronounced [lɛŋ⁶], as can be seen from the spectrogram analysis of the word *laiŋ⁶*'vehicle' (see \wave files\phake\wave images\laing6.bmp).

7)   ᩋ   ᨲᩬᨻ ᩋ   ᨴᩱᨦ ᨾ   ᨧᩮᩬᩢ

    khī¹   lɔ³   khī¹   laiŋ⁶ ma¹ caü⁶

    ride   cart ride   vehicle NEG be

   ᩃᩩᨦᨻᩣ   ᨲᩣᨦ ᩉᩬᩢ ᨿᩮ ᨠᩣ ᨠᩣ ॥

    luŋ⁴ phā¹ tāŋ² hɔ² yɛ⁴ kā¹ kā⁴

    because of this way boat PRT go trade

    'They did not ride their vehicles, and because of this they went to trade by boat.'

    (Phake Story, *The Dolphin, the Crow and the Mosquito* 1-1-1-4, No (21)), told by Aithown Che Chakap)

It is therefore possible that it would be better to notate the diphthongised /ai/ and /au/ in words such as *laiŋ⁶*'vehicle' as simply mid-open vowels. However the diphthongisation is very prevalent in the Phake speech and the advantages of maintaining Banchob's notation are felt to outweigh the disadvantages.

Other dipthongisations may be occurring among Phake speakers. Diller (1992:17) notes that "the vowel -ɔ- in other non-final environments shows the same [-uɔ-] or [-wɔ-] variant noted by Harris for Khamti (1976:120); there is most frequently a contiguous labial *p* or *m*: *kɔp⁶*[kwɔp] 'scoop'; *pɔt¹*[pwɔt] 'to scrub'..."

In February 2002, Yehom Buragohain was recorded speaking the second of Diller's examples, and her pronunciation of it can be heard at \wave files\phake\sounds\pot.mp3. The word (*pɔt⁴*in Banchob's system), seems to clearly show dipthongisation, with a shift to the vowel shape at the end of the vowel, not at the beginning, as can be seen in the spectrogram at \wave files\phake\wave images\pot4.bmp. This word might better be phonetically expressed a [pɔət].

6.2.2.3 Vowel reduction

Although Banchob (1987) and some other writers on the Tai languages write the symbol <ə>, this sound is not a schwa. Rather it is a back unrounded mid-close vowel, which is marked as <ɤ> in this thesis, except in Phake texts which follow Banchob's orthography. The fact that it is not a schwa can be clearly heard in Yehom Buragohain's pronunciation of the word *mɤŋ²*'country' (দেশ dɛh in Assamese) (for the sound file, go to \wave files\phake\sounds\meung.mp3).

In Phake, there are several common words which were originally spoken on other vowels but which are now pronounced with vowels very similar to the back unrounded vowels. These include ᨾᩢ '3Sg', originally pronounced man², but written by Banchob as mün². In modern Phake orthography, the word is often written ᨾᩢ. This is probably a vowel reduction process, which has been reinterpreted as a change to the vowel /ü/ or /ɯ/. The reason why Banchob chose to notate it as <ü> rather than <ə> is probably that in closed syllables /ə/ is usually realised as a long vowel, as in the example of *mɤŋ²*in the previous paragraph.

Several other very common words have undergone vowel reduction in Phake speech. These are listed in Table (16).

Spelling in Manuscripts	Modern Phake Spelling	Putative original pronunciation	in Banchob (1987)	usual pronunciation	gloss
ᨠᩣ	ᨠᩮᩬᩢ	kɔ⁴	kɔ⁴, kə⁴	[kə⁴]	'LINK'
ᨧᩬᨦ	ᨧᩮᩬᨦ	saŋ⁶	süŋ⁶	[səŋ⁶]	'if'
ᨶᨦ	ᨶᩮᩢᨦ	naŋ¹	nüŋ¹	[nəŋ¹]	'like'
ᨲᨦ	ᨲᩢᨦ	haŋ²	haŋ²	[həŋ²]	'to'

Table 16: Vowel reduction in Phake

In connected speech, these words are usually unstressed. The main difference between the vowels of *kə⁴*and *süŋ⁶*, as they are realised in examples (8) and (9), is one of

length, rather than one of height, with the vowel of the open syllable being realised as a longer vowel.

8) ᩃᩢᩢ 𑄟𑄝 ᩮᩬᩳ ᩮᩬᩳ ᩅᩣᩴ ᩇᩬ ‖

   nāŋ² piŋ¹ ñā¹ kə⁴ ma¹ tān³ se⁶

   lady Pingya LINK NEG speak PRT

'Miss Pingya did not speak.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, 1-1-1-3, No. (12), told by Ee Nyan Khet)

9) ᩟ᩢ ᩚ ᩅᩢᩢ ᩗ ᩋᩢᩢ ᩅᩣ ᩇ ᩔ ᩈᩢᩢ

   mün² wā⁵ süŋ⁶ phā⁴ phon⁶ ye⁴ kau² mau¹ pauk⁴

   3Sg say if weather rain PRT 1Sg NEG return

   ᩋ ᩈᩢᩢ ᩃ ‖

   ta¹ sɔn⁴ nā²

   will take shelter field

'She said, "If it rains, I won't return, I'll take shelter in the fields."'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, 1-1-1-3, No. (47), told by Ee Nyan Khet)

6.2.3 Phonotactics

6.2.3.1 Initial consonant clusters

Banchob (1987) does not list possible initial consonant clusters in her brief introduction, but the forms listed in Table (17) are all found in her dictionary:

ky	kl	kw
khy	khl	
	tl	
py	pl	
phy		phr	phw
my		

Table 17: Initial consonant clusters found in Banchob (1987)

Kingcom (1992:36) lists the possible clusters in Phake in Table (18)

ky	kl
	khl
	tl
py	pl
phy	phl
my	

Table 18: Initial consonant clusters in Phake (after Kingcom 1992)

Kingcom specifically excludes the possibility of clusters with /w/ as the second member, whereas Banchob records several such words.

In both cases, virtually every example is a borrowed word from Burmese or Pali. Consider the words which Banchob recorded as having the initial /ky/ cluster in Phake are listed in (10):

10)   ᩢᩢ₁₁ kyām¹ to attempt

    ᩢᩢ₂₁ kyām² rough, fierce

    ᩢᩢᩢ₁₁ kyok¹ the sound of a big cock crowing

    ᩢᩢᩢᩢ₁₁ kyok⁴ pɛn³ a grindstone

    ᩢᩢᩢ₁₁ kyauŋ² a Buddhist monastery

Three of these words are also found in the Shan dictionary (Sao Tern Moeng 1995), where each of them is marked as Burmese loan. Of the other two, ᩢᩢᩢ *kyok¹* "the sound of a big cock crowing" is an onomatopoeic word, and ᩢᩢᩢᩢ *kyok⁴ pɛn³* "a grindstone" is probably related to Shan ᩢᩢᩢᩢ, *kyɔk; taan*, "slate pencil", which is also a Burmese loan.

In Phake, those words which are reflexes of words that had initial consonant clusters are rarely realised with such initial clusters, unlike in Aiton (see 6.3.3.1). One which occasionally appears is ᩢᩢ *kwā¹*, "go", the spelling of which reflects the proto Tai initial */kw/ and which is occasionally pronounced as if there was an initial cluster. This suggests that /w/ final clusters may have been the last to be lost in Phake.

Apart from this exception, the clusters which are permitted today occur only in loan words, or onomatopoeic words, or terms such as names of birds or fish. This is quite unlike the situation in Aiton, where some proto-Tai clusters are preserved.

6.2.3.2 Syllable structure

Kingcom (1992:26f) lists and exemplifies all of the possible arrangements of consonants and vowels in syllables in both Tai Phake and Standard Tai. So, for example, he comes up with the following four possible structures for a monosyllable in Tai Phake:

11) CV
     CVC
     CCV
     CCVC

Whilst (11) is essentially a restatement of the possible syllable structure discussed in 6.1.3, Kingcom goes on to state that only certain combinations of these are found in disyllabic words. The permitted structures for disyllabic words, according to Kingcom, are presented in (12):

12) CV.CV
     CVC.CV
     CV.CVC
     CVC.CVC
     CVC.CCV

If all four patterns found in (11) were available in all possible combinations, there would be 12 possible disyllables. One example, recorded by Banchob (1987), which does not accord with Kingcom's findings is *kyok⁴ pɛn³* 'a grindstone', which has the structure CCVC.CVC. Further investigation will be necessary to establish whether all the possible combinations of syllables listed in (11) are permissible in disyllabic words.

6.2.4 Tone

As discussed above in 6.1.5.3, the tonal system of Tai languages can be derived from a relatively modest number of words, by using the principles laid down by Gedney (1972). The following section will present previous examinations of the Phake tonal system and compare those observations with the present investigation.

6.2.4.1 Tone in Phake as notated by Banchob (1987)

Banchob, like the writer, did most of her research on Phake in the village of ဖၟၵၞၟၚ *mān³ phā⁴ kɛ⁵ taü³* (Lower Phake village) or Namphakey, near Naharkatia, Dibrugarh District.

Based on Banchob (1987), the following tone chart (Table 19) can be derived:

	A	B	C	DL	DS
1	6	1	3	1	
2	2				
3					
4		5	4	4	

1	even tone with a slight rising at the end	mā¹	"shoulder"
2	high tone	mā²	"to come"
3	grave tone, with a glottal stop	mā³	"mad"
4	falling tone	mit⁴	"seeds, knife"
5	grave tone	mā⁵	"not",
6	high rising tone	mā⁶	"dog"

Table 19: Tones in Phake (after Banchob 1987)

6.2.4.2 Tone in Phake reported by Diller (1992)

Banchob's system differs slightly from that reported by Diller (1992:18). Diller's informant, Yehom Buragohain, was from the same village that Banchob worked in, but had lived away from the village for many years away. The following tone box (Table 20) retains Banchob's numbering for ease of comparison, but gives Diller's

descriptions of the tones. The distribution of the tones is the same, but the realisation of them is slightly different:

	A	B	C	DL	DS
1	6	1	3	1	
2	2				
3					
4		5	4	4	

1	mid rising
2	high falling
3	mid falling glottalised
4	mid falling
5	low
6	low rising tone

Table 20: Tones in Phake (after Diller 1992)

6.2.4.3 Tone in Phake as reported by Wilaiwan (1983)

A quite different tonal system is reported by Wilaiwan (1983). There are around one hundred Phake words exemplified in her article and all but two or three are marked for tone. By examining all of these, it has been possible to derive the following tone chart for her data:

	A	B	C	DL	DS
1	1	4 (?)	3	4	
2	2				
3					
4	5	6	5 or 3?	6	

	Description of tones (Wilaiwan 1983:231)
1	low level then rising
2	mid level
3	mid falling glottalised
4	high
5	high level then falling
6	mid

Table 21: Phake tone box, derived from data in Wilaiwan 1983

Unfortunately, the data from Wilaiwan is not comprehensive enough for us to be fully confident that Table (21) is correct. There are enough tokens of words with the A tone to be clear that for this variety A1 is the low level rising tone (Tone 1 in Wilaiwan's system), A2&3 the mid level tone (Tone 2), and A4 the high tone (Tone 5). There are also a reasonable number of tokens for B4, suggesting that this is the mid tone (Tone 6).

There are fewer tokens of the C tones, but all of those for C1-3 are the mid falling glottalised tone (Tone 3). There are quite a few tokens for C4, but some are marked tone 3 and some tone 5 - perhaps suggesting that this is a high falling glottalised tone, such as is found in Khamyang for C4 (see below 6.4.3). It is equally possible that C4 is merged with either C1-3 or with A4.

Finally, there are very few tokens for B1-3. Some of them are marked tone 4 by Wilaiwan, as *luu⁴* "to offer" and *ʔuu⁴* "to be at", and it was this that lead me to suggest tone 4 for that box. However a number of words which are B1-3 are not marked for tone 4, as *yai⁶* "big" and *maa⁵* "shoulder". I do not know of any Tai language in which the tone of 'shoulder' and 'to be at' differ from each other, nor have I met any Tai Phake speakers who use a variety in which the tones are distributed as in Table (21).

Certainly the tonal system of Table (21) is quite different from that of the Phakes in Namphakey village whom I have interviewed. It does however some show similarity to Khamti as described by Harris (see above 3.2.13) and to Khamyang as described below (see 6.4.3).

6.2.4.4 Present investigation

The present investigation found that the distribution of tones was the same as described by Banchob and Diller, with some slight alterations regarding the exact quality of each tone. As with both Banchob and Diller, the principal informants for this thesis are from Namphakey village.

Namphakey is much the largest Phake village, and it may be that tonal systems are less stable in other villages which are smaller and closer to the larger Khamti population of Arunachal Pradesh. This may explain the tone chart reported by Wilaiwan in Table (21). Further investigation of this is needed.

Based on data collected in 1996 - 2000, the following tone chart (Table 22) is presented for Phake:

	A	B	C	DL	DS
1	6	1	3	1	
2	2				
3					
4		5	4		

1	high level tone	mā¹	"shoulder"
2	high falling tone	mā²	"to come"
3	low, falling tone, glottalised	mā³	"mad"
4	low falling tone	mā⁴	"horse"
5	low tone	cā⁵	"for, to"
6	rising tone	mā⁶	"dog"

Table 22: Tones in Phake (present investigation)

Table (23) gives contours for the tones as spoken by Ee Nyan Khet on the syllable *maa*, except the fifth tone which is spoken on the word *me⁵* 'mother'. It is possible that because it was spoken on a different syllable, which has a higher vowel, the pitch might have been slightly higher than if all six tones had been spoken on exactly the same syllable. Unfortunately no such series of six tones can be found.

Tone no.	contour (Hz)	key feature	description
1	200-190	level	mid to high level tone with slight final fall
2	225-250-150	high level then falling	mid to high onset, with rise then sharp fall
3	160-120	glottal constriction	mid falling or level tone with strong glottal constriction

4	175-125	falling	mid to high onset with immediate sharp fall
5	180-160	low	mid to low onset with slight gradual fall
6	220-200-275	rising	low to mid onset with slight fall before sharp rise

Table 23: Tones in Phake

Table (24) lists the Assamese words and gloss for each of the five tones pronounced by Ee Nyan Khet on the syllable *maa*.

Assamese		Gloss	Tai Phake
কুকুৰ	kukur	'dog'	mā⁶
ঘোঁড়া	ghora	'horse'	mā⁴
বলিয়া	bɔliya	'crazy'	mā³
কান্ধ	kandhɔ	'shoulder'	mā¹
অহা	ɔha	'come'	mā²

Audio Link to Pronunciation of these words. (\wave files\phake\sounds\ma.mp3)
Link to Image of pitch contour chart for these words. (wave files\phake\wave images\ma-tones.mp3)

Table 24: Tones 1-4 and 6 pronounced on the syllable *maa* in Phake

In March 2002, Yehom Buragohain was recorded pronouncing words on the syllable *kaa*. This included the two words in Table (25), for the comparison of the first and fifth tones:

Assamese		Gloss	Tai Phake
যোৱা	dʒʊwa	'go'	kā¹
সমান	hɔman	'equal'	kā⁵

Audio Link to Pronunciation of these words (wave files\phake\sounds\ka-tones.mp3)

Table 25: Tones 1 and 5 pronounced on the syllable *kaa* in Phake

Table (26) presents the contours for these two tones:

Tone no.	contour (Hz)	key feature	description
1	205-200-220	level	mid to high level tone with slight fall and then small final rise
5	190-160-180	low	mid onset then gradual fall with slight final rise.

Link to Image of pitch contour chart for these tones. (wave files\phake\wave images \ka1ka5.bmp)

Table 26: Tone contours for tones 1 and 5 pronounced on the syllable *kaa* in Phake

Tone number 1 is clearly a level tone in citation. When informants are required to repeat it, the pitch of the tone may get higher. The key feature of this tone appears to be its level nature.

Frequency analysis of tone number 2 has shown that it rises before reaching a peak and then falls sharply. The key features of this tone seem to be that it is high, and falling, and that this fall does not commence at the beginning of the tone. In the production of the tone, by thinking of it as high level then falling, and repeating it back to the native speakers, the writer has been able to manage to say it correctly.

The most important distinctive feature of tone number 3 is clearly the glottal constriction which is very noticeable in citation form. It is sometimes pronounced without a clear fall, but never without glottalisation, as shown in (13)

13) ᜅᜒᜎᜓᜂᜌ᜔ ᜏ ᜋᜅ ᜏᜓ ᜋᜅ ᜋᜅᜈ ᜋᜅ
nam¹ pɔt¹ sām⁶ kā³ kʰau³ kā³ kon² kā³
number three kā³ rice seedlings rice person dance

ᜋᜅ ᜊᜒᜈᜒᜊ ᜋᜅ ‖
kā³ maiŋ² mī³ kā³
dance Butterfly kā³

'Number three is *kā³*, as in 'rice seedlings', 'a dancer', 'the Butterfly dance', that is *kā³.*'

(Phake text, *The Phake Tones*, 1-10-2-3, No. (3), spoken by Ai Che Let.)

An examination of the pitch contour for the last token of the word *kā³* in (13)[47] shows that beginning at about 115Hz, it falls and then begins to slowly rise, reaching about 125 Hz at the end of the word. From this, we may conclude that glottalisation is the only key feature of this tone, and that contour is not crucial.

Tone number 4 has a clear and direct fall. The most important difference between tone 4 and tone 2 seems not to be the level of the tone, but the fact that tone 2 rises or is level before falling, whereas tone number 4 falls directly. It is interesting to note that Phake speakers hear Khamyang tone number 4 (see 6.4.3) as being the same as Phake tone 2, showing that rise then sharp fall of tone 2 is its crucial feature.

The key feature of tone number 6 is that it rises. It may fall slightly or be level before commencing the rise.

6.2.4.5   Changed Tone in Phake

6.2.4.5.1   The questioning tone in Phake

In addition to the six citation tones discussed above in 6.2.4.4 there is a further tone, with a quite different contour, which has been observed in questions, as can be seen in example (14):

14) ᜋᜂᜃ᜔ ᜃᜒᜈ᜔ᜒᜉ ᜃ᜔ᜌ ᜏᜈ ᜌᜈ᜔ᜌ ‖
mauk¹ kə² kau² su⁶ han⁶⁽²⁾
pipe salt 1Sg 2Pl see.QN
"'Did you find my pipe of salt?'"
(Phake Story, *The Dolphin, the Crow and the Mosquito* 1-1-1-4, No (41)), told by Aithown Che Chakap)

The tone on the last word is not the citation tone for the word *han⁶* "see", but a rising-falling tone, here notated as -6⁽²⁾. I have decided not to use an arbitrary notation (such as -⁷) for this tone in order to show its special characteristics, as it has a quite a different function from the six citation tones.

---

[47] See \wave files\phake\wave images\kaa3.bmp in the electronic version of this thesis.

Aithown Che explained that "generally when there is a question, this type of tone is found." The image of the tone contour for the word $han^{6(2)}$ shows that the pitch rises from 200Hz to a peak of about 250Hz then falls back to about 240Hz.[48]

This tone is even clearer in certain elicited sentences, such as (15), where the questioning tone is found on the word kin^2, which has the second tone in citation.

15) ၮၟ   ၐၟ   က�Ⴡဃ   ॥
     khau³ ta¹   kin⁶⁽²⁾
     rice   will eat.QN
     'Will you eat rice?'
     (Phake sentence elicited from Ai Chanta)
     Audio Link for this sentence example (\wave files\phake\sounds\kin62.mp3)
     Link to image of wave file for kin⁶⁽²⁾ (\wave files\phake\wave
     images\kin62.bmp)

The frequency of the tone of kin⁶⁽²⁾ commences at 140Hz, rises to 180Hz and falls to 115Hz.

When a question is asked on a word which in citation form has the first tone, such as $kā^1$ "GO", the question tone is simply a rising tone, as in (16)

16) ၔ   ၱဃၟ   ၵ   ၐၟ   ၒၣ   ॥
     maü² nɔn²   nī² kā⁶   ma¹ nai³
     2Sg   sleep good GO.QN   today
     "Did you sleep well today."
     (Phake sentence elicited from Ai Chanta)
     Audio Link for this sentence example (\wave files\phake\sounds\ka-qn.mp3)
     Link to image of wave file for kā⁶ (\wave files\phake\wave images\ka-qn.bmp)

It appears that some tone combination is occurring in some words. There may be a question morpheme -⁶, where the rising intonation indicates a question. It may be that this rising tone is transferred from the question particle $nɔ^6$ (see below 8.6.2.1). If this is so, the process might be characterised as in Table (27):

---

[48] See \wave files\phake\wave images\han62.bmp in the electronic version of this thesis.

Word in Citation form	Question morpheme	Word in question form	Tone change process
kin²	-⁶	kin⁶⁽²⁾	-2 + -⁶ = -⁶⁽²⁾
nī²	-⁶	nī⁶⁽²⁾	-2 + -⁶ = -⁶⁽²⁾
kā¹	-⁶	kā⁶	-1 + -⁶ = -⁶

Table 27: The Questioning Tone in Phake

If the suggestions in Table (27) are correct, it still does not explain how a question morpheme -⁶ when combined with a word whose citation form has the sixth tone, is pronounced [han⁶⁽²⁾] as in example (14) above. Perhaps the question form which was originally found only with words that had the 2nd tone in citation is now being generalised to words of all citation tones.

This question tone seems to be an innovation in Phake. It is hoped that further study of this phenomenon will be possible. The syntax of questions is discussed below in 8.6.1.

6.2.4.5.2 The Tai Phake negative tone

In addition to a questioning tone, there appears to be a negative tone which is always expressed with the 6th tone. Banchob (1987) first reported this, as in (17).

17) ၔ   ၐၟ   ၐၭၐၟ   ॥
     mā⁵   kā⁶   lā² lā²
     NEG   go.NEG PRT
     "Not to go at all."
     (Banchob 1987: 359)

Diller (1992:19) also reported this as "a morphophonemic tone sandhi rule, involving a rather spectacular contour reversal, regularly operates after the proclitic negative marker [m-] (but not with the full form maw⁵) in the case of tone 3 verbs." Diller's tone 3 is tone 2 in Banchob's system, but as we can see from (17), this use of the 6th tone for negation applies not only to verbs whose citation form is in tone 2, but to other tones as well. The citation form of 'go', which is negated in (17), is $kā^1$.

Example (18) clearly differentiates the positive *yaŋ²* with the negative *yaŋ⁶*.

18) ကှက် ဃၞၚ် ႃ႞ှက် ၞ ဃၞၚ် ။
    kon²   yaŋ²   hən²   ma¹   yaŋ⁶
    person be   house NEG   have.NEG
    'There are people but no houses.'
    (Phake Story, *Chaw Mahosatha meets Nang Pingya*, 1-1-1-3, No. (26), told by
    Ee Nyan Khet)

Not all negative words have the negative tone. In example (19), a compound is
negated but the tone of the compound is not affected:

19) ႃ႞ှ် ၞ ႃ႟ၞြက်
    mün²   ma¹   tɔ²   huk¹
    3Sg   NEG   weave
    'She didn't weave!'
    Audio Link for this sentence (\wave files\phake\extracts\Aithown Che\2.mp3)

When this compound was split, however, as in (20), the negative tone was observed:

20) ႃ႞ှ် ၞြက် ၞ ႃ႟ၞ
    mün²   huk¹   ma¹   tɔ⁶
    3Sg   loom   NEG   weave.NEG
    'She didn't/doesn't weave.'
    Audio Link for this sentence (\wave files\phake\extracts\Aithown Che\1.mp3)

Further investigation remains to be done on changed tone in Phake, to identify in
which circumstances it is possible and in which it is not. Several reasons were
advanced by Phake informants themselves as to why this changed tone was used.
These are discussed below in the section of the syntax of negation in 8.6.3.

6.2.4.6 Tone assignment in loan words

Tai Phake words which are reflexes of proto Southwestern Tai words usually have a
tone which is the reflex of the proto tone of that word. However, words borrowed into

---

a tonal language have to acquire tone, and the process by which this is done in the Tai
languages of Assam is not yet fully understood.

The late Aimya Khang Gohain very carefully notated the tones in that pai : of
Grandfather teaches Grandchildren which he assisted me to analyse. Several loan
words from Phake occur in that work, and some of these are listed in Table (28)

Tai Phake word	Phonemic transcription	English Gloss
ဃၞၞၟၠႃ႞ၬက်	sak¹ ka¹ rit¹	"era"
ႃ႞ၬၟၠ	tā² ti¹ yā⁴	"third"
ၞၞႃ႞ၬႃ႞	a¹ ni¹ cā⁴	"changing"

Table 28: Tone assignment in loan words from Pali

In each case, the unstressed middle syllable is first tone, raising the question of
whether there are toneless syllables in Tai Phake, and whether the first tone is in some
sense a default tone. In this regard it is interesting to note that in Chau Khouk
Manpoong's system of tonal marking for Khamti, there is a special mark for such
unstressed syllables. This mark is not one of the five citation tones (see 7.3).

Discussion with other informants suggests that the normal pattern of tone assignment
for polysyllabic loans, often from Pali, is that the first few syllables will be 1st tone,
and the last syllable will either be 1st tone. 2nd tone or 4th tone. The 3rd, 5th and 6th
tone seem rarely to be used in such borrowed.

Example (21) includes a number of these (indicated by bold type):

21) ၬၠၞ် ဃၟၠၬကြက်ၩ ဃၟၞၠ ဃၟၞၮ ဃၟၞၟၠ
    woi¹ cem²   si¹ li¹ kuk¹ trā⁴   ya¹ ma¹ tɔ²   ya¹ ma¹ kā²   ya¹ ma¹ māŋ⁴
    bless   **Silikuktra**   **Yamato**   **Yamaka**   **Yamama**

ဝၒ် ႁၮ် ၯၞ်ၞ် ၔ�528 ၒၒ် ၓ၁ၞ် ၯၞ်ၝ် ॥

tī⁵	an²	pum¹	məŋ²	phāi²	het¹	cau³
which	CLF	reign	country	demon	make	RESP

'Blessings to Silikuktra, to Yamaka and Yamama, who reign in the Kingdom

of the Demons.'

(Phake Prayer, *Prayer of Blessings* 1-3-1-3, No. (25), told by Sam Thun Wingkyen.)

6.3 Aiton

The following information about Tai Aiton phonology is largely based on a word list containing 1924 items collected in ဝၒ်ၞၨၓၕၙ (baan³ nam³ thum³) Duburoni village, Golaghat district, Assam in December 1999 and January 2000. The complete word list appears as the *Tai Aiton Dictionary*[49]

The informant for this word list was Chaw Bidya Thoumoung, a 35 year old male teacher who had been ordained as a monk for several years, is fully literate in Tai script and is educated in the literature of the Tai Aiton. His parents are both Tai Aiton and one of his maternal uncles is the Tai pandit Nabin Shyam Phalung, and another is the expert story-teller, Sa Cham Thoumoung.

In some cases the data from the word list has been supplemented by information gained from the analysis of the texts which form the basis of this thesis. In the opinion of the writer, the phonological sketch below represents the present standard in the Aiton villages.

6.3.1 Consonants

The Aiton phoneme inventory is presented in Table (29). The symbols adopted for use in the Aiton-English dictionary have been used here. They are basically the same as those used by Banchob (1987) for Phake. Additionally, however there are two phonemes in Aiton, <b> and <d>, which are not found at all in Phake, and I have

---

[49] See \lexicon\aiton\aitdic.doc in the electronic version of this thesis.

decided to use <ʒ> in place of <y>, because of the strong fricativisation which characterises this consonant in Aiton. This will be discussed in detail below in 6.3.1.1.6.

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiced Stops	b	d			
Voiceless unaspirated stops	p	t	c	k	
Voiceless aspirated stops	ph	th		kh	
Nasals	m	n	ñ	ŋ	
Voiceless fricative		s			h
Voiced fricative			ʒ		
Semi vowel	w				
Rhotic Approximant		r			
Lateral Approximant		l			

Table 29: Consonant Phonemes in Aiton

6.3.1.1 Realisation of the Consonants

6.3.1.1.1 Voiced Stops

The voiced stops, /b/ and /d/, in Aiton are reflexes of the preglottalised initials *ʔb and *ʔd in proto Southwestern Tai. In every other variety of the Shan group which is still spoken, the reflexes of *ʔb and *ʔd have merged with the reflexes of other proto Tai consonants and the distinctions have been lost.

For example, in Phake, the reflexes of *ʔb and *ʔd have merged with /m/ and /n/ respectively. In Khamyang, they have merged with /m/ and /l/ respectively. On the other hand, the reflexes of these two proto phonemes are realised as /b/ and /d/ respectively in both Standard Thai and Lao, and possibly to some extent in Ahom, which had a distinctive glyph for /d/, see below 6.5.1.

In Bidya Thoumoung's word list, most words with /b/ or /d/ initials are also realised with the homorganic nasal initial, as in the examples listed in Table (30):

Assamese		Gloss	Tai Aiton
সুঙা	huŋa	to smell	dum², num²
চোৱা	suwa	to see	duu², nuu²

Audio Link to Pronunciation of these words, (\wave files\aiton\sounds\d-.mp3)

Table 30: Initial /d/ in Aiton

In careful speech, both of these voiced stops are realised as slightly prevoiced [ᵐb] and [ⁿd], as they often are in Tai languages where these phonemes are still found. This is clearly audible in Bidya Thoumoung's pronunciation of *dum²*.

In present day Aiton, almost all of the words with initial /b/ or /d/ have variants in initial /m/ and /n/ respectively. This may be due to the influence of the writing system, in which there is no separate character for /b/ or /d/ (see below 7.2.1). Alternatively it may be due to the influence of the other Tai varieties in which these sounds do not occur contrastively.

This variation with the homorganic nasal stops, means that there is a reduced functional load for both voiced stops, at least in regard to words of Tai origin. The extent to which these two phonemes are perhaps increasing in use for words borrowed from Assamese, such as names, is not clear. One example of the persistence of /b/ with borrowed words is the name of the principal informant, Bidya Thoumoung, where *bidya* is an Assamese name based on the Sanskrit word *vidyā* 'knowledge', (Sanskrit *v > *b in eastern Indo-Aryan).

Banchob Bandhumedha, in her *Aiton–English–Thai Dictionary* (ms) also includes six words with initial /g/. All of these words have initial /gr-/ clusters and are all onomatopoeic. They include *griŋ² grɛŋ²* 'the sound of a thunder bolt' and *grɔ² grɔ²* 'the sound of children sleeping'. Since all Aitons are bilingual in Assamese and Assamese has a distinctive phoneme /g/, it is clear that they can pronounce this sound. In this study, however, no evidence has been found of /g/ as a phoneme.

6.3.1.1.2 Voiceless unaspirated stops.

The bilabial, dental/alveolar and velar stops, /p/, /t/ and /k/ respectively are found both in initial position and in syllable final position. In initial position each of them is articulated with slight force. For Khamti, Harris (1976) said of this, referring to /t/ that it is "probably always accompanied by a simultaneous weak glottal stop [t]."

Assamese		Gloss	Tai Aiton
উজ্জল	udʒɔl	bright	tɔŋ¹
পাত	pat	big leaf, banana leaf	tɔŋ²
পেট	pɛt	stomach	tɔŋ³

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\t-.mp3)

Table 31: Initial /t/ in Aiton

In final position, each of these phonemes is generally unreleased as [p̚], [t̚] and [k̚]. Table (32) gives some examples of final /-k/:

Assamese		Gloss	Tai Aiton
বেৰা মৰিওৱা	bɛra mɔriuwa	to hammer	tɔk¹
জোখা	dʒukʰa	to weight	tɔk³

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\k-final.mp3)

Table 32: Final /k/ in Aiton

The palatal stop /c/ is usually released as an affricate [tɕ]. Harris (1976:117) reported that the cognate sound in Khamti "represents a voiceless laminal-prepalatal (or alveolo-palatal) affricate [tɕ]. The narrow grooved fricative release is very short in duration." The initial realisation of the palatal stop is exemplified in Table (33). It does not occur in syllable final position.

Assamese		Gloss	Tai Aiton
মুচা	musa	to wipe	cit³
আত	at	seven	cit¹

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\c-.mp3)

Table 33: Initial /c/ in Aiton

6.3.1.1.3 Voiceless aspirated stops.

Voiceless aspirated stops are often realised as voiceless fricatives at the same place of articulation, as shown in (22). The exception to this is /th/, which is unable to have a alveolar fricative allophone because there is a separate phoneme /s/.

22)   /ph/   →   [ɸ], [pʰ]
       /kh/   →   [x], [kʰ]

For the corresponding phonemes in Khamti, Harris stated that /ph/ "represents a voiceless aspirated bilabial affricate [pɸʰ]" (1976:116) and /kh/ "represents a voiceless aspirated dorso-velar affricate [kxʰ]" (1976:118). In Aiton, at least, whilst there may be some affrication of the fricative realisation of these phonemes, it does not appear to be as strong as Harris' statements suggest. Furthermore, this phoneme certainly also has a voiceless stop allophone.

Table (34) presents examples of the phonemes /ph/ and /kh/.

	Assamese	Gloss	Tai Aiton
বিচৰা	bisɔra	to seek	khaa¹
চাকৰ	sakɔr	servant	khaa³
টিকা	tika	thigh	khaa¹
কাপোৰ	kapʊr	cloth	phaa³
বেৰ	bɛr	wall (of the house)	phaa¹
আকাশ	akʰaĥ	sky	phaa³

Audio Link to Pronunciation of the words with initial /kh-/ (\wave files\aiton\sounds \kh-.mp3)
Audio Link to Pronunciation of the words with initial /ph-/ (\wave files\aiton\sounds \ph-.mp3)

Table 34: initial /kh/ and /ph/ in Aiton

When praying or intoning Pali language, Tai Aiton speakers follow the Burmese practice of dentalising the reflex of Pali *s. In Burmese this reflex is pronounced as [θ], which is not a phoneme in any of the Tai languages. Among the Tai Aitons, this

reflex of Pali *s is pronounced as a dental stop which is slightly fricativised. It would probably be regarded as /t/ by the Aitons.

Banchob Bandhumedha, in the *Aiton-English-Thai Dictionary*, distinguishes two phonemes /f/ and /ph/. Many of the words in the /ph/ section of her dictionary are transcribed with both initial phonemes as variants, as *phā⁴/fā⁴* 'the sky'. Those words listed under /f/, on the other hand, do not have a /ph/ initial as a variant, such as *fɛ²*'to be born, as cattle' and *fāt⁶*'to be astringent', the latter word being a reflex of a word with Proto Southwestern Tai initial *f-. No evidence has been found in the present investigation for suggesting that /ph/ and /f/ are different phonemes.

6.3.1.1.4 Nasals

Of the four nasal phonemes, /ñ/ is the least common, and the most restricted in its distribution, in that it only occurs in word initial position. It is frequently found in final position in writing, following the vowel /i/ or /u/. After /i/ it is pronounced /n/ and causes the vowel to be pronounced /ɛ/. When it follows /u/ it is pronounced /i/ as discussed below in 7.2.2. Final -ῳ /-ñ/ in writing is never pronounced as /ñ/.

Words with initial /ñ/ are more frequent in Aiton than in Phake. Several of the Aiton words with initial /ñ/ have variants, beginning with other letters, as shown in Table (35).

Assamese		Aiton		Alternate		
দৰব	dɔrɔb	ဢၣ	ñaa³	ဢၣ	ʒaa³	medicine
ঊনা	ñɔna	ဢၣS	ñin²	Ƨɔʃ	ŋin²	hear
ডাঙৰ	daŋɔr	ဢ	ñauɯ¹	ဢ	ʒauɯ¹	big

Audio Link to Pronunciation of the first two words in Table (35) (\wave files\aiton \sounds\ny-.mp3)

Table 35: some variations with initial /ñ/

When comparing initial /ñ-/ and initial /ŋ-/, with reference to the words for 'wisdom', and 'draw a line', Bidya Thoumoung spoke sentence example (23):

23) ကꩻ ꩡꩻ ꩠꩻ ꩠꩻ ꩠꩻ ꩡꩻ ꩠꩻ ꩠꩻ ꩡꩻ ꩠꩻ ॥

an² nɛ³ ŋān¹ ŋān¹ an² nɛ³ ñān¹ an² nɛ³ ŋān¹
CLF DEF draw draw CLF DEF wisdom CLF DEF draw

'This one is 'to draw a line', *ŋān¹*, and this one is 'wisdom', this one is 'to draw a line.'

(Aiton sentence spoken by Bidya Thoumoung)

Audio Link for this sentence example (\wave files\aiton\sounds\ny-ng.mp3)

The other three nasal consonants are very widely distributed in both initial and final positions. In final position. ... are sometimes pronounced somewhat longer than in initial position, such as the various words pronounced on the syllable *tɔŋ* in Table (31) above. A spectrogram analysis of *tɔŋ¹* 'bright' shows that the vowel lasts for about 0.3 seconds and the final nasal coda for another 0.3 seconds[50].

6.3.1.1.5 Fricatives.

Of the many fricatives posited in proto Southwestern Tai (see above 6.1.1.1) two have been maintained unaltered in Aiton. These are the voiceless alveolar fricative /s/ and the voiceless glottal fricative /h/. However, as will be argued in 6.3.1.2 below, some of the other consonants are becoming fricativised.

For example, as mentioned above, in 6.3.1.1.3, the voiceless aspirated stops are often released as fricatives in the same place of articulation. Furthermore there is strong fricativisation of both semivowels, which is so strong in the case of the palatal semivowel /j/ that I have chosen to notate it as <ʒ> rather than <y>.

There is sometimes palatalisation of the fricative /s/ when it precedes a front vowel. In example (24.1), the /s/ in *sii¹* 'four' is palatalised, whilst that in *saa¹* 'PRT' in (24.2) is not.

---

[50] The spectrogram can be viewed at \wave files\aiton\wave images\tong1.bmp. The sound file is at \wave files\aiton\sounds\tong1.mp3.

24.1) ꩠꩻ ꩫ ꩬꩠꩬ ꩦ ꩫ ꩬ ꩠꩻ

an² nai³ pha¹ ta¹ wii¹ hau² nai³ sii¹ cun²
CLF this earth 1Pl this four part

ꩬ ꩠꩻ ꩫ ꩠꩠꩻꩨ ꩫ ॥

sii¹ cun² nai³ u¹ taŋ¹ ku³ ru³ nai³
four part this Utangkuru this

'This earth of ours has four parts, and one of those parts is Utangkuru.'

24.2) ꩠꩻꩨ ꩬꩠ ꩦꩨ ꩬꩠ ꩬꩠꩠꩻ ꩬꩠ ꩦꩨ ꩬ ॥

tun³ mai³ kɔ³ luŋ¹ saa¹ laŋ¹ din² kɔ³ luŋ¹ saa¹
tree LINK yellow PRT ground LINK yellow PRT

'(There) the trees were yellow and the ground was also yellow.'

(Aiton text, *Why Buddha was born in this world>*, 2-4-3-1, Nos. (2)-(3), told by Sa Cham Thoumoung.)

However in Bidya's list, there is little appreciable difference in initial /s/ when followed by /a/ or by front vowels, as indicated in Table (36):

Assamese		Gloss	Tai Aiton
ভৱৰা	rʊwa	to stay	sau²
গাভৰু	gabʰɔru	female	saau¹
বিংশ	bîh	twenty	saau²
খোটা	khʊta	post	sau¹
শীল	hil	morality	sin¹
মেখেলা	mɛkhɛla	skirt	sin³
চিঞৰা	sinɔra	to roar	sɛn¹

Audio Link to Pronunciation of the words on the syllables *sau* and *saau* (\wave files\aiton\sounds\sau.mp3)

Audio Link to Pronunciation of the words on the syllables *sin* and *sɛn* (\wave files\aiton\sounds\sin.mp3)

Table 36: Initial /s/ in Aiton

6.3.1.1.6 Semivowels.

As discussed above, both semivowels are often realised as homorganic voiced fricatives, with /j/ being realised as [ʒ] and /w/ as [β]. This tendency is so strong in the

case of the palatal semi-vowel that the symbol <ʒ> has been chosen for use in the Aiton dictionary and in the Aiton transcriptions in this thesis.

Assamese		Gloss	Tai Aiton
আচোৰা	asura	to scratch	wak³
ধুনা	dhuna	wax, gum	waak¹
নাও বোৱা	nau buwa	to row	wak³
আছে	asɛ	Be	ʒaŋ²
দাড়ি ধৰা	daɲi dʰɔra	to shave	ʒaŋ¹
এঠা	ɛtʰa	Gum	ʒaaŋ²

Audio Link to Pronunciation of the words with initial /w-/ (wave files\aiton\sounds \w-.mp3)
Audio Link to Pronunciation of the words with initial /ʒ-/ (\wave files\aiton\sounds \yang.mp3)

Table 37: Initial /w/ and /ʒ/ in Aiton

In the other Tai languages, variations in the phonetics of these two semivowels have been noticed. For example, Harris stated that in Khamti:

"/w/ represents a voiced labio velar approximant [w] or a voiced bilabial approximant [ʋ]" (1976:116) and

"/y/ represents a voiced front-palatal approximant [j] or a voiced front-palatal fricative [ʝ]. In syllable final position this symbol represents a close front unrounded vowel [i]" (1976:117).

The symbol /v/ (voiced labio-dental fricative) is frequently used for the reflex of proto *w in some Tai varieties. Jerry Edmondson (pers. comm.) has assured me that the realisation of this phoneme in some Shan languages is indeed the labio-dental /v/, not the bilabial /β/ which is observed for Aiton.

A good example of the variation in the use of the palatal semivowel (/j/) phoneme is found in the following example, where the text has been transcribed narrowly:

25) ꨓꩻ ꩺ ꨕ ꨔ ꨣꩻ

ʒaŋ²   uu¹   tii²   thai³   yaŋ²
have   STAY   place   here   have
'Yes there is one here.'

(Aiton text, *Explanation of the Spirit house>*. 2-8-1-6, No. (69), spoken by Sui Khong Thoumoung)

The phoneme /ʒ/ is clearly a reflex of the proto Southwestern Tai /*j/ (see 6.1.1.1) and as such it might be argued that it should be written as <y>. In the various Aiton texts which form the basis of this thesis, it is pronounced [ʒ] in most spoken texts although Aitons will pronounce [y] when reading from manuscripts. The decision to write this phoneme as <ʒ> was taken in order to reflect current spoken practise. However the symbol <y> is used for this phoneme when it appears in initial consonant clusters (see below 6.3.3.1). All of the words in which /y/ appears as the second member of an initial consonant cluster are borrowed words, mostly from Burmese.

6.3.1.1.7 Approximants

Two approximants are found in Aiton, the lateral approximant /l/ and the rhotic approximant /r/. The rhotic approximant is often not found as a full phoneme in languages of the Tai family, although it is present in the writing of almost all the Tai varieties and some type of rhotic is often pronounced in careful speech. In Standard Thai, for example, trill [r] is realised in careful speech where in everyday speech the /r/~/l/ contrast has been largely neutralised.

In Aiton, however, /r/ is definitely present as a full phoneme both for initial position and as the second member of an initial consonant cluster (see below 6.3.3.1). Unlike in Standard Thai, the rhotic in Aiton is not a trill but is definitely an approximant with fairly strong retroflexion. It would be best phonetically described as either [ɹ] or [ɻ]. In this thesis, however, it will notated phonemically as /r/.

Assamese		Gloss	Tai Aiton
ভগা	bʰoga	broken	raak[1]
ঘিণকৰা	gʰin kɔra	to hate	raŋ[3]
যিকোনোবাহ	dʒikʊnʊbah	nest	raŋ[2]

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\r-.mp3)

Table 38: Initial /r/ in Aiton

When reading the words in Table (38), Bidya Thoumoung stated that there is a variant in initial /h/ for several of these words, as shown in Table (39):

Aiton word		Alternate		Phake	Stanard Thai	Gloss
ၵၢၵ်	raak[1], rak[1]	ႁၢၵ်	hak[1]	hak[1]	rák	'broken'
ၵၢင်	raŋ[2]	ႁၢင်	haŋ[2]	haŋ[2]	raŋ	'nest'

Table 39: Variations of initial /r/ and /h/ in Aiton

It is not, however, the case that all words which are /r/ initial in Standard Thai have /r/ initial in Aiton, or that all words which have initial /r/ in Aiton are found in Standard Thai. Furthermore, there appear to be are a number of words which are /r/ initial in Aiton for which no variant with initial /h/ is tolerated by the community.

The preservation of initial /r/ in Aiton as a full phoneme is most significant, especially in light of the preservation of /b/ and /d/, discussed above in 6.3.1.1.1. As was the case with those voiced stops, /r/ is also a phoneme in Ahom. On the basis of the evidence presented thus far, it appears that Aiton is indeed very close to the ancient Ahom language. In every other language of the Shan group, the reflexes of proto /*r/ have merged with /h/ or some other phoneme. It is only Ahom and Aiton in which this does not occur.

6.3.1.1.8 Glottal Stop

As discussed above in 6.1.4, I do not propose to mark a glottal stop in front of initial vowel segments. Aiton does not require the positing of a glottal stop phoneme in

order to be able to explain its phonemic system or its system of tones, and I believe that it would not be notated by scholars except that they have taken into account the theories of syllable structure 6.1.3 and the proto Tai tonal system as described by Gedney (1972) see 6.1.5.3.

The marking of a glottal stop in Aiton would not appear to add anything to our knowledge of the language. In vowel initial words, a slight glottal onset is present for phonetic reasons, but no more than in many languages in which no initial glottal stop phoneme is posited. This glottal onset is apparent in the words given in Table (40):

Assamese		Gloss	Tai Aiton
দ্বিতীয় জী	dvitiyɔ zi	second daughter	ii[1]
মৰম	mɔrɔm	pity	ii[1] duu[2], ii nuu[2]
পঞ্চম জী	paɲcɔm zi	fifth daughter	uk[1]
খাল	kʰal	centre of the palm of the hand	uŋ[5]

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\v-init.mp3)

Table 40: Vowel initial words in Aiton

In connected speech any glottal onset is hardly audible, as demonstrated in example (26):

26) တၢၼ်ႃလုင် ၵ တ ႁၢမ် ၼၢ ... ၵ တ ႁၢမ် ၼ ၼ်ႂၢ
   taŋ[1] luŋ[1] hauɯ[3] nai[3] uu[1] khyɛm ... hauɯ[3] nai[3] uu[1] nii[2] khyɛm[2] saa[1]
   all   GIVE get live good ... GIVE get live good good
   'May all be happy.'
   (Aiton manuscript, *Calendar for the year 2001/2002*, No. (2.1), told by Sui Khong Thoumoung.)

6.3.1.2 Possible reassignment of phoneme symmetry

It is possible that the weakening of voiceless aspirates to voiceless fricatives and the strengthening of semi vowels to voiced fricatives has so changed the phoneme inventory that it would be better to express it as in Table (41):

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiced Stops	b	d			
Voiceless unaspirated stops	p	t	c	k	
Voiceless aspirated stops		th			
Nasals	m	n	ñ	ŋ	
Voiceless fricative	ɸ		ʃ	x	h
Voiced fricative	β		ʒ		
Rhotic Approximant		r			
Lateral Approximant		l			

Table 41: Realigned Aiton Phoneme Chart

Note that to make this table more symmetrical, one might think that /th/ should fricativise to [θ], but there is no evidence of this at this stage. As noted above in 6.3.1.1.3, in intoning prayers the reflex of Pali *s is pronounced as a dental stop which is slightly fricativised. Whether this fricativisation and dentalisation might ever extend by analogy to /th/ remains to be seen.

Table (42) presents some minimal pairs of initial /ph/ - reanalysed as /ɸ/ and /w/, reanalysed as /β/ to show the way in which, synchronically at least, it might be possible to regard these two phonemes as a voiceless~voiced pair.

Assamese		Gloss	Tai Aiton	Reanalysed
আতি	dʒati	caste	phan²	ɸan²
দুখীয়া	dukʰiya	poor	phaan¹	ɸaan¹
হগৰী পহু	hukɔri pɔhu	type of deer	phaan²	ɸaan²
মিঠা	mitʰa	sweet	waan¹	βaan¹
দিন	din	day	wan²	βan²

Audio Link to Pronunciation of the words with initial /ph/~/ɸ/ (\wave files\aiton\sounds \phan.mp3)
Audio Link to Pronunciation of the words with initial /w/~/β/ (\wave files\aiton\sounds \wan.mp3)

Table 42: A new voicing pair in Aiton, /ɸ/ and /β/

6.3.2 Vowels

It is the claim of this thesis that the Aiton vowel system contains seven contrastive vowels, as in Table (43):

i		ɯ	u
ɛ			ɔ
	a, aa		

Table 43: Aiton vowels

It is interesting to observe that this is the same distribution of vowels as that posited for the vowels of Proto Southwestern Tai by Jonsson (see above 6.1.1.3). Before turning to the evidence from the present study, we will briefly examine previous work on Aiton phonology by Banchob Bandhumedha and Diller (1992).

6.3.2.1 Previous descriptions

6.3.2.1.1 Banchob

Based on her fieldwork in the Aiton villages in the early 1970s, Banchob Bandhumedha wrote a manuscript *Aiton–English–Thai Dictionary*. Unfortunately her work was never completed and has never been published. Part of this dictionary has been analysed.

Table (44) lists the vowel symbols which have been found in the *Aiton–English–Thai Dictionary*.

i,ī	ü,ǖ	u,ū
e,ē		o
ɛ		ɔ
	a,ā	

Table 44: Aiton vowels found in the *Aiton–English–Thai Dictionary* (Banchob ms)

Length is only contrastive with the vowels <a> and <ā>. In the case of all the other vowels, Banchob has used a long vowel symbol only in vowel final or open syllables. However it does not appear that the contrast between <e> and <ɛ> is phonemic. At

least for words beginning with /k/, the symbol <e> is used, with only one exception, for open syllables, and the symbol <ɛ> in closed syllables.

Furthermore, all of the words which have the vowel /e/ in Phake are notated in Banchob's *Aiton–English–Thai Dictionary* with <i>, and the words which are notated with /e/ all have /ɛ/ (or /ai/ before a velar, see above 6.2.2.2) in Phake. Some of these are listed in Table (45):

Aiton–English–Thai Dictionary	Phake Banchob (1987)	Gloss
kiŋ²	keŋ²	iron tripod
kɛŋ⁹	kaiŋ³	to use sticks to clean the bowels
ke⁶	kɛ¹	old, mature
ke⁹	kɛ³	to loosen
khiŋ⁸	a¹ khiŋⁱ	time
khiŋ⁸	kheŋ¹	to compete with
khiŋ¹	kheŋ¹	shelf
khɛŋ⁷	khaiŋ⁶	hard
khɛŋ⁹	khaiŋ⁵	shin

Table 45: Comparison of Aiton front vowels in Banchob's *Aiton–English–Thai Dictionary* with the *Phake–English–Thai Dictionary* (Banchob 1987).

There are some contrasts observed between the back vowels <o> and <ɔ>, such as *ko⁹* 'a measure word for persons' and *kɔ⁹* 'too, a word for emphasising'. A similar contrast was found in the word list elicited from Bidya Thoumoung (see below 6.3.2.2.3). However in both Banchob's *Aiton–English–Thai Dictionary* and in Bidya Thoumoung's list, the only words which appear with [o] are those which in Phake are notated with <ɔ>, and all of the words which in Phake are notated with <o> are notated <u> by Banchob. It appears therefore that the Aiton recorded by Banchob showed a merge of */u/ and */o/, suggesting a seven vowel system, as given above in Table (43).

Fortunately, some recordings made by Dr. Banchob have survived[51]. Her principal informant, the late Ai Ne Phalung, or Mohendra Shyam, told the *Story of the ghosts and the precious opium pipes* which includes the words *siŋ³* 'end' (*siŋ³* in Phake) and *tiŋ³* 'hit' (*teŋ³* in Phake). The spectrogram analysis (see \wave files\aiton\wave images \ting3.mp3)shows that these two words are clearly spoken on the same vowel, namely [i].

On the other hand, Banchob's observation that the words for 'human classifier' and 'LINK' were not the same, seems to be confirmed by examining *Story of the ghosts and the precious opium pipes*. Example (27) contains a token of *kɔ³* 'human classifier' and example (28) *kɔ³* 'LINK'.

27) ꩡꩳ ꩬ ꩍ ꩬ ꩫ ꩡꩳ ꩃꩰ ꩉꩰ ॥
    pɔ²   thau³   mɛ²   thau³ sɔŋ¹ kɔ³ kun²   phaan¹
    father old   mother old   two   CLF person poor
    '(Once upon a time) there was an old man and an old woman.'

    (Aiton Story, *Story of the forest ghosts and the opium pipes*, No. (1), told by Mohendra Shyam).

28) ꩃꩰ ꩬ ꩪꩰ ꩫꩰ ꩡꩰꩃꩰ
    kun²   thau³   an²   nan³   khun¹ kɔn¹
    person old   CLF   that   again

    ꩡꩰ ꩡꩳ ꩳꩶ ꩩ ꩪꩰ ꩉꩰ ꩫꩰ ꩡꩰ ॥
    hau²   kɔ³   saŋ¹ daa¹   kan²   phan² nai³ sii¹
    1PI   LINK if   quarrel RECIP like   this PRT

    ꩡꩰ ꩡꩳ ꩬꩰ ꩮꩰ ꩯꩰ ꩪ ꩡꩰ ॥
    hau²   kɔ³   ta¹   maak¹ kai³ kai³ wai³ sii¹
    1PI   LINK will   rich   quickly   KEEP PRT

    'That neighbour said, "If we quarrel like this we too will become rich very soon.'

    (Aiton Story, *Story of the forest ghosts and the opium pipes*, No. (25), told by Mohendra Shyam).

An audio file with the words [ko³] 'CLF and [kɔ³] 'LINK' has been made using the tokens of these words from examples (27) and (28). It can be found at \wave files\aiton\sounds \ko-mohen.mp3, and shows that, phonetically at least, the vowels are not the same. This is confirmed by the spectrogram analysis (\wave files\aiton\wave images\ko-mohen.bmp). So far this is the only minimal pair that can be found suggesting a contrast between the mid-close and mid-open back vowels.

6.3.2.1.2 Diller (1992)

Diller (1992:17) presents a ten vowel system, stating that "Ten vowel contrasts common in many Tai varieties apply to Khamti and also to Phake and Aiton:

i   ɯ   u   e   ɤ   o   ɛ   a   aa   ɔ ".

This is a different finding from that in this thesis. Diller presents a story, *Ghosts of the jungle* (1992:26-38), which is very similar to *The story of the forest spirit and the first daughter*. Table (46) lists all the words notated by Diller with [u] or [o] in the *Ghosts of the jungle*:

Tai word	Phake form	Form in Diller (1992)	Proto form	Gloss
ကွၚ်	kon²	kon	*konA2	person
လꩳၚ်	loŋ⁶	loŋ³	*hluaŋA1	big
ꧧꩫ်	khut¹	khut³	*khutDS1	to dig
လꩳꩫ်ꩬꩶ	luk³ sãu⁶	luk-saaw³	*lukDL4	daughter
ꧨ	ũ¹	ʔuᵇ	*ʔuuB3	to stay
လꩳၚ်	loŋ²	luŋ	*loŋA4	descend
လꩳꩫ်	luk³	luk		from
တꩭ	to²	tu	*tuaA2	CLF
ꧭꩬ်	ho⁶	ho³	*huaA1	head
ꧩꩬ	um³	ʔom²	*umC3	to hold
ꧭꩳ်	huŋ⁶	hoŋ³	*huŋA1	to cook
ꧨ	nũ⁶	nu³	*hnuA1	rat

Table 46: Words notated by Diller with [u] or [o] in the *Ghosts of the jungle* (Diller 1992:26-38)

There are number of irregular correspondences here. Of particular interest is the fact that some words which are have /*u/ as the nucleus being notated by Diller with [o] as *hoŋ³* 'to cook'., while some words which have /*o/ as the nucleus are notated with [u], as *luŋ* 'descend'. This suggests that the distinction between [u] and [o] might not be phonemic.

6.3.2.2 Present Investigation

As discussed above in 6.3.2, the finding of this thesis is that there are seven contrastive vowels in Aiton, with one of the contrasts being one of length between /a/ and /aa/. This is restated in Table (47):

i	ɯ	u
ɛ		ɔ
	a, aa	

Table 47: Vowel phonemes in Aiton

The relation between Aiton and Phake vowels (see 6.2.2) is shown in (29)

29) Phake [i,e]   >   Aiton [i]
     Phake [ɛ]   >   Aiton [ɛ]
     Phake [ɯ,ə]   >   Aiton [ɯ]
     Phake [a,aa] >   Aiton [a,aa]
     Phake [u,o]   >   Aiton [u]
     Phake [ɔ]   >   Aiton [ɔ], [o]

The symbols <ɛ> and <ɔ> have been chosen for the Aiton mid vowels rather than <e> and <o>, because <ɛ> and <ɔ> in Aiton correspond with <ɛ> and <ɔ> in Phake.

6.3.2.2.1 Front vowels

Given the divergence between the various reconstructions of proto Southwestern Tai vowels, the following discussion assumes Fang Kuei Li's reconstruction (see 6.1.1.3) of a nine vowel system. In Aiton, proto */e/ has been raised to /i/, thus neutralising a

contrast which is still very clear in Phake. Table (48) illustrates this merger for closed syllables.

Assamese		Aiton		Phake		Gloss
আদা	ada	ঞ্চ€	khiŋ[1]	ঞ্চ€	khiŋ[6]	'ginger'
চাং	saŋ	ঞ্চ€	khiŋ[1]	ঞ্চ€	kheŋ[1]	'shelf'

Audio Link to pronunciation of these words in Aiton (\wave files\aiton\sounds\khing.mp3)
Audio Link to pronunciation of these words in Phake (\wave files\phake\sounds \khing.mp3)

Table 48: Merger of /i/ and /e/ in Aiton

The raising of /e/ > /i/ may have been happening for a long period of time. Table (49), based on Grierson (1904:214-233), gives transcriptions of some words with front vowels from his comparative word list. It appears to show the raising of /e/ to /i/.

Grierson's Aiton	Phake	Aiton (this thesis)	Gloss
Lik	lek[1]	lik[1]	iron
Lin	lin[4]	lin[3]	tongue
Kin	kin[2]	kin[2]	eat
Liñ (len)	lɛn[5]	lɛn[2]	run
Chich (chet)	cet[1]	cit[1]	seven
Pit (pet)	pɛt[1]	pɛt[1]	eight
Pin	pen[2]	pin[2]	be

Table 49: Proto */i/, */e/ and */ɛ/as recorded by Grierson (1904)

We must be wary of Grierson's transcriptions, which seems to be based only on the written forms, as shown by the form for 'run', which would be written ငလ္ဢ၅, which Grierson notates as <liñ>. This does not suggest that the vowel /i/ is present, rather that the symbol ${}^{\circ}$, which can indicate [i], [e] or [ɛ] in closed syllables, was employed (for the notation of front vowels in Tai script, see below 7.2.3).

Banchob, on the other hand, was a very careful observer. As discussed above in 6.3.2.1.1, where she writes /e/ in Phake, Banchob usually writes /i/ for Aiton. Table

(50), based on Banchob (1977), which was originally written in Thai orthography, illustrates this:

Aiton (Banchob 1977)	Phake (Banchob 1987)	Putative proto form (after Li 1977)	Gloss
cit[1]	cēt[1]		'seven'
cip[1]	cēp[1]	*cepD2	'to be ill'
pīk[1]	pik[1]	*pi(i)kD2	'wing'
pit[1]	pɛt[1]	*petD2	'duck'
bit[1]	mit[1]	*bitD3	'to twist'
bit[1]	met[1]	*betD3	'fishhook'
cit[4]	cēt[4]	*chetD4	'to wipe'

Table 50: Proto */i/ and */e/ as recorded by Banchob (20/3/1977)

In general, *e in open syllables is realised as [i] in the data from Bidya Thoumoung, such as ɕī[2]'granary' and ɕī[3]'eldest daughter'. These two words are recorded by Banchob in the *Aiton–English–Thai Dictionary* as yī[4] and yī[9] respectively, showing that in open syllables the /*e/ > /i/ merge had occurred by the time that Banchob researched the Aiton language in the late 1960s and early 1970s.

Bidya Thoumoung also provided some examples in open syllables of the merge of /e/ to /ɛ/, listed here in Table (51). However these words do not merge with /ɛ/ in Banchob's *Aiton–English–Thai Dictionary*.

Assamese		Aiton - recorded by Bidya Thoumoung		Aiton - Banchob	Phake		Gloss
নকৰ	nɔgɔr	ငလ	cɛ[2]	-	ငလ	cɛ̄[5]	'city'
কাগছ	kagɔs	ငလ	cɛ[3]	cī[9]	ငလ	cɛ̄[3]	'paper'
তিয়াই থোৱা	diyai tʰʊwa	ငလ	cɛ[2]	cī[9]	ငလ	cɛ̄[5]	'to soak in water'

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\ce.mp3)

Table 51: Merger of */e/ and */ɛ/ in Bidya Thoumoung's list and comparison with Banchob's *Aiton–English–Thai Dictionary*.

One early example of the merge of /e/ > /i/ in open syllables is provided in the *Treaty between the Aitons and the Turungs*, No. (44), where the word ဪ *ʒii³* '1st daughter' is found. The manuscript was written about 1950 and although it is now somewhat damaged and the consonant can no longer be fully made out, the vowel ᵒ [ii] can still clearly be seen on the top line of the second half of the page, to the right (see photograph of manuscript, \photos\books\Lik Kheu Meung\, p2). It is fortunate that Nabin Shyam Phalung made a transcription of this manuscript before it was damaged by worms. The Phake equivalent is ဪ *yē³*, showing clearly that this is an example of the merge of /e/ to /i/ in Aiton.

6.3.2.2.2 Central Vowels

As with the front vowels, it appears that in Aiton there is complete merger of the central vowels. The close back spread vowel symbol [ɯ] has been chosen to represent this, just as the close front [i] and close back [u] are being used.

Table (52) shows this merger:

Assamese		Aiton		Phake		Gloss
চোলা	sola	ဪ	suɯ³	ဪ	sə̄³	'shirt'
পোন	pon	ဪ	suɯ²	ဪ	sǖ⁵	'straight'
বাঘ	bagh	ဪ	suɯ¹	ဪ	sə̄⁶	'tiger'
কিনা	kina	ဪ	suɯ³	ဪ	sǖ⁴	'to buy'

Audio Link to pronunciation of these words in Aiton (\wave files\aiton\sounds\seu.mp3)
Audio Link to pronunciation of these words in Phake (\wave files\phake\sounds\seu.mp3)

Table 52: Merger of */ɯ/ and */ɤ/ in Aiton

The Aitons and the Phakes are well aware of the difference between their languages in this regard. In Phake, the words *mü²* 'hand' and *mə²* 'return home' are clearly different (see Table (13) in 6.2.2). On the other hand, Nabin Shyam Phalung clearly that in Aiton these two words are clearly the same, here notated as *mɯ²*.

6.3.2.2.3 Back Vowels

Whereas in the case of the front vowels, there has clearly been a merger of three phonemes into two, and in the case of the central vowels a clear merger of two phonemes into one, there remains some evidence for three back vowels in Aiton.

In general, as with the front and central vowels, the mid-high back vowel has merged to the high vowel, especially in closed syllables. This is exemplified in Table (53):

Assamese		Aiton		Phake		Gloss
পকঁৰ	pɔkʰɔr	ရ	kun³	ရ	kon³	'bottom'
মিছা মাছ	misa mas	ရ	kuŋ³	ရ	kuŋ³	'shrimp'

Audio Link to pronunciation of these words in Aiton (\wave files\aiton\sounds \kun-kung.mp3)

Table 53: Merger of */o/ and */u/ in Aiton

It appears that this merger of /o/ > /u/ in closed syllables has been occurring for many years. Grierson (1904) records the Aitons as pronouncing the word for 'person' as <Kun>, although, as discussed above (see 6.3.2.2.1) we cannot be sure that this is not simply following the written form.

In addition, Banchob (1977) recorded the following (Table 54)

Aiton (Banchob)	Phake (Banchob)	Gloss
kū²	ko²	to fear
kun²	kon²	person
hū²	hū²	hole
suŋ³	soŋ¹	to send
sut⁴	sot⁴	to sip
phun⁶	phon⁶	rain
tuk¹	tok¹	to fall

Table 54: Merger of */u/ and */o/ in Aiton, after Banchob (20/3/1977)

An examination of all the words beginning with the letter /k/ in Banchob's *Aiton–English–Thai Dictionary* shows that in closed syllables, some Aiton words with reflexes of proto *o are written with <u>, and some with <o>. This is presented in Table (55):

Aiton (Banchob Dictionary MS)[52]	Phake (Banchob 1987)	Gloss	Merge
kuŋ²	koŋ²	'bow'	*o > u
kuŋ⁹	kuŋ³	'prawn'	*u > u
kuŋ⁸, koŋ⁸	kuŋ¹	'hunchbacked'	*u > u, o
kuŋ², koŋ²	koŋ²	'hollow'	*o > u, o

Table 55: Merge of *u and *o in the *Aiton–English–Thai Dictionary* (Banchob ms).

Banchob only records six words beginning with /k/ that have <o> in a closed syllable. Two of these are onomatopoeics and the others are also all found with <u>, as in the examples for 'hunchbacked' and 'hollow' in Table (52). The fact that Banchob wrote <o> as the vowel of the word 'hunchbacked', when the vowel is in fact a reflex of *u, suggests that her Aiton informant freely mixed [o] and [u], but that they had been merged into a single phoneme.

As mentioned earlier, (see 6.3.2.1.1), it appears that in open syllables there may still be a distinction between /o/ and /ɔ/ in a small number of cases. Table (56) exemplifies this contrast.

Assamese		Aiton		Phake		Gloss
বন্ধু	bɔndhu	ကော	[kɔ³]	ကော	kɔ³	'friend'
উঠোৱা	utʰuwa	ကော	[ko¹]	ကော	kɔ¹	'heap'
জন	dʒɔn	ကော	[kɔ³], [ko³]	ကော	kɔ⁴	'CLF'

Audio Link to Pronunciation of these words in Aiton (\wave files\aiton\sounds\ko2.mp3)
Link to spectrogram image of repeated pronunciation of 'friend' and 'classifier' (\wave files\aiton\wave images\ko2.bmp)

Table 56: [ɔ] and [o] in open syllables

---

[52] Banchob notates ten different tones in the *Aiton–English–Thai Dictionary*. Her numbering is maintained in this table. Her findings on tone are discussed at 6.3.4.1.1.

In the audio for Table (56), Bidya Thoumoung repeats the words for 'friend' and 'classifier'. The repetition makes it clear that that the pronunciation of the word *kɔ³* 'Classifier for humans' varies between [ko³] and [kɔ³].

In the texts, there is evidence of the vowel in *kɔ³* 'CLF' being pronounced with a higher vowel than /ɔ/ in a word like *kɔ³* 'LINK', just as was observed by Banchob (see 6.3.2.1.1). This can be seen by comparing the vowels of *kɔ³* in (30.1) and (30.2), both of which are from the same text and the same speaker.

30.1) ကော ဝီ ကော ရှိ ၃ ။
    lɛ¹ khaa¹ kɔ³ uŋ¹ dai³
    roam search LINK NEG get
    'She went to look for it but could not get it.'
    (Aiton Story, *The story of the forest spirit and the first daughter*, No (7), told by Sa Cham Thoumoung)

30.2) ရှိ ဝီ ဝင်း ကော ရှိ ဝင်း ဝီ ။
    huŋ¹ maa² luk³ kɔ³ luŋ² pin² maa²
    long time come child CLF one be come
    'After a long time a child was born.'
    (Aiton Story, *The story of the forest spirit and the first daughter*, No (29), told by Sa Cham Thoumoung)
    Link to Spectrogram comparing *kɔ³* 'LINK' and *kɔ³* 'CLF' (\wave files\aiton\wave images\ko3.bmp)

Laura Tollfree (pers. comm.) pointed out that to her ear the words for 'friend' and 'CLF for humans' have a different vowel, and she felt that rounding was a key difference between the two. This may suggest that /kɔ³/ 'human classifier' is weakening to /kɔ³/ in Aiton.

The use of [o] as an allophone of phoneme /ɔ/ is also found in manuscripts, where the orthography for /o/ is used for words with /ɔ/, as shown in example (31):

31) ကန်တော
    kan¹ tɔ²
    pardon
    (Aiton Manuscript, *Calendar for the year 2001-2002*, No. (35))

This use of လၘၖ for ꨁꨯꨮ seems to show that /o/ is no longer a separate phoneme. This spelling would never be countenanced by the Phakes. Further evidence from the same manuscript has the word head, *huu¹* in Aiton, spelled as ꨁꨯ, which would be the spelling in Phake (*ho⁴*) and in older manuscripts. The use of the same orthography for both /u/ and /o/ strengthens the case that the old phoneme /o/ is no longer contrastive.

All of the words which have [o] in Aiton are reflexes of words whose vowel is *ɔ, such as *kɔ³*'CLF' or *tɔ²*'to weave'. The possibility of a separate phoneme /o/ does not change the fact that proto *o and *u appear to have merged as /u/ in Aiton.

We can summarise this as (32):

32) Phake   Aiton
     /u/   >   /u/
     /o/   >   /u/
     /ɔ/   >   /ɔ/ and perhaps /o/

6.3.2.2.4 Vowel length distinction.

In general vowels in open syllables are long, and those in closed syllables are shorter. Vowel length is not distinctive, except in the case of /a/~/aa/ where length contrast is present in closed syllables.

It may be that the vowel length distinction /a/ versus /aa/ is being lost in dead (stop final) syllables. Table (57) gives the measurements in seconds for the vowels of several words with /a/ or /aa/ as the nucleus. It is interesting that the short vowels followed by stops are longer than the short vowels followed by nasals in these examples.

Aiton			Length of Vowel in Aiton	Phake		
ꨓꨮꨬꨮ	taan¹	'like'	0.355"	ꨓꨮꨯ	tān¹	'as valuable as'
ꨓꨮꨬꨮ	taan³	'to say'	0.275"	ꨓꨮꨯ	tān³	'to speak'
ꨓꨮꨬꨮ	taan²	'cancer'	0.46"	-		
ꨓꨮꨬꨮ	tan²	'to bar'	0.11"	ꨓꨮꨯ	tan²	'to close, to bar'
ꨓꨯꨮ	tap¹	'liver'	0.28"	ꨓꨯꨮ	tap¹	'the liver'
ꨓꨯꨮ	tap³	'need'	0.225"	ꨓꨯꨮ	tap⁴	'need not, must not'
ꨓꨯꨮ	tap³	'a town'	0.28"	ꨓꨯꨮ	tap⁴ sük¹	'military camps in the battlefield'

Audio Link to Pronunciation of these words in Aiton (\wave files\aiton\sounds \tan-tap.mp3)

Table 57: Contrastive vowel length in Aiton

As can be seen, there is a very clear difference in the length of the vowels which have nasal finals. The word *tan²*'to bar' is phonemically short and is clearly much shorter than the phonemically long words. In the case of *taan³*'to say', the vowel length is shorter than for *taan²*'cancer'. This may reflect the fact that the 3rd tone was originally glottalised and therefore shorter than the other tones.

In Table (57), all three words with final stops have nuclei which consist of phonemically short vowels, yet they are all somewhat longer the vowel in *tan²* 'to bar'. It may be that the length contrast is being reduced for dead syllables, and that if no contrast is present (no word *taap* was recording in Aiton), length is no longer important. Where there is a phonemic contrast, contrastive length remains, as in Table (58):

Aiton			Length of Vowel	Phake		
ꨀꨯꨮ	kaap³	'to bite'	0.17'	ꨀꨯꨮ	kāp⁴	'to bite'
ꨀꨯꨮ	kap³	'tight'	0.09'	ꨀꨯꨮ	kap⁴	'tight'

Audio Link to Pronunciation of these words in Aiton (\wave files\aiton\sounds\kap.mp3)

Table 58: Contrastive vowel length in dead syllables in Aiton

It appears that in words with final stops, the length of /aa/ is shorter in words with the 3rd tone (low falling), than those with the 1st tone (high level). A neutralisation of length contrast might then be expected to show up first in words with the 3rd tone.

6.3.2.2.5 Diphthongs

Table (59) lists the diphthongs identified in Aiton:

Finals Initials	i	ɯ	u
i	–	–	iu
ɛ	–	–	ɛu
ʊ	ui	–.	–
ɔ	ɔi	–	–
a	ai, aai	aɯ	au, aau

Table 59: Diphthongs in Aiton

The decision to categorise these as diphthongs has already been discussed in 6.1.2 above, and the length distinction between /ai/ and /aai/ exemplified in Table (7) above. The diphthong /ɔi/ is exemplified in Table (60):

Assamese		Gloss	Tai Aiton
শামুক	ħamuk	shellfish	hɔi¹
চিন	sin	symbol	hɔi²
মালা ওঠা	mala gutʰa	to hang	hɔi³

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\hoi.mp3)

Table 60: The diphthong /ɔi/ in Aiton

The final /-i/ is longer in both hɔi¹ 'shellfish' and hɔi² 'symbol', than it is in hɔi³ 'to hang'. This is due to the fact that the third tone is shorter than the others, because it was originally a glottal tone.

Table (61) exemplifies the dipthong /aɯ/.

Assamese		Gloss	Tai Aiton
চফা	sɔpʰa	clear	saɯ¹
ভবোৱা	bʰɔrɔwa	put in	saɯ¹

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\saeu.mp3)

Table 61: The diphthong /aɯ/ in Aiton

6.3.2.2.6 Vowel reduction

Apart from the vowels listed in Table (43) above, there is vowel reduction, to something approaching schwa, in some unstressed syllables. As discussed above in 6.2.2.3, this is interpreted in Phake as being either the vowel /ü/ [ɯ] or /ə/ [ɤ]. In Aiton, however, these two phonemes are merged as single phoneme (see above 6.3.2.2.2.

The reduced vowel of haŋ² 'to', as pronounced by Sa Cham Thoumoung and the back unrounded vowel of huŋ² 'glare', as pronounced by Bidya Thoumoung are not the same. The spectrogram for huŋ² 'glare' shows a second formant of around 1500Hz, suggesting a high back vowel, whereas the spectrogram for haŋ² 'to', although rather weak, shows a second formant around 1000Hz. The sound files for these two words can be found at \wave files\aiton\sounds\ as hang.mp3 and heung.mp3, and the spectrograms at \wave files\aiton\wave images.

6.3.3 Phonotactics

6.3.3.1 Initial consonant clusters

Banchob's *Aiton–English–Thai Dictionary* records the following initial clusters for Aiton, mainly in onomatopoeic words:

33)   br   pr   phr   dr   cr   sr   kr   khr   gr
           pl   kl

Several of these clusters appear to be two syllable words where the first syllable is unstressed and probably pronounced with a unstressed schwa. For example Banchob

records sra⁶'Buddhist monk', which is written သျၢၺ and transcribed as sa¹ ra¹ in this thesis.

Diller (1992:15) mentions only four clusters, namely:

34)   pr   phr   kr   khr

From the word list of Bidya Thoumoung, the clusters listed in (35) were recorded.

35)   pr   phr   tr   khr
             pl
             py   ky

Of these, several such as /tr/, /py/ and /ky/ are only found with words that are obviously borrowed from Pali or Burmese.

Initial clusters are exemplified by the word *khraa³*'die' (see \wave files\aiton\sounds \khra.mp3).

6.3.3.2 Epenthetic syllables:

In the names of the days Sunday and Monday, a syllable *ka¹* is inserted in pronunciation, as in example (36).

36)   ၁.   စ႑ၼၵ္ကၢၨဎၵ္ ဂ္ဂေၵ႑   ॥
         (luŋ²)   ta¹ naŋ¹ ka¹ nui¹ kam³ kɔ¹
         one   Sunday   *Mesua pedunculata**
         'First: on Sunday, *Mesua pedunculata* flowers.'
         (Aiton Manuscript, *Calendar for the year 2001-2002*, No. (41), read by Sui Khong Thoumoung)

The Shan spellings of the word for Sunday is ဝၼ်းဢၢတိတ်ꧣꧤꩪꩫ, (ta naaŋ: noi,) and Banchob wrote *ta⁴ nãŋ¹ noi¹* for Phake, both of which lack the epenthetic *ka¹*. This *ka¹* probably arises as a result of a velar nasal being followed by /n/ in the next syllable.

6.3.4 Tone

6.3.4.1 Previous Investigations of Aiton

6.3.4.1.1 Banchob

6.3.4.1.1.1 Banchob (24/4/1977)

Banchob (1977) lists seven tones which she identified for Aiton. Using the methodology of Gedney (see 6.1.5.3), Table (62) has been derived from the words listed by Banchob (24/4/1977):

	A	B	C	DL	DS
1	6	3 (also 1,5)	7	1	
2	2				
3					
4		7 but very little data	4		

Description of Tones:		
1.	high tone	2. high tone with final falling
3.	low tone with final rising	4. falling tone
5.	low tone	6. rising tone
7.	falling tone with a high tone at the beginning	

Table 62: Tone Box for Aiton, derived from Banchob (24/4/1977)

It is clear from the above description that

A1   is a rising tone, and not merged with B1-3

A2-4   is a high falling tone, as it is today

D1-3   is a high level tone, as it is today

C4 & D4   is a falling tone, probably lower than A2-4, as it is today.

Unfortunately Dr. Banchob's data does not give a clear enough description of the B tone or C1-4 for a clear picture to emerge.

6.3.4.1.1.2 Banchob (Aiton–English–Thai Dictionary)

In the *Aiton–English–Thai Dictionary*, Banchob notates ten tones, but mo t of them are very infrequently used. Only tones 1,2,8,9 & 10 are used widely throughout the manuscript. It appears that Tones 1 and 8 refer to the same toneme, because words like kai^8 'chicken' and $k\bar{i}^1$ 'how many' belong to the same tone box (B1) and have the same tone in other Tai varieties. Tone number 10 is usually reserved for dead or checked syllables. Table (63) compares Banchob's tones 1, 8 and 10 with Phake:

Aiton - Banchob (*Aiton–English–Thai Dictionary*)	Phake - Banchob (1987)	Gloss
$k\bar{i}^1$	$k\bar{i}^1$	how many
ko^8	ko^1	heap
$k\bar{a}\eta^1$	$k\bar{a}\eta^1$	crossbow
$k\ddot{u}k^{10}$	$kək^1$	to try
$kɔk^{10}$	$kauk^1$	a glass
$kɔŋ^8$	$kauŋ^1$	curve
$kɔt^8$	$kɔt^1$	to embrace

Table 63: Comparison of Banchob's tones 1,8 & 10 from the *Aiton–English–Thai Dictionary*.

If we assume that Banchob's tones 1, 8 & 10 correspond with the 1st tone as notated in this thesis, that Banchob's tone 2 corresponds with the 2nd tone as notated in this thesis, and Banchob's 9th tone corresponds with the 3rd tone in this thesis, then with very few exceptions, a three tone system can be posited from Banchob's data.

6.3.4.1.1.3 Banchob (Tape recordings)

Old reel-to-reel tape recordings made by Dr. Banchob in Bargaon in about 1970 have come to light. Her methodology for finding out the tones was to read a series of words from each of the Standard Thai tones and get equivalents in Aiton. She pronounced the word in Thai first and then Ai Ne Phalung pronounced the Aiton equivalent. Because the B4 box is merged with C1-3 in Standard Thai, there are no examples of B4 in her list.

Table (64) lists ten words collected by Banchob:

Aiton script	Aiton phonemicisation (after *Aiton–English–Tai Dictionary*)	Aiton phonemicisation (following this thesis)	Tone Box	Gloss
ဤၵၖ	$pa^3\ c\bar{a}i^2$	$pa^1\ caai^2$	A2	'male'
ၵၖ	kin^2	kin^2	A2	'eat'
ၵၜၟ	cit^8	cit^1	D(S)2	'seven'
ၵ	kai^8	kai^1	B2	'chicken'
ၵၟ	$b\bar{a}n^9$	$baan^3$	C3	'village'
ၵ	$khai^9$	$khai^3$	C1	'fever'
ၵ	ma^4	maa^3	C4	'horse'
ၵ	hai^9	hai^3	C4	'dry ricefield'
ဤၵ	$pa^3\ s\bar{a}u^7$	$pa^1\ saau^1$	A1	'female'
ၵ	$sɛŋ^7$	$sɛŋ^1$	A1	'diamond'

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\ban-tone.mp3)

Table 64: Aiton Tones from the Tapes of Dr Banchob Bandhumedha.

Although Banchob notates the two A1 tones as 7th tone, and the two B2 tones as 8th tone, both are rising tones and it is hard to see why they are differently notated. Banchob uses the 9th tone to notate both the word for 'dry ricefield' and for 'fever', suggesting a C1-3 and C4 merge.

There is considerable variation in the way that individual Aiton informants pronounce the tones in citation, and this may explain the large number of tones notated by Banchob.

6.3.4.1.2 Diller (1992)

Diller (1992:18) reported two different systems, in Tables (65) and (66):

Aiton₂:

	A	B	C	DL	DS
1	4	6	5	6	
2	3				
3					
4		1	2	1	

Description of Tones:		
1. low	2. mid falling	3. high falling
4. low rising	5. mid falling glottalised	6. mid rising

Table 65: Aiton₂ Tone Box after Diller (1992)

This was described by Diller (1992:19) as "undoubtedly more conservative." Considerable effort during the research for this thesis failed to find any speaker who produces a tone box of this type.

A second tone box, for which Nabin Shyam Phalung was the informant, is given in Table (66). It closely relates to the current situation in the Aiton villages:

Aiton₁:

	A	B	C	DL	DS
1	1		2	1	
2	3				
3					
4				3	

Description of Tones:		
1. rising	2. mid falling	3. high falling

Table 66: Aiton Tone Box after Diller (1992)

6.3.4.1.3 Current situation

In preliminary discussions with Bidya Thoumoung and other Aiton informants in November 1999, five tones were recognised. These were described and named as indicated in Table (67):

				examples:		
Tone 1	ဘ၆ငၥ၆	siŋ¹ suɯ²	straightforward tone	ၾ� (ၾ)	baa¹	'shoulder'
Tone 2	ဘ၆ဆၟၵ	siŋ¹ luŋ²	falling tone	ၾၡ	maa²	'to come'
Tone 3	ဘ၆ဆၞ	siŋ¹ sau²	resting tone	ၾၡ	maa³	'horse'
Tone 4	ဘ၆ၶ၆ၥ	siŋ¹ khuɯn³	rising tone	ၾၡ	maa⁴	'dog'
Tone 5	ဘ၆ဝ၆ၵၥ	siŋ¹ tɯk³	stopped tone	ၾၡ (ၾ)	baa⁵	'mad'

Table 67: Tones in Aiton

Despite the fact that it appeared that there were 5 tonemes, on the basis of the above list, while working through the word list, it became clear that Bidya Thoumoung assigned tone 5 to dead or checked syllables, and tone 3 to syllables with final nasals or final vowels, and that therefore Tones 3 & 5 were probably allotones of the same toneme.

Tone 4 was rarely assigned, and in particular was not assigned to the word for 'dog' when that word was encountered in the word list. It therefore appeared that tone 4 was a sporadic allotone of Tone 1.

The following tone box, then, is proposed for contemporary Aiton (Table 68)

	A	B	C	DL	DS
1	1 (&4)		3 (&5)	1	
2	2				
3					
4					

Description of Tones:		
1. mid/high level	2. high level then falling	3. mid falling
4. mid rising	5. mid falling glottalised	

Table 68: Aiton Tone Box

Examples of the three tones are given in table (69):

Assamese		Aiton		Gloss
মানুহ	manuh	ကွၟ်ꩻ	kun²	'person'
পকৰ	pakɔr	ကွၟ်ꩻ	kun³	'bottom'
উভলা	ubhɔla	ကွၟ်ꩻ	kun¹	'to uproot'

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\kun-tone.mp3)

Table 69: The Aiton tones

A very clear example of the tone merger of A1 and B1-4 is indicated in Table (70):

Assamese		Aiton		Phake	Gloss
আদা	ada	ၶိင်	khiŋ¹	khiŋ⁶	'ginger'
চাং	saŋ	ၶိင်	khiŋ¹	kheŋ¹	'shelf'

Audio Link to pronunciation of these words in Aiton (\wave files\aiton\sounds\khing.mp3)
Audio Link to pronunciation of these words in Phake (\wave files\phake\sounds \khing.mp3)

Table 70: Merger of tones in Aiton

The relationship between the tones of Aiton and those of Phake is shown in (37):

37) Phake tones   Aiton tones
    Tones 1 & 6   Tone 1
    Tones 2 & 5   Tone 2
    Tones 3 & 4   Tone 3

6.3.4.2 Some examples of tone variations:

6.3.4.2.1 Variation between tones 1 & 4

In mid 2000, Bidya Thoumoung was asked by letter to assign tones to a series of words, after being sent a copy of Table (67) above. He sometimes assigned the rising tone, Tone 4, to words which have the level tone in Phake, and the level tone, Tone 1, to words which have the rising tone in Phake. This suggests that the rising tone (4) is an allotone of the level tone (1), or vice-versa.

Some of Bidya Thoumoung's tone assignments are listed in Table (71), in which, for ease of comparison, Shan forms are phonemicised as if they were Phake, where tone 6 refers to the rising tone. In the Aiton columns, tone 4 refers to the rising tone. Tone 1 refers to the level tone throughout.

Aiton (tones as assigned by Bidya Thoumoung)	English	Phake	Shan	What the Aiton tone would be following Phake & Shan
lung¹	'big'	loŋ⁶	loŋ⁶	luŋ⁴
tang⁴	'all'	tüŋ¹		taŋ¹
kaa¹ thaa¹	'Buddhist prayer'		kaa¹ thaa¹	kaa¹ thaa¹
kho⁴	'separately'	khɔ⁶	khɔ⁶	khɔ⁴
kheun² kon⁴	'also'		khün² kɔn¹	kɔn¹

Table 71: Variation between the Aiton level tone and rising tone.

6.3.4.2.2 Possible survival of separate B4 tone

The only evidence of a possible separate B4 tone was in ကွၚ်ၚ် [kuu² muɯ⁶] 'every day' (see \wave files\aiton\sounds\tone6.mp3) where the number 6 has been arbitrarily given to refer to a low tone. The word *muɯ²* 'time' is usually pronounced with the high level falling tone.

6.3.4.3 Short tones in Aiton

There were a small number of words which Bidya Thoumoung pronounced with very short high tones. These words were always open syllables (vowel final). One of these is given in Table (72), and is compared with a vowel final word with a normal long tone.

Aiton		Aiton		Phonetic form	Gloss
তিয়াই থোরা	diyai tʰuwa	ၸၢဲ	cɛ²	[tɕêe]	'to soak in water'
ঘিন কৰা	gʰin kɔra	ၸၢဲ	cɛ¹	[tɕé?]	'to hate, to dislike'

Audio Link to Pronunciation of these words (\wave files\aiton\sounds\ce-short.mp3)

Table 72: Short tones in Aiton

Another example pronounced by Bidya Thoumoung with a short tone was *taa³*'wipe' (hona in Assamese) (\wave files\aiton\sounds\ta.mp3). This word is not always pronounced with a short tone, as can be which can be heard in example (38), spoken by Nabin Shyam.

38) ဟ႑ ဟ႑ ၌ ႑ှၞ် ႑ၭၫ် ႑ၣ် ၫ႒ၫ်

ta¹ taa³ haw³ khon¹ an² lum² nan³
will wipe GIVE khon CLF one that
'Will wipe (it) for that one *khon*.'

(Aiton text, *Explanation of Hong Khon ceremony*, No. (57), told by Nabin Shyam.)

Words with short tones do not occur frequently in the texts, and when they do occur it seems to be younger speakers who use them. One example is (39), where there is a short tone on *paa³*'basket':

39) ႑ ၯ႒ၫ် ႑ၞၣ် ၃ ႑ ၯ႑်   ॥

pa³ man² khɔt¹ pai² kaa¹ yau³
basket her leave behind go GO FINISHED
'She left her basket behind.'

(Aiton Story, *Story of the monkey and the fox*, No. (5), told by Bidya Thoumoung).

Further research needs to be done into short tones in Aiton.

6.3.4.4 Tone assignment in loan words

Most loan words in the texts presented in this thesis are polysyllabic Pali loan words. In general, the 1st tone is used for these words, although the third tone can be found on the last syllable. An example of Pali loan words is (40):

40) ႑ၯၣၵၣ႑ ... ၯၟးၐၐၯ႒ဏ ႑ၣ် ၯ႒ဏၯ႑် ႑ၫ် ၃၂

u¹ taŋ¹ ku¹ ru¹ ... pu³ pa³ wi¹ tɛ¹ haa¹ kan² wan² tuk¹ an² nai³
Utangkuru ... Pubbavideha direction west CLF* this

ၯ၌ ႑ ၯ႒ဏၐၣ ॥

khau¹ paa³ wan² puṃ²
they side east**

'And there is Pubbavideha in the west whilst they (= Utangkuru) are in the east.'

(Aiton Text, *Why Buddha was born in this world*, 2-4-3-1, No. (13), told by Sa Cham Thoumoung)

The assignment of tones to Pali loan words appears to be largely arbitrary.

There are another group of loan words, this time from Burmese, which have been used by the Tais for many generations. The rules of tone assignment for these words has not been investigated as part of this study.

6.3.5 The place of Aiton in Southwestern Tai

Throughout the discussion on the Aiton tonal system, there has been an underlying assumption that the Aiton tones are a merged system. With only three tonemes, Aiton has fewer contrastive tones than almost any other language in the Tai family. It is certainly plausible that a tonal system like that of the Phake could have merged into what we find in Aiton now.

Yet that is not necessarily so. In the light of the Aiton tonal system, it is worthwhile to observe that the reconstruction of proto Southwestern Tai tones posits three tones, as discussed above in 6.1.5.3. A similar observation can be made with the vowels, where the number and distribution of vowels in Aiton agrees with the reconstruction for proto Southwestern Tai in Jonnson (see above 6.1.1.3).

Indeed it is only the consonant inventory of Aiton that is markedly different from the posited proto Language. Since no Tai language has a consonant inventory anywhere near as large as that proposed in 6.1.1.1, it may well be that Aiton, with its preservation of certain consonant distinctions, its preservation of certain clusters, its seven vowel system and three tones, is closer to proto Southwestern Tai than the other daughter languages in Assam.

6.4 Some preliminary observations on Khamyang phonology

As discussed above in 2.3.4.2, Khamyang is now only spoken in the single village of Pawaimukh, in Margherita subdivision of Tinsukia district. Over several visits to Pawaimukh, I have made a number of recordings in Pawaimukh, mostly of Chaw Sa Myat Chowlik.

However, according to Deben Chowlik, Chaw Sa Myat Chowlik mixes Khamyang speech with Phake. There are several reasons for this. The first is that the Khamyangs are used to converging to other dialect norms because they are few in number. Secondly, since I am much more familiar with Phake than Khamyang, the Khamyang speakers are more likely to converge to Phake forms when recordings are being made. Furthermore, at least some of the Khamyang speakers feel that their own variety is more difficult to understand that the other varieties. In order to assist speakers from other varieties, they tend to converge to them.

For example, Chaw Sa Myat spoke line (41.1) with Phake phonology and then repeated it as (41.2) with Khamyang phonology:

41.1) ၮၖ္ ၮၖ္ ၑ ၮ္ၕ ၮ ၮ ၮ
cau^3 khau^6 yɛ^4 thaɯ^6 ta^1 kaa^1 nai^3
RESP 3Pl PRT where WILL go get
'His father said: 'Where will you get (them)"

41.2) ၮ္ၕ ၮ ၮ ၮ
thaɯ^6 ti^1 kaa^1 iai^3
where WILL go get
'Where will you get it?'
(Khamyang Story, *Of children and kings*, 3-1-1-3, Nos. (8) & (9), told by Sa Myat Chowlik.)

The differences between the two relate to the tones and the use of initial /l/ in (41.2), for the reflex of proto Southwestern Tai *d, which is always /n/ in Phake. Furthermore, in the word *thaɯ^6* 'where', there is a full realisation of [ɯ], the second element of the

diphthong, in Phake, whereas [ɯ] is hardly audible in the Khamyang pronunciation. This is further discussed below in 6.4.2.1.

Chaw Deben Chowlik, otherwise known as Chaw Kyun, was born around 1940 and never learned to read Tai, although he is fully proficient in reading Assamese and also knows English. His spoken forms are quite distinct from Phake forms, and he believes that they represent a more 'authentic' Khamyang. Most of the phonological information, therefore, is based on the speech of Chaw Deben Chowlik, rather than Chaw Sa Myat Chowlik.

6.4.1 Consonants

Based on several interviews with Chaw Deben Chowlik, Table (73) has been drawn up showing the initial consonants of Tai Khamyang:

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiceless unaspirated stops	p	t	c	k	(ʔ)
Voiceless aspirated stops	ph	th		kh	
Voiced stop	b				
Nasals	m	n	ñ	ŋ	
Voiceless fricative		s			h
Semi vowel			y		
Lateral Approximant	l				

Table 73: Consonant Phonemes in Khamyang

The distribution of /b/ and /l/ differs from that in other Tai varieties, as shown in Table (74):

English	Khamyang	Phake Banchob 1987	Aiton	Khamti Harris 1976
good	lii^5	nī^2	dii^2	ni^1
star	laau^5	nāu^2	daau^2	naaw^1
finger	liu^4	ñiu^4	niu^3	liw^2
cold	laau^6	nāu^6	naau^1	naaw^4
speak	baa^5	wā^5	waa^2	

comb	bii⁶	wī⁶	wii¹	wi⁴
fan	bii²	wī²	wii²	wi³
shoulder	maa¹	mā¹	baa¹	
village	maan³	mān³	baan³	maan⁵

Table 74: Comparison of Khamyang initial /b/, /l/ and /m/ with other Tai varieties

This use of initial [b], rather than [w] is apparently found for all words with reflexes of initial *w and *hw in proto Southwestern Tai. It is interesting to note that Grierson (see 3.2.6.6) observed that in Nora, "the letter ၀ is sometimes *ba*, but more usually *wa*." (1904:179). Chaw Deben Chowlik always pronounces this phoneme as [b], such as in example (42), but other Khamyangs do not.

42) ကၟ်   ဧၟ   ꩦဵꩧ်   ꩦꩨ   ၈   ꩩꩬ်   ꧡ   ၈
     kau⁵ baa⁵ uŋ¹   cau⁵ bii⁶ ho⁶ ɛ¹ bii⁶
     1Sg   say   NEG   be   comb head PRT comb
     'Didn't I say *bii⁶ ho⁶*, a comb?'

(Khamyang Sentence spoken by Deben Chowlik.)

Audio link for this sentence example (\wave files\khamyang\extracts\3-10-2-6 \44.mp3)

Sometimes this realisation of /b/ is somewhat fricativised, as in (43):

43) ꩤꩨ   ꩩ   ꧤ   ꩨ   ꩤꩨ   ꩦꩨဵ
     khaam² boi¹ phraa² tuu¹ tuu¹ khaam² nuŋ⁵
     word   bless Buddha same   word   one

     ꩩꩬ်   ꩴꩬ်   ꩤ   ꩤꩨ   ꩤꩨ   ꩩ   ꧤ ||
     lɔt¹   an²   khau⁶ yaau² khaam² boi¹ phraa²
     short CLF   3Pl   long   word   bless Buddha
     'The words of blessing the Buddha are the same, although ours is short and theirs is long.'

(Khamyang Sentence spoken by Deben Chowlik.)

Audio link for this sentence example (\wave files\khamyang\extracts\3-10-2-6 \56.mp3)

The pronunciation of words which are the reflexes of those that had initial *d in proto Southwestern Tai is a much more widespread Khamyang feature. Most of the Khamyang speakers pronounce such words with initial /l/ at least some of the time.

For example, Chaw Sa Myat pronounces the word for little finger as *liu⁴ kɔi³* (from proto *niuC4) (see \wave files\khamyang\sounds\liu.mp3).

Some of these words which are initial /l/ are not found with initial /l/ in any other Tai varieties in India. Examples of this include *laau⁶* 'cold', which has initial /n/ in Aiton, Phake, Khamti and in the reconstruction of proto Southwestern Tai. Moreover, I have not been able to identify any words with initial /n/ in Phake that are never pronounced with initial /l/ in Khamyang.

It is not clear why the Khamyangs pronounce so many of these words with initial /l/. It is a distinguishing feature of Khamyang to use of initial /l/ for words which are reflexes of those with proto initial *d, and which are pronounced with initial /n/ in Phake. It may be that in the terminal stages of the Khamyang language, hypercorrection to /l/ is occurring.

Both the pronunciation of initial /b/ and the pronunciation of initial /l/, as discussed in the previous paragraph, alternate with the Phake norms of /w/ and /n/ respectively. Chaw Sa Myat Chowlik almost always uses /w/ and /n/, following the Phake pattern, even when he is using Khamyang tones.

In example (4) in section 6.2.1 above, it was pointed out that certain function morphemes in Phake with initial /k/ show consonant initial lenition. A similar process is observed in Khamyang. The initial consonant of the morpheme *kaa¹* 'GO', marking past tense, is lenited to [ʔ] or zero. This can be seen in example (44), which also shows lenition to [ɣ] on the initial /k/ of *nan⁴ kɔ¹* 'moreover'.

44) ꩫꩬ်ꧠ   ꩩꧠ   ꩬꩨꩬ်   ꩬꩬ်   ꩬ   ꩨ
     nan⁴ kɔ¹   lum²   caŋ¹ nai³   luk⁴   nɛ⁴   khau⁶
     moreover   forget now   child   DEF   3Sg

 လွတ်လွှာ့ ၀ှ် ကှိင် တၢၵ် ၇ွ ကှ ॥

luk⁴ ɔn¹   khau⁶ uɯŋ¹ taan³ haai⁶ aa⁴
child 3Pl   NEG speak disappear PRT.GO
'Moreover, they've forgotten, those children now, those children can't speak, its all gone.'

(Khamyang text, *The future of the Khamyang language*, No. (20), spoken by Deben Chowlik.

6.4.1.1 Consonant clusters.

Although it has not been possible to make a comprehensive study of initial consonant clusters in Khamyang, no words of proto Southwestern Tai origin have been recorded with initial consonant clusters. There are some words with clusters having /r/ as the second member, such as *phrā²*'Buddha in example (43) above.

6.4.2 Vowels

The ten vowel contrast reported by Diller (1992:17) for the Tai varieties appears to be maintained in Khamyang. The vowels are listed in Table (71):

i		ɯ	u
e		ɤ	o
ɛ			ɔ
	a, aa		

Table 75: Vowel phonemes in Khamyang

There is a clear contrast in height between the close vowels /i,ɯ,u/ and the corresponding mid close vowels /e,ɤ,o/, as demonstrated for the Central Vowels in Table (76):

Assamese		Khamyang example	English gloss
চোলা	sula	sɤ³	'shirt'
বাঘ	bagʰ	sɤ⁶	'tiger'
কিন	kina	suɯ⁴	'buy'
বিসনা পাৰা	bisɔna para	sɤ¹	'cover, spread'
পোনে পোনে	pune pune	suɯ⁵	'straight'

Audio link to Pronunciation of these words (\wave files\khamyang\sounds\seu.mp3)

Table 76: Central vowels in Khamyang, pronounced by Deben Chowlik

Sa Myat Chowlik clearly distinguishes between /u/, in *huɯ⁶*'ear' and /o/ in *ho⁶*'head', as can be heard in the audio file \wave files\khamyang\sounds\hu-ho.mp3. However, this distinction is also present in Phake, and Sa Myat's speech might be influenced by Phake.

In the speech of Deben Chowlik, these distinctions also seem to be clear. In examples (45) and (46), the mid-close vowels /o/ and /e/ are heard in both in open syllables (*to⁵* 'body' and *me²*'wife') and in closed syllables (*taŋ¹ loŋ⁶*'all' and *cep¹*'hurt'). The realisation of /e/ is clearly mid-close, but the /o/ is less clear, especially *to⁵*in example (46). A clearer pronunciation of /o/ is found in example (42) above, on the word *ho⁶* 'head'.

45)

lɤ⁴ pɤ⁵ nai⁴   sii⁶   to⁵   taŋ¹ loŋ⁶   cep¹   na¹ huɯ⁶
therefore   PRT   body all   hurt   how

knaŋ⁵   naa⁶   ta¹   au⁵   pɔk⁴ ॥
direction PRT.QN WILL take return
'So we are distressed, and do not know how we will get (our traditions) back.
(Khamyang text, *The future of the Khamyang language*, No. (40), spoken by Deben Chowlik.

46)

haŋ² me² haŋ² luk⁴ taŋ¹ loŋ⁶ hauɯ³ taan³ khaaɯ² tai²
to   wife to   child all   GIVE   speak word   Tai

sɔn⁶   khaam² tai² ॥
teach word   Tai
'We should get out wives and children to learn and speak the Tai language.'
(Khamyang text, *The future of the Khamyang language*, No. (42)>. spoken by Deben Chowlik

6.4.2.1 Vowel reduction:

A number of vowel reduction processes with function words are apparent in the speech of Chaw Deben Chowlik. For example, the word (/luk⁴/ in Phake) 'from' is pronounced by him as [lu̯k⁴], and is distinguished from the word for 'child', (also /luk⁴/ in Phake), as can be heard in examples (47) and (48):

47) လုက် ၌လုင် မှ ၢၣ် ၢၢင် လဲ ...
    luk⁴   maau² loŋ⁶ maa²   co⁵   an⁵   lai⁴ ...
    from   Mau Lung come generation CLF so ...
    'From Mau Lung, generations ago ...

    (Khamyang text, *The future of the Khamyang language*, No. (26), spoken by Deben Chowlik

48) ၢၢင်ကော ၼ်ဳ ၢၢၣ်ဳ လုက် ၢ ၌
    nan⁴ kɔ¹ lum² caŋ¹ nai³ luk⁴   nɛ⁴ khau⁶
    there   forget now   child   DEF 3Sg

    လုက်ၢၢင် ၌ ၢ၌င် ၢၢင် ၢ ၢၢ ॥
    luk⁴ ɔn¹   khau⁶ uɯ¹ taan³ haai⁶   aa⁴
    child   3Pl   NEG speak disappear PRT.GO
    'Moreover, they've forgotten, those children now, those children can't speak, its all gone.'

    (Khamyang text, *The future of the Khamyang language*, No. (20), spoken by Deben Chowlik.

When *luk⁴* 'from' in example (47) and the two tokens of *luk⁴* 'child' in example (48) are compared closely, it is found that there are three different vowels, as can be heard on the audio file (\wave files\khamyang\sounds\luk.mp3). The vowel in the second token of *luk⁴* 'child' is clearly lower. This might be attributed to assimilation to the following mid-open vowel. Phonetically the second *luk⁴* is [lu̯ᴷɔn¹]. The spectrogram (\wave files\khamyang\wave images\luk.bmp) of these three vowels clearly shows the presence of second formant at around 1000Hz for the words for 'child', which is not present in the word for 'from'.

There is also a process of reduction of the second member of the diphthong /aɯ/. The latter was already commented above in the speech of Chaw Sa Myat (see above example (41)), and is one of the more noticeable differences between Phake and Khamyang.

The words *thaɯ⁶* 'where' and *haɯ³* 'give' occur quite frequently but in each case the final /ɯ/ is barely audible (see for example *haɯ³* in example (46) above). This could be notated phonetically as [aaˇ]. This is so striking that a casual observer might initially hear the vowels as simply [aa], and an [aa] pronunciation is accepted by the Khamyangs.

There are a number of other vowel reduction processes that have been observed, such as *lai⁴ 'so' being realised as [lɤ⁴] in example (45) above.

6.4.3 Tones

The Khamyang tonal system has six tones with a different distribution from the six tones of the Phake (see 6.2.4.4). The tone box in Table (77) is based on an interview and some words recorded in April 2001. The informant was Chaw Deben Chowlik[53]. This tone box differs from the Phake, where A2-3 is merged with A4, and B4 is a separate tone, It also differs from Khamti where A1 is merged with B1-3.

	A	B	C	DL	DS
1	6	1	3	1	
2	5				
3					
4	2	5	4	4	

---

[53] I am very grateful to Jerrold Edmondson for discussing these tones with me.

Tone	1	mid level with slight fall
	2	high level then falling
	3	low, falling and glottalised
	4	high level then falling with glottal ending (weaker glottalisation than tone 3)
	5	low falling
	6	level, with possible rise at the end

Table 77: Tones in Khamyang

The numbering of the Khamyang tones in this thesis follows the Phake numbering as closely as possible. Because of the different distribution of the tones, A2-3, which are tone 2 in Phake, are numbered as tone 5 in Khamyang. All other boxes have the same tone number as Phake, although the contour and quality of the tones differs.

Table (78) exemplifies five of these tones on the syllable *maa*, as spoken by Deben Chowlik.

Assamese		Gloss	Tai Khamyang
কান্ধ	kandhɔ	'shoulder'	maa[1]
কুকুৰ	kukur	'dog'	maa[6]
বলিয়া	bɔliya	'crazy'	maa[3]
অহা	ɔha	'come'	maa[2]
ঘোঁড়া	ghʊra	'horse'	maa[4]

Audio Link to pronunciation of these words (\wave files\khamyang\sounds\ma-tones.mp3)
Link to Image of pitch contour chart for these words (\wave files\khamyang\wave images
\ma-comp.bmp)

Table 78: Tones 1-4 and 6 pronounced on the syllable *maa* in Khamyang

The contour of the 2nd tone and 4th tone are almost exactly the same. Both are high level and falling, but the 4th tone has an abrupt glottal ending. The contour of these two tones differs significantly from the 5th tone, which is low falling. Table (79) compares the 1st, 4th and 5th tones on the syllable *kaa*.

Assamese		Gloss	Tai Khamyang
যাব	dʒabɔ	'go'	kaa[1]
কাউৰী	kauri	'crow'	kaa[5]
বেপাৰ	bɛpar	'business'	kaa[4]

Audio Link to Pronunciation of these words (\wave files\khamyang\sounds\ka-tones.mp3)
Link to Image of pitch contour chart for these words (\wave files\khamyang\wave images
\ka-comp.bmp)

Table 79: Tones 1, 4 and 5 on the syllable *kaa* in Khamyang

Unlike in the other varieties of Tai spoken in Assam, the C4 tone (tone 4) is a high falling tone, with only an abrupt glottal closure clearly distinguishing it from the A4 tone (tone 2). The contours of tones 2 and 4, however, are almost identical.

The difference between the C4 tone and the cognate tone in Phake is exemplified in (49), where Sa Myat Chowlik pronounces a sentence first in Phake, then rather more quietly in Khamyang phonology:

49)

khauɯ³ kin² nɤ⁴ khauɯ³ kin⁵ lɤ⁴
want eat meat want eat meat
'Want to eat meat.'
(Khamyang Story, *Of children and kings*, 3-1-1-3, No (65) told by Sa Myat Chowlik.)

The A1 tone (tone 6) and the B1-3 tone (tone 1) are similar. They are both level, and both at about the same pitch, but the key difference is the possible final rise in tone 6, which is not found in tone (1). This final rise is much less pronounced than in Tai Phake, and was another difference between examples (41.1) and (41.2) above. Clear examples of the rising tone in connected speech can be heard with the words *sii⁶* and *loŋ⁶* in example (36) above.

Where there is no contrast required, the 6th tone can even fall, as in example (42) above, where the word *bii⁶* at the end of the sentence clear falls. This can be seen in the pitch contour diagram (see \wave files\khamyang\wave images\bi.bmp).

The first tone can be either level, as with *ū¹* in (50.1) or falling, as with *ū¹* in (50.2):

50.1) ᧠ ᧠ ᧠ ᧠ same
    kii¹   ban²   ta¹   uu¹
    how many day   will   stay   same
    'How many days will you stay, it's the same (as in Phake).'

50.2) ᧠ ᧠ ᧠ ᧠
    kii¹   ban²   ta¹   uu¹
    how many day   will   stay
    'How many days will you stay.'
    (Khamyang sentences spoken by Deben Chowlik.
    Audio link for these sentence examples (\wave files\khamyang\extracts
    \3-10-2-6\79-80.mp3)

This has been only been a brief sketch of Khamyang phonology. Moreover it is largely based on the speech of a single informant, and it is not clear whether this sketch represents the Khamyang speech community adequately. Readers who wish to hear the Khamyang, as spoken by Deben Chowlik, are enjoined to examine the story of the *Hunter's Parrot*. The text contains some other features of Deben Chowlik's speech, such as word fusion, where two words are fused as one. This is demonstrated in (51):

51) mɤn⁴   ...   cau³ phaa⁴ kɔŋ⁴   ñaŋ⁵ bai⁴
    at that time ...   king   CLF.one have KEEP
    (Khamyang story, *The Hunter's Parrot*, No. (1), spoken by Deben Chowlik.

The first morpheme in (51) is a combination of *mur⁵* 'time' and *nan⁴* 'that', with the whole word being in the fourth tone. A second fused morpheme is [kɔŋ⁴] which is a combination of *kɔ⁴* 'CLF' and *nɤŋ⁵* 'one'. It also takes the fourth tone. In both cases the fourth tone (high falling glottal) seems to override the fifth tone (low falling).

Perhaps the contour and other features of the fourth tone are stronger than those of the fifth tone, and this is why both of these fused forms bear the fourth tone.

6.5 Some preliminary observations on Ahom phonology

6.5.1 Consonants

Based on the Ahom alphabet, and assuming that there is a one to one correspondence between the glyphs for the initial consonants and their phonemes, the following phoneme table can be presented:

	Bilabial	Dental / Alveolar	Palatal	Velar	Glottal
Voiceless unaspirated stops	p	t	c	k	(ʔ)
Voiceless aspirated stops	ph	th		kh	
Voiced stop	b	d			
Nasals	m	n	ñ	ŋ	
Voiceless fricative		s			h
Semi vowel			y / ʒ		
Rhotic Approximant	r				
Lateral Approximant	l				

Table 80: Possible Consonant Phonemes in Ahom

Because Ahom is no longer spoken as a mother tongue (see above 2.2), any discussion of the Ahom consonants must be based on written sources. We are fortunate to have access to an eighteenth century Ahom-Assamese word list, the *Bar Amra* (manuscript No. 31 in the Department of Historical and Antiquarian Studies, Guwahati.). This manuscript, written entirely in Ahom script, gives definitions for several thousand Ahom words in Assamese. By analysing this manuscript, we may be able to establish the status of at least some of the Ahom phonemes.

For example, in Assamese the reflex of proto Indic *j, the palatal semi vowel, is realised as [ʒ], [dʒ] or [z] and phonemicised as /z/[54]. In the *Bar Amra*, this consonant is usually written with Ahom <w> which is the reflex of proto Tai *j, also a palatal

---

[54] I am very grateful to Dr. Jyotiprakash Tamuli of Guwahati University for his assistance on Assamese phonology.

semi vowel. Given that Aiton shows fricativisation of the palatal semi vowel (see 6.3.1.1.6 above), it might be that, at least in the late 18th century, the reflex of the the palatal semi vowel in Ahom was [ʒ] or [dʒ].

Example (52) from the *Bar Amra* exemplifies its use

52) ᨾᨠ ᩅᨶ ᨷᩣ ᨿᩣᨷᩬᩁ ॥
kaa cam baa yaa bɔr / ʒaa bɔr
go PRT say go (Assamese zabɔ)
'*Kaa* means 'go'.'

The reflex of proto Southwestern Tai *c is still pronounced as [c] in both Aiton and Phake. However in Assamese, the reflex of proto Indic *c is always [s]. Present day Ahoms generally pronounce Tai words with initial /c/ as if they had initial /s/. If this were so in the late 18th century, we might expect to find Assamese words with /s/ being written with both Ahom <vɔ> /c/ and <w> /s/. However, in that portion of the *Bar Amra* so far examined, there is no examples of Ahom <vɔ> being used for /s/ in any Assamese word. This suggests that Ahom /c/ was maintained as a phoneme towards the end of the Ahom era.

There is one interesting orthographical usage in the *Bar Amra*, and that is the use of subscript /r/ to mark the voiced aspirate series. For example, in (53), the Assamese word is pronounced [mukʰ bʰɔŋibɔr]. The voiced aspirate /bh/ is written with Ahom <w> /ph/ and the subscript /r/. In Ahom this cluster would have been pronounced /phr/.

53) ᨡᩬᩴ ᩅᨶ ᨷᩣ ᨾᩩᨡᨷᩁᩬᩴ ॥
khɛ cam baa mukh phraŋi bɔr
khɛ PRT say remove the evil eye (Assamese mukʰ bʰaŋibɔ)
'*Khɛ* means 'to remove the evil eye'.'

As Diller (1992:14) observed, Ahom maintains at least one of the proto preglottalised stops, namely /d/, which is a reflex of proto *d. The Ahom symbol for /d/, <ɗ> is very

similar to that for /n/, and it is not clear which letter is intended. This may suggest that in Ahom, as in Aiton, words with initials that are reflexes of proto *d varied between /n/ and /d/ in initial position.

In written Ahom it can sometimes be difficult to distinguish <w> /m/ from <ɔ> /b/. Ahom /b/ is itself a merge of proto *b and proto *w, and since in many other Tai varieties the reflexes of proto *b merge with proto *m, the orthographical similarity of these two symbols is a source of some problems.

Finally, the printed edition of the *Bar Amra* (B.K. Barua and Phukan 1964), together with the earlier *Ahom–Assamese–English Dictionary* (G.C. Barua 1920) contained a number of words with initial consonant clusters. Some of these do appear in Ahom manuscripts, but it seems that many do not. The extent to which B.K. Barua and Phukan (1964) can be relied upon has not yet been fully determined. Historical reconstructions which rely on G.C. Barua (1920) or B.K. Barua and Phukan (1964) should be re-examined in the light of these findings.

6.5.2 Vowels

Only six vowels are written in Ahom, discussed below in 7.1. Therefore it will be difficult to reconstruct the Ahom vowel system from the script alone. One clue, however, will come from the marking of front vowels in open syllables. In Phake, where there are three contrastive front vowels, words like ye^2'granary' and me^2'wife' are written with initial soi^1 tho^2 (ᨧ-), just as words with final /-ɛ/ would be. In Aiton, on the other hand, where /e/ and /i/ have merged, these words are written with $luŋ^1$ ki^5 $teŋ^1$ ($-^ʒ$) [ii].

The Aiton pattern is also found in Ahom, at least in the *Bar Amra*, where the word for wife is written as ᩅ (mii) and granary as ᨿ (yii), as in examples (54):

54.1) ... ꯀ ...

mii cam baa tii rɔ tak
mii PRT say female (Assamese tirɔta)

'*Mii* means 'a female'.'

(*Bar Amra*, p 8 top line)

Link to a photograph of this page of the Bar Amra (\photos\Bar Amra\p 8.jpg)

54.2) ... ꯀ ...

yii cam baa phrɔralɔk
yii PRT say granary (Assamese bʰɔral)

'*Yii* means 'a granary'.'

(*Bar Amra*, p 13 third line)

Link to a photograph of this page of the Bar Amra (\photos\Bar Amra\p 13.jpg)

It is therefore probable that the Ahom vowel system, at least in the final stages of the language, was the same as Aiton (see above 6.3.2.2). This is given in Table (81):

i		ɯ	u
ɛ			ɔ
	a, aa		

Table 81: Possible vowel phonemes in Ahom

6.5.3 Tones

Almost nothing is known of the original tonal system of the Ahoms. Grierson, having indicated that there should be tones in Ahom, and that tradition is silent on the matter, states:

> "Moreover, in the one word, the tones of which I have been able to ascertain, they differ from those of the Khāmtī and Shān. This is the word *mǎ*, which, when it means 'a horse', has in Āhom a long tone, and in Khāmtī and abrupt tone, while *mǎ*, 'a dog', has in Āhom an abrupt tone, but in Khāmtī and Shān a rising inflection." (1904:90)

It would be most curious if this statement reflected the situation. The abrupt tone referred to is presumably a glottalised or creaky tone, a feature of C tones and occasionally B tones, but rarely if ever of A tones (see the discussion of tones in

6.1.5.3 above). Grierson gives no information about his source for this claim, and it must be treated with some caution.

The only attempt to reconstruct Ahom tones has been Weidert (1979), who compared the Shan, Standard Thai and Khamti tonal systems in order to make a reconstruction of the Ahom tonal system (see above 3.2.14.2). I do not believe that this methodology was valid. First of all it fails to take into account the tonal systems of the other Tai languages remaining in Assam, the Aiton, Khamyang and Phake. Secondly, Weidert's methodology gave equal weight to each of the three languages, even although Standard Thai is part of the Shan group and on linguistic grounds appears to have been separated from Ahom for a long time. Thirdly, not enough languages are examined.

Future attempts to reconstruct the original Ahom tone system might involve:

- Exploring foreign sources, particularly from China, which have data about the Ahom. It is likely that Chinese missions visited Assam over the centuries, and if they did, they may have collected some linguistic information about Ahom. Any recording of the Ahom language in Chinese would encode information about the Ahom tones.
- Comparing the tonal systems of the existing Tai languages of Assam and trying to explain how they came about. Existing Tai tone systems are quite different from those in, for example, Shan state of Burma. This may be due to the influence of the Ahom.

6.5.4 Ahom and Aiton

Interestingly, the putative reconstructions for the inventories of Ahom consonants (see above Table (80)) and Ahom vowels (see above Table (81))are very similar to those found for Aiton, (see above 6.3.1 and 6.3.2 respectively). In addition, Aiton and Ahom both preserve some initial consonant clusters.

Given this, and the Aiton traditions about contact with King Suhungmung as far back as 1500 and Grierson's statements on the Aitons being interpreters for the Ahoms (see above 2.3.1), it might be well to rethink the relationship between Aiton and Ahom.

It may be that some aspects of contemporary spoken Aiton, particularly those that differ from the other Tai languages, are features that might have been present in the Ahom language.

It is to be hoped that further research on the relationship between Aiton and Ahom will be possible in future years.

7. Writing

All the various Tai communities in India are literate. The scripts that they use are all related to the other scripts of mainland Southeast Asia, and like them are descended from Indic scripts. There are two distinct systems in use amongst the Tai of Assam: the Ahom script, and the scripts of the other Tai groups. The living Tai scripts are known in Phake as ၂၆၀၀၀ရဝ�3 *to² lik⁴ tai²* 'Tai writing', or simply Lik Tai.

Only a brief survey of the Ahom script will be presented here, so that it can be referred to and compared with the script in use amongst the other Tai groups. The Ahom script was discussed in very great detail by Ranoo (1986).

7.1 Ahom Script

Scholars differ as to the origins and date of the Ahom script. The leading scholars among the Ahom community, for example J.N. Phukan, believe that the Ahom script was brought with King Sukapha (ၯၐၭၵ suɯ kaa phaa), suggesting a date for the genesis of the script in the 12th or 13th century. This, however, would appear to predate the time at which the merger of proto voiced and voiceless consonants (see 6.1.1.2 occurred[55], and had the script been devised at that time, it would have been expected that it would have marked this distinction, as the Thai, Lao and Lanna scripts all do. This issue will be discussed in detail below in 7.6.

Table (1) lists the initial consonants of Ahom in two fonts, the first of which is based on the printed Ahom font developed for the publication of Ahom texts in the first half of the 20th century, and the second of which is based on an Ahom manuscript. The transcription into Roman script is that of B.K. Barua and Phukan (1964). The order of the consonants in Table (1) is given by B.K. Barua and Phukan (1964:198), itself based on the order in which the consonants first appear in the *Bar Amra* (see 3.2.9)

---

[55] The script used in Thailand dates from around the 12th or 13th century. The fact that there are 44 consonant glyphs in the Thai script is thought to suggest that the distinction of voiced and voiceless consonants was still in place at that time in Thailand.

Ahom letter	Ahom letter (manuscript font)	Transcription into Roman, B.K. Barua and Phukan (1964)	Reconstructed phoneme	Cognate phoneme in Phake	Cognate phoneme in Aiton
m	ᾓ	k	/k/	/k/	/k/
ᾝ	ᾝ	kh	/kʰ/	/kʰ/	/kʰ/
ᾢ	ᾢ	ng	/ŋ/	/ŋ/	/ŋ/
ᾗ	ᾗ	n	/n/	/n/	/n/
ᾤ	ᾤ	t	/t/	/t/	/t/
ʋ	ʋ	r	/p/	/p/	/p/
ᾦ	ᾦ	ph	/pʰ/	/pʰ/	/pʰ/
ᾧ	ᾧ	b	/b/	/m/, /w/	/b/, /w/
ᾨ	ᾨ	m	/m/	/m/	/m/
ᾩ	ᾩ	j	/j/~/dʒ/	/j/	/ʒ/
ᾫ	ᾫ	s	/tɕ/	/tɕ/	/tɕ/
ᾬ	ᾬ	th	/tʰ/	/tʰ/	/tʰ/
ᾭ	ᾭ	r	/r/	/h/	/h/, /r/
ᾮ	ᾮ	l	/l/	/l/	/l/
w	ᾯ	ch	/s/	/s/	/s/
ᾰ	ᾰ	ñ	/ɲ/	/ɲ/	/ɲ/
ᾱ	ᾱ	h	/h/	/h/	/h/
ᾲ	ᾲ	–	–	–	–
ᾳ	ᾳ	d	/d/	/n/	/d/
ᾴ	ᾴ	dh	/dʰ/	–	–

Table 1: Initial Consonants in Ahom

The column headed 'reconstructed phoneme' posits the pronunciation of these consonants at the end of the period of the Ahom Kingdom. This is discussed at 6.5.1 above.

The last two consonants may have been added later. The glyph ᾳ /d/ almost certainly represents an Ahom phoneme, but the symbol is very similar to ᾗ /n/. As discussed above in 6.5.1, the phonemes /n/ and /d/ may have been in variation in Ahom, as they are in Aiton today. The glyph ᾴ /dh/ is listed in B.K. Barua and Phukan (1964), but it is not accepted as an Ahom phoneme. The glyph was developed from ᾬ /th/ and was probably mostly only for borrowed words.

In addition to these consonants, all of which occur regularly in the manuscripts, there are a number of consonants used for borrowed words, most of which are derived from existing Ahom consonants. These are shown in Table (2)

Ahom letter	Ahom letter (manuscript font)	Transcription into Roman, Barua and Phukan (1964)
᾵	᾵	g
	ᾶ	g
ᾷ		g
Ᾰ		gh
Ᾱ		jh
Ὰ		bh

Table 2: Ahom Consonants used for borrowed words

Table (3) shows the vowels of Ahom. As in most of the alphabets based on old Indic alphabets, vowels cannot stand alone, but have to be attached to a consonant. In Table (3) they are all shown below with initial /k/ (ᾓ), and with a final /-k/ when showing the method of writing medial vowels.

Ahom Symbol	Recon-structed Phoneme	Cognate sound in Phake	Cognate sound in Aiton	Ahom Symbol	Recon-structed Phoneme	Cognate sound in Phake	Cognate sound in Aiton
Initials				Medials			
[symbol]	/a/, /aː/	a, ā	a, aa	[symbol]	/a/, /aː/	a, ā	a, aa
[symbol]	/i/	i (ī)	i (ii)	[symbol]	/i/, /ɛ/	i, e, ɛ	i, ɛ
[symbol]	/ɛ/	e, ɛ	ɛ				
[symbol]	/u/	u (ū), o	u (uu)	[symbol]	/u/	u, o	u,
[symbol]	/ɔ/	ɔ	ɔ	[symbol]	/ɔ/	ɔ	ɔ
[symbol]	/ɯ/	ü, ə	ɯ	[symbol]	/ɯ/	ü, ə	ɯ,

Table 3: Ahom Vowels

The phonemic values given for the vowels in Table (3) are discussed above in 6.5.2.

In addition to the consonants and simple vowels there are three symbol which are used for diphthongs and vowels with final consonant combinations. These are listed in Table (4):

[symbol]	/auɯ/
[symbol]	/am/, /ãm/
[symbol]	/ai/, /ãi/

Table 4: Combined vowel and consonant signs or diphthong signs in Ahom

7.2 Aiton & Phake

7.2.1 Initial Consonants

The Tai alphabet, ([symbols] *tuu² lik³ tai²* in Aiton and [symbols] *to² lik⁴ tai²* in Phake), contains 18 initial consonants. The symbols used vary from manuscript to manuscript and from scribe to scribe, with those presented in Table (5) being those in the Aiton and Phake fonts (see below 7.8). The Phake transcription in this table is that of Banchob (1987), with the Aiton transcription being that proposed in this thesis.

Aiton Symbol	Phake Symbol	Transcription in Aiton	Aiton Phoneme	Phonetic realisation in Aiton	Transcription in Phake	Phake Phoneme	Phonetic realisation in Phake
[symbol]	[symbol]	k	/k/	[k]	k	/k/	[k]
[symbol]	[symbol]	kh	/kʰ/	[kʰ], [x]	kh	/kʰ/	[kʰ], [x]
[symbol]	[symbol]	ŋ	/ŋ/	[ŋ]	ŋ	/ŋ/	[ŋ]
[symbol]	[symbol]	c	/c/	[tɕ]	c	/tɕ/	[tɕ]
[symbol]	[symbol]	s	/s/	[s], [ʃ]	s	/s/	[s]
[symbol]	[symbol]	ñ	/ɲ/	[ɲ]	ñ	/ɲ/	[ɲ]
[symbol]	[symbol]	t	/t/	[t]	t	/t/	[t]
[symbol]	[symbol]	th	/tʰ/	[tʰ]	th	/tʰ/	[tʰ]
[symbol]	[symbol]	n & d	/n/, /d/	[n], [d]	n	/n/, /d/	[n]
[symbol]	[symbol]	p	/p/	[p]	p	/p/	[p]
[symbol]	[symbol]	ph	/pʰ/	[pʰ], [ɸ]	ph	/pʰ/	[pʰ], [ɸ]
[symbol]	[symbol]	m & b	/m/, /b/	[m], [b]	m	/m/, /b/	[m]
[symbol]	[symbol]	3	/j/	[ʒ], [ʝ]	y	/j/	[j]
[symbol]	[symbol]	r	/r/	[r]	-		
[symbol]	[symbol]	l	/l/	[lʲ]	l	/l/	[l]
[symbol]	[symbol]	w	/w/	[w], [β]	w	/w/	[w]
[symbol]	[symbol]	h	/h/	[h]	h	/h/	[h]
[symbol]	[symbol]	-	(ʔ/)	-	-	(ʔ/)	-

Table 5: Initial Consonants in Aiton & Phake.

As can be seen in Table (5), there are two Aiton phonemes not expressed in the script, namely /d/ and /b/. The Aiton sometimes use the Burmese symbols ɜ for /d/ and ʊ for /b/, and these have been used from time to time in the transcription of Aiton texts.

In Phake, /r/ is not a phoneme, although it may be pronounced in careful speech. The symbol is used in manuscripts for Pali words.

7.2.1.1 Shapes of the Consonants

Many of the consonants are based on circular elements. The symbol for /w/, which is a plain circle in Burmese and Shan, was in the opinion of villagers of both communities preferable if there was a small black dot on the inside of the left hand side of the circle, as in Ⓞ. This black dot is a feature of the writing of the Tais of Northeast India.

In some manuscripts the black dot is much larger than in others. The large black dots are exemplified in the Aiton *Creation of the World* which may be several hundred years old (see \photos\books\Lik Kheu Meung\Creation of the World 2.jpg). The smaller type is exemplified in the *Book of history from the time of the ancestor Chaw Tai Lung up until Sukapha* (see \photos\books\Lik Kheu Meung\History of Chaw Tai Lung to Sukapha 1.jpg)

Many of the consonants are based on one or two circles, some of which may be open. The glyph ∞ /th/, for example, is formed from two adjacent circles, whereas ∞ /t/ is formed from a circle on the left and a circle open at the bottom on the right.

The small black dots were added to most of the consonants, although in the case of /c/ and /ph/ – both consonants based on a full circle and an incomplete circle – the Aiton consultants asked for a dot on the inside of the circle on the right of the character and the Phakes preferred it not to be there. It appears that the Phakes preferred to have a dot only when it was on the left hand circle, whereas the Aitons would allow it for any one circle, but not in both, as would be the case in /th/.

The only major difference between the two scripts was with the character for /n/, which has a different form for Aiton and Phake, although both are reflexes of symbols which were historically the same.

In the case of /ph/ there is an alternate version where the open circle of the first half of the character is flipped. This was the form preferred by the late Aimya Khang Gohain, as in, as in Table (6):

Aiton	Phake	Form proposed oy Aimya Khang Gohain	Phoneme
∞	∞	ഥ	ph

Table 6: Forms of the consonant /ph/

7.2.2 Final consonants

Table (7) lists those consonants which are found in final position, where they are generally marked with the sign -*, called *sat'*. This sign is now required, but is often absent in old manuscripts, (see for example *Book of calling back the khon*, \photos \khamyang\books\Lik Hong Khon\Hong Khon p3.jpg)p3, showing Nos. (40) ff. in the analysis.)

Aiton Symbol	Phake Symbol	Transcription	Notes
ന്ദ്	ന്ദ്	-k	
ξ	ξ	-ŋ	
ന്ദ്	ന്ദ്	-t	used only in certain words after -ʔ /i/, /e/, /ɛ/
ന്ദ്	ന്ദ്	-n, -i	/n/ after -ʔ, causing the vowel to be realised as /e/,/ɛ/. /i/ after ꞈ, realised as /ui/ or /oi/
ഠ്	ഠ്	-t	
ꞌ	ꞌ	-n	
ട്	ട്	-p	
ട്	ട്	-m	
ട്	ട്	-u, -Ø	zero after ꞈ, which is realised as /ɯ/ (/ü/) or /ɤ/ (/ə/).

Table 7: Final consonants in Aiton and Phake

7.2.2.1 Subscripted consonants

Several consonants are also found in subscripted form. There are two types. Firstly, the consonants that can be the second member of an initial consonant cluster, that is the approximants /r/, /w/ and /y/. These have a different form in clusters from their initial form, listed in Table (8), in combination with the letter /k/.

Tai symbol	Name in Phake	Transcription	Notes
၆	yā¹ lɛt¹	kr	
၇	hā¹ tho¹	kw	
၇	ya¹ pɛn¹	ky	This sign is sometimes used to indicate the vowels /e/ or /ɛ/.

Table 8: Subscripted consonants in Aiton and Phake

A second series of subscripted consonants is used for multisyllabic borrowed words, in which there may be doubled consonants in the language from which the words were borrowed. The subscripted consonants that have been recorded in Aiton manuscripts are listed in Table (9):

Tai Character	Transcription	Examples
ၵ	k	ဎၵၢၵ sak¹ ka¹ rit¹ 'era'
ၸ	c	
ၲ	t	ၮ aa¹ tii¹ pa¹ tii¹ 'king, lord'
ၳ	th	ၯ puk¹ ta¹ huu² 'Wednesday'
ၺ	p	ၡ pu³ pa³ wi¹ tɛ¹ ha³ 'Pubbavideha'
ၽ	y	
ၿ	l	ၠ wɛ¹ pu¹ la¹ 'Vepula'

Table 9: Subscripted final consonants in Aiton and Phake

7.2.3 Vowels

Whilst the consonant system is fully capable of expressing the range of phonemes, the vowel system is underspecified, as is shown in Table (10) where the various vowel symbols are combined in open syllables with the dummy consonant ၹ.

Aiton Symbol	Phake Symbol	Transcription in Aiton	Transcription in Phake	Phonemic in IPA script (Aiton)	Phonemic in IPA script (Phake)
ၹ	ၹ	a, aa	a, ā	[a] [aː]	[a] [aː]
ၹ	ၹ	a, aa	a, ā	[a] [aː]	[a] [aː]
ၹ	ၹ	ii	ɪ	[iː]	[iː]
ၹ	ၹ	uu	ʊ	[uː]	[uː]
ၹ	ၹ	ɛ	e, ɛ	[e]	[e] & [ɛ]
ၹ	ၹ	ɔ	ɔ	[ɔ]	[ɔ]
ၹ	ၹ	o	o	[o]	[o]
ၹ	ၹ	ɯ	ü, ə	[ɯ]	[ɯ] & [ɣ]/[ə]

Table 10: Vowel symbols in open syllables

By far the most common symbol for /a/ is ၙ, but the alternate symbol ·ː is used in several circumstances. In manuscripts it is often used for the unstressed syllables in Pali words such as ၹ tuk¹ kha⁴ "suffering" from Pali *dukkha* (*Grandfather teaches Grandchildren* Proverb No. 15).

However, the symbol is also found for Tai words. Table (11) lists all of the uses of this symbol in the Aiton manuscript, *Treaty between the Aitons and the Turungs*:

Aiton script	phonemic transcription	English	Reference in text
ၹ	khaa²	straw	(4)
ၹ	pa¹ caɯ³	'messenger' - literally 'person-employ'	(25)
ၹ	pa¹ kɛ¹	old men	(31)

ဟ္ခြင္းလား	cau³ phaa³	king	(32)
လားက်	la¹ aam²	Proper name	(35)
ကားၛၚလားၕ	kan² naa³ taa² paai²	in future, literally 'side-face-eye-side'	(51), (67)
လား	ta³	PRT - hortative	(56)
ကျိၕလားၔၜ	kum¹ mai³ la¹ wɔ¹	Proper name - probably Singpho	(102)
ကၛၔလၚ	aa¹ ma¹ khaam¹	accept on behalf of	(103)
ၔား	ma¹	NEG	(111)

Table 11: Use of the symbol –ː in *The Treaty between the Aiton and the Turung*

Referring to the symbol –ː, Diller (1992:19) observed that "occasionally a visarga-like symbol" was used in words which have or once had the C tone (see 6.1.5.3). There are some examples in Table (11) which are the original C tone, such as *cau³ phaa³* 'king' (> *faaC4 'sky'), but most of the examples of the use of –ː are for short /a/. In *The Treaty between the Aiton and the Turung*, No. (25), we also find the word လားၔ္ဈ *pa¹ cau³* 'messenger', which is arguably a case of marking the C tone with the symbol –ː.

Another use for the symbol –ː is in the special combination ᷆ː, in words like ᷆ː *nui³* 'with', also a word with a reflex of the C tone. (see below, Table (15)),

The symbol –ː is still used by in writing as shown in example (1):

1) က္ဈ   လား   ၝၚၔ   မၚ   ၕၚ႙၂ၕ   တ်   ၜၕၕ   ၕၝ
    kau² khaa¹ ʒaŋ² maa²   nɔ¹ bin¹ tɛm³ laai² maai¹ sii¹
    1Sg   be   be   COME   Nabin   write letter   PRT

    ၜၕၕ   ၝၚၕ   ၔား   တၝ   ၕၝ   ၔၕ   ၕၝ ॥
    phaak¹ hot³ maa² tii² suu¹ cau³ ʒau³
    send arrive COME at   2Pl   RESP FINISHED
    'It is Nabin who has written this letter to you.'
    (Phake letter, No (2), written by Nabin Shyam Phalung, 1997)

When translating example (1), Nabin Shyam stated that the word *khaa¹* meant 'be'. However there is reason to think that this might actually be the word *khaa³* 'slave', which would be yet another example of this symbol being used in a word the tone of which is a C tone.

There is a further symbol for /a/, –ɔ which is only found in older manuscripts. This is the usual symbol in Burmese script for /a/. In Tai orthography, it is used both for Pali loan words, as in *tā² ti¹ yā⁴* in example (2), and in native Tai words, as in *cā⁴* in example (3):

2) က္ဈ   လၚ   လၕ   ၕၔ   လားၕၚ   ၕၝ   ၝၕ
    kau² sī⁶ (sē⁶) pan¹ khɔi⁵   tā² ti¹ yā⁴   tāu⁵ khɔp⁴
    1   PRT   whirl meet again third   turn complete whole circle

    ၝၕ   ၝၝၕ   ၕၕ   ၕၝ ।
    ho⁶ hauk¹ khiŋ² yɔm²
    head grey hair body be decreased
    'I have completed the third stage of my life, my hair is grey and my body is decreasing.'
    (Tai Phake Text, *Grandfather teaches Grandchildren*, Intro. No. 5, read by Yehom Buragohain)

3) ၕၕ   ၕၚ   ၝၕ   ၕၝ   ၕၕ ।
    pɔn⁵ cā⁴ hau² yā¹   phün⁴
    other bad we   don't return

    ၝၕ   ၕၝ   ကၕ   ၕၕ   ၕ ॥
    hun² caü² kan³ khan⁶ pɛ⁴
    restrain mind control quick can
    'Others may be angry with you but you are to endure it and not return it. Restrain your mind as quickly as possible.'
    (Tai Phake Text, *Grandfather teaches Grandchildren*, Proverb No. 36, read by Yehom Buragohain)

Whilst the system of open vowels described in Table (10) covers the existing phoneme inventory of Aiton reasonably well, it clearly underspecifies for Phake. In Phake, the contrasts between [e] and [ɛ] and that between [ɯ] (written <ü>) and [ɣ] (<ɔ>) are not specified.

For closed syllables, the writing system is even less specific. Table (12) shows the marking of vowels in closed syllable with initial dummy consonant ᴄᴏ and final /ŋ/, written as ᴇ.

Tai Symbol	Transcription in Aiton	Transcription in Phake	Phonemic in IPA script (Aiton)	Phonemic in IPA script (Phake)
ᴄᴏᴇ	a, aa	a, ā	[a] [aː]	[a] [aː]
ᴀᴏ	i, ɛ	i, e, ɛ	[i], [ɛ]	[i], [e], [ɛ]/[ai]
ᴄᴏᴇ	u	u, o	[u]	[u], [o]
ᴄᴏᴇ	ɔ	ɔ	[ɔ]	[ɔ]/[au]
ᴄᴏᴇ	ɯ	ü, ə	[ɯ]	[ɯ] & [v]

Table 12: Vowel symbols in closed syllables

The lack of contrast in the front vowels in closed syllables means that in Phake there are potentially 18 different pronunciations of the glyph written as ᴄᴏᴇ <kiŋ>. Table (13) lists the various words written as ᴄᴏᴇ <kiŋ> found in Banchob's 1987 *Phake–English–Thai Dictionary*:

> ᴄᴏᴇ₁₁ kiŋ¹ a small branch of the tree
>
> ᴄᴏᴇ₂₁ kiŋ³ to roll,to tumble down
>
> ᴄᴏᴇ₃₁ kiŋ⁴ green (mostly for dye or paint)
>
> ᴄᴏᴇ₄₁ keŋ² a kind of grass, softer than the others
>
> ᴄᴏᴇ₅₁ keŋ² an iron tripod to support cooking pots or pans over a fire
>
> ᴄᴏᴇ₆₁ keŋ² a period, time
>
> ᴄᴏᴇ₇₁ keŋ³ decent or nice to look at (mostly the fruits)
>
> ᴄᴏᴇ₈₁ kaiŋ² dwarfed, stunted
>
> ᴄᴏᴇ₉₁ kaiŋ³. to use sticks to clean one's self after evacuating the bowels
>
> ᴄᴏᴇᴄᴏ₁₁ kaiŋ⁴ khe⁶ a shuttle used for weaving a net

Table 13: Forms of the glyph ᴄᴏᴇ <kiŋ> in Phake (after Banchob 1987)

Had all 6 tones been realised for each of *kiŋ*, *keŋ* and *kaiŋ* in Phake, there would have been eighteen different pronunciations of the same glyph. As it is, there are eight.

7.2.3.1 Special symbols for vowel–consonant or diphthong combinations.

There are a number of special symbols for combinations of vowel and consonant, or for diphthongs. These are listed in Table (14):

Tai Symbol	Transcription in Aiton	Transcription in Phake	Alternative
ᴄᴏ	am, aam	am, ām	ᴄᴏ
ᴀᴏ	im, ɛm	im, em, ɛm	ᴀᴏ
ᴄᴏ	ai, aai	ai, āi	
ᴄᴏ	ɔi	ɔi	
ᴄᴄᴏ		oi	ᴄᴏ
ᴄᴏ	ui	ui, oi	ᴄᴏ
ᴄᴏ	au, aau	au, āu	ᴄᴏ, ᴄᴏ, ᴄᴏ
ᴄᴏ	aɯ	aü	

Table 14: Special signs for vowel–consonant combinations and diphthongs.

7.2.3.2 Names of the vowel symbols

Apart from the consonants, the names of which are pronounced as the consonant followed by a short /a/, as in ᴄᴏ [ka?¹], all of the symbols in the orthography of the Tai Phake have names based on Burmese. These names, elicited from Ai Che Let Hailung of the Tai Phake, are listed in Table (15):

Symbol	Sound	Name	Transcription of name in Phake
ᴄ	–a, ā	ᴄᴏ	khyā²
ᴄ	–i	ᴄᴏᴇ	luŋ¹ kɪ⁵ teŋ¹
ᴄ	–ɪ	ᴄᴏᴇᴄᴏ	luŋ¹ kɪ⁵ teŋ¹ saiŋ¹ khāt⁴

—	-u, o	(Tai script)	ta¹ khyauŋ¹ ɛn⁴
—	-ū	(Tai script)	nā¹ khyauŋ¹ ɛn⁴
ɕ —	-e, ɛ	(Tai script)	soi¹ tho²
—ʾ	-ai, āi	(Tai script)	nāu¹ pet⁴
ɕ—ʾ	c	(Tai script)	soi¹ tho² khyā²
—˙	-am, ām	(Tai script)	yam¹ tāŋ¹
—˙	-um	(Tai script)	ta¹ khyauŋ¹ ɛn⁴ yam¹ tāŋ¹
—ɔ	-ɔm	(Tai script)	wa¹ sɔi¹ yam¹ tāŋ¹
—ʾ 6	-ɯ, ɤ	(Tai script)	luŋ¹ kɪ̄⁵ tɛŋ¹ ta¹ khyauŋ¹ ɛn⁴ wa¹ sāt⁴
— ẟ	-au	(Tai script)	wa¹ sāt⁴
subscripted /w/	subscripted /w/	(Tai script)	hā¹ tho¹
—ɔ	-ɔ	(Tai script)	wa¹ sɔi¹
—ʾɔ	-ɔi	(Tai script)	wa¹ sɔi¹ nāu¹ pet⁴

Audio Link for the pronunciation of the names of the vowels (wave files\phake\texts \i-10-2-7.mp3)

Table 15: Names of vowel symbols in Tai Phake

When spelling out words, these names are used, as in example (4), which is the spelling of the word (Tai script) *kin²*'eat'.

4) (Tai script)
    ka¹   luŋ¹ kɪ̄⁵ tɛŋ¹   na¹   sāt⁴   kin²
    /k/   letter /i/   /n/   final consonant   'eat'
    '*Eat* is spelled ka i final na.'

7.2.4 Ligatures

In addition to the system described above, there are two types of ligature in Tai script. One is the use of a special symbol to write a word, often a very common word. This will be called the *əi²type* because the most common example of it is the word *əi²* 'VOC'. These are listed in Table (16) and appear to be a closed class.

Form in Phake	phonemic	Form in Aiton	phonemic	gloss
(Tai)	əi²	(Tai)	uri²	"VOC", "EXCL"
(Tai)	niŋ⁵	(Tai)	luŋ²	"one"
(Tai)	sauŋ⁶	(Tai)	sɔŋ¹	"two"

Table 16: *əi²*type ligatures in Tai

The symbol (Tai) or (Tai) is only found with the meanings given above, but the ligature for "two" can also be used for homphonous words which do not mean 'two', as in (5):

5) (Tai script)
    khaaŋ²   sɔŋ¹   cɔi³   saai¹ huŋ¹   mau¹   phaak³
    near   bright   Queen   glorious   NEG   separated
    'He sat close together with his queen, (they were) never separated.'
    (Aiton History, *Book of history (from) the time of the ancestor Chau Tai Lung up until Sukapha*, No. (4)).

The second type of ligature will be termed the *hɔ⁶ kham²type*, after one very common word in this class. In this class, compound words are expressed with the initial consonants of the second member of the compound elided, as is shown in Table (17):

Full form	transcription	Ligature form	transcription	phonemic	gloss
(Tai)	e h a kh am	(Tai)	e h a am	hɔ⁶ kham²	palace
(Tai)	kh y am s a	(Tai)	kh y ɛm a	khyaam² saa¹	good, well

Table 17: *hɔ⁶ kham²*type ligatures in Tai

In Table (17), the columns headed 'transcription' present a romanised transcription of each of the symbols. Some of the transcriptions use digraphs, such as <kh>, to refer to a single symbol. Because of this, a space is used between the transcription of each symbol. The ligature is formed by the omission of the initial consonant of the second syllable.

7.2.5 The writing of borrowed words in Tai script

Since the Tai languages are using an increasing number of words borrowed from Assamese and other languages, including English, some new traditions are arising for the writing of borrowed words. As indicated above in 7.2.3, the use of alternate symbols for /a/ is common in borrowed words.

Since both Assamese and English allow for final consonants which are not permitted in Tai, some interesting traditions have arisen. Example (6), from a poem composed in 2000 and written down in a manuscript, shows the use of writing of borrowed proper names:

6)   လိက်   သုံ   ယျဝီလိက်   မုဝ်လော ॥   ကွယ်ကြွဝီယျ

lik⁴   com²   saˡ tīˡ fɛn⁴   moˡ lɛ⁴   ɔsˡ trɛˡ līˡ yāˡ

book   glad   Stephen   Morey   Australia

"The book of thanks to Stephen Morey, Australia."

(Tai Phake Song, *Song in honour of Stephen Morey*,. (1), read by Ai Che Let Hailung)

The word ɔsˡ trɛˡ līˡ yāˡ has final /s/ in the first syllable, which is contrary to Tai phonology. The use of sāt⁴ to mark a consonant like /s/ that is not a permitted in Tai, is not accepted by all Phakes. For example, Ai Chanta insisted that the Assamese word মূল্য *mul yɔ* "value" would have to be written with *taˡ sāt⁴*, as in (7):

7)   ဟိုင်လဝိင်   လောင်   ဟဝ်   ကော   ဟိုင်   လျာ   ဝုဉ်   ၍   ကော

khəŋ⁵ laiŋˡ   lāŋ⁶   nan⁴   nɛ⁴   nüŋˡ   pɔˡ   khauˡ   wā⁵   nɛ⁴

important   CLF   this   PRT   as   others   3Pl   say   PRT

မုဝ်လျာ   ဗိုဝ်   ဝာ   မ   ဝဝ်   ကိုဝ်   ကာ ॥

mulˡ yɔ⁴   mün²   taˡ   maˡ   yaŋ²   kin²   kāˡ

value   3Sg   will   not   have   eat   GO

"The value, as the Assamese call it, of these important things will be lost."

(Tai Phake *Speech*,. No (11), given by Ai Chanta)

Ai Chanta explained that as *laˡ sāt⁴* was not possible in Phake orthography, *taˡ sāt⁴* would have to be written, as in ဟိုဝ်လကျော် *iˡ saˡ kuˡ⁴* "school", which is also pronounced as *iˡ saˡ kuˡ⁴* or simply *isˡ kuˡ⁴*.

In Aiton, final /l/ is more likely to be written with a final nasal, as in တၟမုဉ် *taˡ mun¹*, betel nut (tamul in Assamese), which occurs in the *Manuscript of the Tai Calendar for the Year 2001/2002 (Sakkarit)*, No (89).

7.3 A note on the revised Khamti Script.

In the early 1990s, Chaw Khouk Manpoong, together with a committee of other Khamti people in Arunchal Pradesh, produced the revised Khamti script. The script is based on the historical forms of the Khamti alphabet, which is virtually identical to the Aiton and Phake. In addition, voiced and voiced aspirate consonants have been devised in order to write words borrowed from English and Assamese. Some of these consonant shapes, such as ဂ /ga/ and ဃ /gha/ are based on Burmese forms, but others, such as ၒ /ja/ are of unknown origin. Furthermore, the script fully marks all the vowel distinctions in modern spoken Khamti, including long and short /a/ in closed syllables.

A two volume Primer teaching the revised Khamti script (Chaw Khouk Manpoong 1993) has been published. This is the first comprehensive attempt to mark tones in a Tai script in India, as well as the first comprehensive attempt to standardise the writing of borrowed words, such as ၦလ bɔl³ 'ball'.

The revised Khamti script has nine tone marks. Of these, five are for marking the citation tones (see the discussion of Khamti tones in 3.2.13), and the other four are found in various prosodic situations. Tone mark No. 1 is used to mark the pronunciation of the consonant names, tone mark No. 2 for unstressed syllables in polysyllabic loan words, and No. 9 is used only in utterance final position.

Table (18) lists the consonants in the revised Khamti script, and table (19) the vowels:

Velars:	ᯂ	ᯄ	ᯅ	ᯆ	ᯇ
	k	kh	g	gh	ŋ
Palatals					
	c	s	j	jh	ñ
Dentals:					
	t	th	d	dh	n
Labials:					
	p	ph	b	bh	m
Semivowels					
	y	r	l	w	
Others					
	h	-			

Table 18. Consonants in the revised Khamti script

	◌ᩣ	◌ᩦ	◌ᩥ	◌ᩩ	◌ᩮ	ᥱ◌	ᥰ◌
	a	i	ii	u	ε	ε	e
	ᥰ◌ᩣ	◌ᩣ					
	o	ɔ	ai	aɯ	ɯ	ɤ	

Table 19: Vowel symbols in the revised Khamti script

It appears that these reforms have not been accepted by most Khamti. The objections to the revised script include aesthetic objections, that the symbols are not sufficiently beautiful, and practical objections, that use of the new script would necessarily involve a loss of knowledge of the traditional scripts.

7.4 Comparison of the Tai Scripts

Table (20) is a comparison of the consonants of the Aiton and Phake scripts with the revised Khamti, a Phake font derived from the manuscript Lama Mang in the library of Namphakey temple, the Ahom printing font, an Ahom font based on an early

manuscript and Shan. Table (20) uses the fonts produced in this project, which are discussed below in 7.8.

Phake	Aiton	Revised Khamti	Phake Script in Lama Mang	Ahom	Ahom Chicken bone MS	Shan	Sound in Phake
							k
							kh
							ŋ
							c
							s
							ñ
							t
							th
							n
							p
							ph
							m
							y
							(r)
							l
							w
							h
							-

Table 20: Comparison of consonants in several Tai scripts

7.5 Some notes on tone marking

Tone is not marked in any of the Tai scripts in Assam, although, as discussed above in 7.2.3, the symbol –: sometimes appears to mark words that are reflexes of those that had the C tone.

Historically, tone was not marked in any of the Shan languages. However in the last 50 years the other Shan languages, Tai Mau (Young 1985) or Dehong (Luo 1997) and Shan (Sao Tern Moeng 1995) have adopted tone marking. Some Tai speakers in India, such as Chau Khouk Manpoong (see above 7.3) and the late Aimya Khang, have been working toward the marking of tone in the Tai scripts of India.

7.5.1 Proposals for marking tones in Tai:

Unfortunately, because the distribution and quality of the tones differ from variety to variety, marking tone quality would lead to different marking for different varieties. As indicated above (see 7.3), a tone marking system has been developed for Khamti, but it has been decisively rejected by the Phakes and the Aitons.

Several people in the Phake community have developed proposals for marking tone, although none of these are in use. One possible system for marking tone in Phake would be to follow the lead of the Shan script reform which was established in Shan state in 1958 (Sao Tern Moeng 1995:i). The tone marks of Shan are given in Table (21).

Shan tone number	Tone quality	Shan tone mark	Example	Gloss
1	no tone mark	rising tone	ꩡ	dog
2	,	low level tone	ꩡ,	shoulder
3	;	mid level tone	ꩡ;	mad
4	:	high level tone	ꩡ:	come
5	°	falling tone with glottal stop	ꩡ°	horse
6	⁚	emphatic tone		

Table 21: Tone marking in Shan (after Sao Tern Moeng 1995)

Table (22) presents a proposal for marking tone in Phake, using the symbols from the Shan system and applying them to the 6 tones of Phake. For example, in Shan, the symbol <:> marks a high level tone. In Phake this would be used to mark $m\bar{a}^1$ 'shoulder', but in Shan the high level tone is spoken on the word for 'come', as shown in Table (22):

Tone number in Banchob (1987)	Tone quality	Phake Example	phonemic	Gloss	Shan
1	high level tone	ꩡ:	$m\bar{a}^1$	shoulder	ꩡ,
2	high falling tone	ꩡ:	$m\bar{a}^2$	come	ꩡ:
3	grave tone with glottal stop	ꩡ°	$m\bar{a}^3$	mad	ꩡ;
4	falling tone	ꩡ:	$m\bar{a}^4$	horse	ꩡ°
5	grave tone	ꩡ°	$k\bar{a}^5$	value	ꩡ;
6	high rising tone	ꩡ	$m\bar{a}^6$	dog	ꩡ

Table 22: Comparison of possible tone marking in Tai Phake with Shan cognates

Unfortunately this type of marking inevitably leads to the Shan forms being marked differently from the Phake forms, because the quality of Shan tones are different. Furthermore, if this system, or one like it, were to be adopted, it would mean that the Aiton would either develop a different system of marking, or would have to learn by rote that certain words were marked in a particular, and arbitrary, way. Table (23), developed from Table (22), adds possible marking of Aiton words:

Mark	Tone No. Banchob (1987)	Possible Phake marking	Gloss	Tone quality in Aiton	Possible Aiton marking	Gloss
:	1	ꩡ:	shoulder	high level tone	ꩡ:	shoulder
;	2	ꩡ:	come, to	high falling tone	ꩡ:	come, to
°	3	ꩡ°	mad	falling tone	ꩡ:	mad
⁚	4	ꩡ:	horse	falling tone	ꩡ:	horse
,	5	ꩡ°	value	high falling tone	ꩡ:	value
–	6	ꩡ	dog	high level tone	ꩡ:	dog

Table 23: Comparison of possible tone marking in Phake and Aiton

As can be seen in Table (23), the Phakes mark the words "dog" and "shoulder"
differently, whereas for the Aiton, these two tones have merged. Presumably the
Aitons would want to write these two words using the same graph. If such a tone
marking system were to be adopted, Aitons could perhaps come to learn to read Phake,
and apply the tone merger rules in (8):

8)   tones 1 & 6   >   1

        tones 2 & 5   >   2

        tones 3 & 4   >   4

However, the distribution of tones for Phake and Khamti is very different. Khamti
splits A2&3 from A4 and merges them with B4. This would create significant
ambiguity. Table (24) lists the comparison across the three varieties:

Boxes	Tone in Aiton	Tone in Phake	Tone in Khamti	Box group designation
A1	high level	rising	high level	a
A2-3	high falling	high falling	mid falling	b
A4	high falling	high falling	high falling	c
B1-3	high level	high level	high level	d
B4	high level	low level	mid falling	e
C1-3	low falling	low falling glottalised	mid rising, glottalised	f
C4	low falling	low falling	low falling, glottalised	g
D1-3	high level	high level	high level	h
D4	low falling	low falling	mid falling	i

Table 24: Comparison of tones in Aiton, Phake and Khamti

In Table (24), lower case letters have been assigned to show all of the possible
distinctions of tone. For example, in Aiton tone boxes designated by the letters b, c
and e have the same tone, the high falling tone. However in Phake there is one tone
for the boxes b and c and another for the box designated e. In Khamti, there is one
tone for boxes b and e and another for the box designated c.

Another alternative might be to mark every box of the Gedney tone chart (see 6.1.5.3),
and simply learn the appropriate tone for that box. This however is too fussy. It would
require 20 different tone marks, for languages in which no more than six tones are
distinguished.

Professor E. Annamalai (pers. comm.) suggested that it might be possible to mark just
the nine groups of boxes in Table (24), with each variety then assigning tone
according to its rules. By this system, The Aitons would know that the boxes
designated as b, c and e would be realised as having a single tone, the high falling
tone, whereas the Phakes would know to distinguish the tone of boxes b and c from
that of box e.

It may be that tone marking is necessary if the Tai scripts are to survive in usage;
however many Tai people remain of the view that it is possible to continue using the
traditional scripts. Any process leading to the marking of tones would have to be
undertaken with wide consultation.

7.6 Development/ History of the Tai scripts

It is generally believed that each of the Tai groups brought their writing systems into
India with them. The fact that both the Ahom and the other Tai scripts are clearly
more closely related to the Mon/Burmese based scripts of Southeast Asia than to any
of the north Indian scripts would certainly suggest that is so.

However, there are considerable problems with accepting that the Ahom script was
brought into India with Sukapha in 1228. As Diller pointed out:

> "... the fact that no Ahom texts examined yet show any trace of distinguishing
> etymological correlates of "low series" from "mid" or "high series" consonants
> appears to indicate that the script was adopted *after* pan-Tai consonant
> mergers of the sort in which sounds of the "low series" consonants presumably
> fell together with certain of the others." (1992:11) (see above 6.1.5.3)

He goes on to speculate whether it might have been possible that there was previously a script which did differentiate low and high series, and was "subsequently modified to accord with the sound changes".

The date of the development of the first Thai scripts is around the time of the Sukhothai kingdom in the late 13th century.[56] Since that writing system apparently preserves the voicing contrasts in proto-Tai, between the voiced (line 4) and voiceless (lines 1 or 2) series, it is presumed that the contrast still existed when the writing system was developed (see above 6.1.5.3). Yet, if the Ahom script was brought to India with Sukhapha in 1228, that contrast would have to have been lost already in the Ahom language.

As to the Aiton/Phake/Khamti script, this is a clear development from Shan, itself developed from Burmese. It borrows Burmese vowel combinations, such as the combination of /i/ and /u/, which in Burmese [ou] and in Phake is [ɯ] or [ɤ]. Similarly, in Burmese is [aw] (as the cognate sign is in Standard Thai), but in Phake it is [ɔ]

There are no manuscripts that can be dated to earlier than the 18th century in any of the Tai Aiton, Tai Phake or Tai Khamyang villages that I have visited. A deeper discussion of the origin of these scripts is beyond the scope of this thesis.

7.7 Alphabetical order of the Tai scripts

The alphabetical order of the consonants as listed in Table (5) above, follows the traditional Indic alphabetical order (velar, palatal, dental, labial). This order is fixed and accepted by all literate members of the Tai community.

---

[56] The traditional date for the invention of the Thai script is 1283. Whether or not one accepts the authenticity of the 'First Inscription', the pillar of King Ramkhamhaeng, the large number of other inscriptions from around this time indicate that the 13th-14th century saw the genesis of the Thai script.

The order of the vowels on the other hand, is not so clear. In a letter written in 1997, Nabin Shyam Phalung gave the order as in Table (25). He later informed me that this order was taught by a visiting monk who came from Burma in the 1970s or 1980s.

ka	kaa	ki	kii	ku	kuu	kɛ	kai	kɔ	kɔɔ
kam	kɯ	kiu	kau	kau	kaɯ	kɔi			

Table 25: Alphabetical order according to Nabin Shyam Phalung

The late Aimya Khang Gohain suggested that the vowel ordering should be slightly amended as in Table (26):

ka	kaa	ki	kii	ku	kuu	kɛ	kai	kɔ	kɔɔ
kam	kum	kɔm	kə/kü	kau	kɔ	kaɯ	kɔi		

Table 26: Alphabetical order according to Aimya Khang Gohain

In Aimya Khang's system, *kiu* will be immediately after *kim* rather than after *ku* as suggested by Nabin Shyam Phalung. The principle is that consonant symbols always precede vowel symbols, and that final consonant symbols are in the same order as initial consonants. Therefore the first word in the Tai alphabet will be *kak*.

Alphabetical ordering is not standardised amongst the Tai in Northeast India. Since there are not yet dictionaries available and since a large proportion of the community are not able to read the Tai script, there has been little discussion of the alphabetical order. The order adopted for the Phake and Aiton lexicons presented in this thesis is that of Table (25).

7.8 Development of the Tai Fonts

The Tai fonts used here were developed in 1997-2000 as part of this PhD project. As noted above in 1.1, the developments of the Tai fonts was the first part of this project to be undertaken.

The fonts were created using Macromedia Fontographer 3.5, and three different processes were used to produce them. The Phake and Aiton fonts were built from existing Burmese fonts, taking those Burmese letters which are similar to the Tai characters and altering them to include the small black circles that are a distinguishing feature of these fonts. Of the 18 letters of the Tai alphabet, 11 could be built in this way. The others, which are quite different from Burmese, were built by combining sections of one letter with sections of another.

When I visited India for the second time, in January 1998, I took with me a laptop computer and demonstrated the shape of the fonts to gatherings of villagers. In Namphakey village, a meeting was held of the older men who are the guardians of the manuscripts in the village. Every letter was discussed and in some cases alterations were made. A similar process was undertaken with Nabin Shyam Phalung for the Aiton script.

When this process was complete, several villagers tried out the computer. The first to do so was Yenow Than Gohain, who was the first person of the Tai community to type her own name in her own language. This was recorded in a photograph (see \biographies \Yenow.doc).

The font called *Phake Ramayana*, in which the shape of the glyphs is based on an old Phake manuscript, was later developed from the Phake Script font.

In the case of the Ahom font, the printed characters from the *Ahom Primer* (G.K.Barua 1936) were scanned and the scanned bitmaps imported into the Fontographer programme. The characters were then adjusted.

The Ahom manuscript font is based on the shape of characters in a manuscript of Ahom augury, a copy of which was made available to me by Sri Atul Borgohain, of Diburgarh. This font was drawn using fontographer, without borrowing from any of the other fonts or importing bitmaps. Characters were often built from others of similar shape, so that, for example, the character for /ph/ ʋɔ was built from the character for /y/ ʋʋ.

Table (27) lists the Tai fonts which have been produced as part of this project:

Font name	Font description	Font file name & Location	Explanatory document
Ahom font	Font based on the printed form of Ahom scripts (after Barua 1987)	ahomfont.ttf	\fonts \Ahomfont.doc
Ahom Manuscript	Font based on Ahom Chicken Bone manuscript	ahom-ms.ttf	\fonts \Ahom-ms.doc
Aiton	Font based on usage in the Aiton community	aiton.ttf	\fonts \Aiton.doc
Khamti	Font based on the revised Khamti script (Chaw Khouk Manpoong 1993)	khamti_.ttf	\fonts \Khamti.doc
Old Shan	Font based on the Shan script in use before the reforms since 1950	oldshan.ttf	\fonts \Oldshan.doc
Phake Modern	Font based on the ideas of the late Aimya Khang Gohain for Phake	phakmod.ttf	\fonts \Phakmod.doc
Phake Ramayana	Font based on the Lama Mang, the Phake Ramayana.	phakeram.ttf	\fonts \Phakeram.doc
Phake Script	Font based on usage in the Phake community	phake.ttf	\fonts \Phake.doc
Shan	Font based on the reformed Shan script	shan.ttf	\fonts \Shan.doc

Table 27: Fonts produced for the Tai scripts

In addition to these nine fonts, this thesis also employs several other non-standard fonts which were developed or altered for the purpose of this thesis. The most important of these is the Banchob font, used for romanisations of the Tai scripts. This font is an adaptation of a phonetic font developed by Richard Horsey.

7.9 On writing the Tai language in Assamese script

Almost all native speakers of the Tai languages can read and write Assamese, but many cannot read their own script. There are a number of manuscripts of songs, stories and other texts which individual villagers have written down in Assamese

script. There is no standard method of realising the Tai language in Assamese script, but after studying several manuscripts, the correspondences listed in Table (28) were found to be most common:

Aiton Symbol	Phake Symbol	Transcription in Aiton	Transcription in Phake	Assamese character	Phonetic realisation in Assamese	Phonemic realisation in Aiton	Phonemic realisation in Phake
ကာ	ကာ	k	k	ক	/k/	/k/	/k/
ဃ	ဃ	kh	kh	খ	/kʰ/	/kʰ/	/kʰ/
၎	၎	ŋ	ŋ	ঙ	/ŋ/	/ŋ/	/ŋ/
ၜ	ၖ	c	c	চ	/s/	/tɕ/	/tɕ/
ၸ	ၸ	s	s	ছ	/s/	/s/	/s/
ည	ည	ñ	ñ	ঞ	/ɲ/	/ɲ/	/ɲ/
တ	တ	t	t	ত	/t/	/t/	/t/
ထ	ထ	th	th	থ	/tʰ/	/tʰ/	/tʰ/
န	န	n	n	ন	/n/	/n/	/n/
၇, ၃		d		দ	/d/	/d/	
ပ	ပ	p	p	প	/p/	/p/	/p/
ဖ	ဖ	ph	ph	ফ	/pʰ/	/pʰ/	/pʰ/
မ	မ	m	m	ম	/m/	/m/	/m/
ဗ, ဘ		b		ব	/b/	/b/	
ယ	ယ	y	y	য়	/j/	/j/	/j/
ယ		ʒ		য	/dʒ/,/ʒ/,/z/	/ʒ/	
ရ		r		ৰ	/r/	/r/	/r/
လ	လ	l	l	ল	/l/	/l/	/l/
ဝ	ဝ	w	w	ৱ	/w/	/w/	/w/
ဟ	ဟ	h	h	হ	/h/	/h/	/h/
အ	အ	-	-	অ	(/ʔ/)	(/ʔ/)	(/ʔ/)

Table 28: Assamese letters used for writing Tai consonants

Initial vowels are listed in Table (29):

Aiton Symbol	Phake Symbol	Transcription in Aiton	Transcription in Phake	Assamese character	Phonetic realisation in Assamese	Phonemic realisation in Aiton	Phonemic realisation in Phake
အာ	အာ	a, aa	a, ā	আ	/a/	/a/, /aa/	/a/, /aa/
ဣ	ဣ	ii	ī	ই	/i/	/ii/	/ī/
ဥ	ဥ	uu	ū	উ	/u/	/uu/	/ū/
	ဩ		o	ও	/ou/		/o/
ဧ	ဧ	ɛ	e, ɛ	এ	/e/	/ɛ/	/e/, /ɛ/
ဩ	ဩ	ɔ	ɔ	অ	/ɔ/	/ɔ/	/ɔ/
ဪ	ဪ	ɯ	ə, ü	ঔ	/au/	/ɯ/	/ə/, /ɯ/

Table 29: Assamese letters used for writing initial Tai vowels

Non initial vowels, written in both and Assamese using both initial and final /k/ are listed in Table (30):

Aiton Symbol	Phake Symbol	Transcription in Aiton	Transcription in Phake	Assamese character
ကာက	ကာက	kak, kaak	kak, kāk	কাক
ကိက	ကိက	kik	kik	কিক
ကေက	ကေက	kɛk	kek, kɛk	কেক
ကုက	ကုက	kuk	kuk,	কূক
	ကောက		kok	কোক
ကောက	ကောက	kɔk	kɔk	কক
ကက	ကက	kɯk	kək, kük	কৌক

Table 30: Assamese letters used for writing non initial Tai vowels

Table (31) Lists the Assamese symbols used for writing combined characters in Tai:

Aiton Symbol	Phake Symbol	Transcription in Aiton	Transcription in Phake	Assamese character
ဢ	ဢ	kai, kaai	kai, kāi	কাই
ဢ	ဢ	kam	kam	কাম
ဢ	ဢ	kaɯ	kaü	কাও

Table 31: Assamese letters used for writing combined characters in Tai

This system for writing Tai language in Assamese script was used in the publication of the *Tai Aiton History* (Morey 1999b), the *Book of calling the khon* (Morey 2001a) and *Grandfather teaches grandchildren* (Morey 2001b). A similar system was adopted by Ngi Kheng Chakap and Ai Che Let Hailung when they published *The Good Way of Teaching* in 2001.

8. Syntax

8.1 Typological Profile of the Tai Languages of Assam

Chapter 8 presents an analysis of the syntax of the two closely related varieties, Tai Aiton and Tai Phake. Reference will also be made to Tai Khamyang and Tai Khamti. Together these four will be referred to as the Tai languages of Assam, or sometimes simply the Tai languages, or even just Tai.

Before proceeding to present a detailed syntactic analysis of these various Tai languages, a short overview of the morphological structure (see 8.1.1) and the syntactic structure (see 8.1.2) will be presented.

8.1.1 Morphological Structure

In common with many languages of Southeast Asia, the Tai languages are isolating in structure. Furthermore, the languages of the Tai family share a common propensity for monosyllabic morphemes (Edmondson and Solnit 1997a:7), and both 'content words' and 'function words' are generally expressed by such monosyllabic morphemes. This is demonstrated in example (1):

1) hā¹   nɛ⁴   luk⁴ ɔn¹   nɛ⁴   pɛ⁴   kün²   kā¹   trā²   nɛ⁴ ||
    time   DEF   child   DEF   win   RECIP GO   case   DEF
    'So the boy won the case.'
    (Phake Story, *Story of the kum⁴ bird*, No. (104), told by Ee Nyan Khet)

As can be seen in (1), almost all of the words are monosyllabic and there are no affixes or clitics marking functions such as person and number on verbs or case on nouns. Grammatical functions such as Tense, Aspect and Modality (see below 8.5.7) are expressed by independent words, such as *kā¹* 'go' in (1). There are, however, some cases in which the realisation of some of these TAM morphemes approaches that of postclitics (Diller 1992:24).

In fact, the Tai varieties in Assam are not exclusively isolating nor is the vocabulary exclusively monosyllabic. Compounds are commonly formed, such as *luk⁴ ɔn¹* in (1), which is formed from two elements. *luk⁴* 'child' and *ɔn¹* 'small'. Moreover there are polysyllabic morphemes which cannot be analysed as a compound of two monosyllables. This is demonstrated by the word *pɛ⁴ yã³* 'goat' in example (2) below:

2) ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦ ꩦ

    pɛ⁴ yã³   saɯŋ⁶   to²   luk⁴ se⁶   kan⁵   pa¹ nan⁴ kə⁴   mã¹   mã²

    goat   two   CLF   from   direction   there   LINK   come   COME

    ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ

    luk⁴ se⁶   kan⁵   pa¹ nai⁴   kə⁴

    from   direction   this   LINK

    'Two goats came, from this direction and that.'

    (Phake Story, *The story of the two goats*, No. (3), told by Yehom Buragohain)

Example (2) also includes two polysyllabic forms, *pa¹ nan⁴* 'there' and *pa¹ nai⁴* 'here, this', both of which include the morpheme *pa¹* that could be analysed as a proclitic. The form *pa¹* appears to be an unstressed form of the word *paa³/pã³* 'side', but in this unstressed form it never stands alone.

Even clearer examples of non-isolating morphology are seen with the use of changed tone in Phake to express negation, imperative and questions. This was already discussed above in 6.2.4.5, and will be discussed further in 8.6.

Despite these exceptions, the Tai varieties in Assam are usually isolating and thus most morphemes are also words, at least in the pretheoretical sense of 'word'. Since these are written languages, all of these elements can be written as words. For example, when speech is transcribed using the Tai script, hesitation phenomena and other particles are regarded as words, such as ꩦꩦ *aa¹ saŋ¹* and ꩦ *ai³* in (3):

3.1) ꩦ ꩦꩦ ꩦꩦ ꩦ ꩦꩦ ꩦꩦ ... ꩦꩦ ꩦ ...

    khau¹ aa¹   a¹ saŋ¹   ai³   pin²   kaa¹   ...   luk³   nii²   ...

    3Pl   PRT   HESIT   PRT   be   GO   ...   child   good   ...

    'They, um, became ...'

3.2) ꩦꩦ ꩦꩦ ꩦ ꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦꩦ ꩦ

    ɔ¹   a¹ saŋ¹   khau¹   nai³   pin²   kaa¹   luk³   kan²   pa¹ ʒiŋ²

    EXCL   HESIT   3Pl   get   be   GO   child   RECIP   girl

    'Umm, they got the (local) girls.'

    (Aiton text, *History of the Tai*, Nos. (51-52) told by Sa Cham Thoumoung)

When the *History of the Tai* was first transcribed by Bidya Thoumoung, he wrote both the forms ꩦꩦ *aa¹ saŋ¹* and ꩦ *ai³*, although we might not expect either of these words would be listed in an Aiton dictionary.

Even when a form might be argued to be approaching clitic status, as with *kaa¹* in (3.2), it is still regarded as a separate word by the Tais, both in writing and when discussing the analysis of the texts. In this thesis, none of the morphemes encountered will be treated as clitics or affixes.

8.1.2 Syntactic Structure

The Tai languages of Assam exhibit a basic constituent order of Subject Verb Object (SVO) as exemplified in (4):

4) ꩦꩦ ꩦꩦ ꩦ ꩦꩦ ꩦ ꩦ ꩦꩦ ꩦꩦ

    kon² moŋ²   khau⁶   kã¹   pam³   mai⁴   nan⁴   se⁶

    person city   3Pl   go   fell   tree   that   PRT

    'These city people went and cut that tree down.'

    (Phake Story, *Story of the foolish king*, No. (78), told by Ee Nyan Khet)

As will be discussed below in detail in section 8.4, sentences with a different constituent order very frequently occur in Tai and this is reflected in the corpus of texts which form the basis of this thesis.

The expression of noun phrases is not obligatory in any of these languages. If a participant has already been introduced, it is not necessary to express it either by a full noun phrase or by a pronoun. Therefore a sentence such as (5) would be as

grammatical as (4), provided that the agent had already been introduced into the discourse.

5) ကျ ပံ ဒ္ဓ �168 ဆး ॥

   kā¹   pam³   mai⁴   nan⁴   se⁶
   go   fell   tree   that   PRT
   '(They) cut that tree down.'
   (Phake sentence)

Where a sentence is intransitive, the single noun phrase participant is usually expressed before the verb, as in (6):

6) ကှ168 �168 ၇ှ168 ဆ

   kun²   nan³   nɔn²   suu¹
   person   that   sleep   TOWARDS
   'That person is sleeping.'
   (Aiton sentence)

Adjectives can also appear as the intransitive predicate, as in (7):

7) ပါ်ဆ္ရ္ဌ် က ဆ္ရ္ရ်ဂ် ၁8၆ ॥

   pa¹ sāu⁶   nɛ⁴   sɔp¹   phɛu⁶
   girl   DEF   beautiful   excessive
   'That girl was very beautiful.'
   (Phake Story, *Story of the foolish king*, No. (19), told by Ee Nyan Khet)

The core arguments of a predication are expressed either as noun phrases or as prepositional phrases. Noun phrases are headed by nouns or pronouns, the former of which can be postmodified in a variety of ways, such as in example (8) which shows a noun modified by an adjective and example (9) which shows a noun modified by a possessor:

8) ကှ168 ဆ္ရ္၆

   kon²   suŋ⁶
   person   tall
   'a tall person'
   (Phake sentence)

9) ပ၆ဆ၆က် ဆ္ရုက်ကှက် �168

   pāp⁴ lik⁴   luk⁴ ɔn¹   nan⁴
   book   child   that
   'the child's book'
   (Phake sentence)

Noun phrases are discussed in detail below in see 8.3.2.

As mentioned above, some arguments are marked in prepositional phrases. Animate core participants which are not agents are frequently found in prepositional phrases marked by the preposition *haŋ²*'to', as in (10).

10) ၇ှ က ဆ္ရ္၆ က ကဆ၆ စာ၆ ၄၆ဆ် ၇၆ ၆ု က ॥

   hā¹   nɛ⁴   sɔ̄⁶   nɛ⁴   kâp⁴   tüŋ¹   wɔt⁴   haŋ²   mǖ⁶   nɛ⁴
   time   DEF   tiger   DEF   bite   all   PRT   to   pig   DEF
   'The tiger ate the pig up.'
   (Phake Story, *Story of Deception*, No. (39), told by Ee Nyan Khet)

The use of the prepositional phrase in (10) might be a type of non-obligatory case marking, and is discussed further in 8.3.3 below. This is a type of dependent marking in the terms presented by Nichols (1986). In most syntactic constructions in the Tai languages, however, neither head nor dependent is marked.

The remainder of this chapter on syntax is divided as follows:

8.2   Words

This section has been titled 'words', using the term in a pre-theoretical and perhaps intuitive sense. As discussed above in 8.1, for Tai speakers in India, morphemes are words and most of these words are monosyllabic. Of the words that are not monosyllables, many are compounds of two or more monosyllabic Tai elements. There are a smaller, but not insignificant, number of borrowed words, particularly from Pali, Assamese and Burmese, and these are often polysyllabic. These will be discussed in 8.2.8 below.

8.2.1   Word classes in Tai languages

In his study of Standard Thai, Noss (1964:79) divided words into two classes, 'free lexemes' and 'bound lexemes'. Free lexemes are those which occur 'in isolation' and can occur as 'minor sentences', a category similar to the non-initiating sentences as defined by Vichin Panupong (1970). Bound lexemes are those which cannot so occur and thus include a large number of words which would not normally be regarded as bound morphemes, such as prepositions. For example, the word *thัŋ* 'reach, arrive' is included by Noss as a preposition, one of the subclasses of bound lexeme, with the meaning 'to, all the way to, reaching' in (11).

11)   กลับมา   ถึงบ้าน   เวลา สอง ทุ่ม
      klàp-maa   thัŋ-bâan   welaa sɔ̌ɔŋ tʰûm
      return-come   reaching-home   time   two   evening o'clock
      'Got back to the house at eight p.m.' (Noss 1964:149)[57]

In addition, Noss also categories *thัŋ* both as a conjunction (1964:173), another form of bound lexeme in his analysis, and as a completive verb (1964:127), which is a free lexeme.

Noss's analysis seems to suggest that native speakers of Standard Thai can compartmentalise the different senses of a word like *thัŋ*. It appears that he regards

---

[57] Only the phonetic script and translation is from Noss. The morpheme gloss and Thai script were supplied by me.

---

words like this as having polysemous senses as both a free and a bound lexeme, or perhaps even as more than one separate homophonous words. This, however, fails to take into account either the diachronic factors, by which it can be shown that *thัŋ* 'to, reaching' is related to and derived from *thัŋ* 'to reach', or the native speaker's intuition, which would normally suggest that this is a single word.

Noss' free lexemes are themselves divided into three discrete categories: isolatives, substantives and predicatives. He defined substantives as "any free lexeme which occurs as co-constituent of a predication of which it is not the predicator." (1964:88) This group includes nouns, complementives, pronouns, numerals, classifiers and demonstratives. Predicatives, including modal verbs, adjectives, transitive verbs and completive verbs are defined as "any free lexeme which occurs as a predicator" (1964:114).

There are difficulties with Noss's approach. For example, his characterisation of adjectives as a subclass of predicatives (verbals) is not uncontroversial. It is followed by Vichin Panupong (1970:125) who describes adjectives in Thai as 'intransitive adjectival verbs'.

On the other hand, Dixon (2001b) maintains that three word classes, nouns, verbs and adjectives are "implicit in the structure of each human language", and that each has "(a) a prototypical conceptual basis and (b) prototypical grammatical function(s)."

Let us examine for a moment how we might attempt to establish that a particular word was a verb in the Tai languages. Semantic, syntactic and morphological criteria might be used. Being isolating languages, there are no formal criteria to distinguish verbs from other word classes and no bound morphology peculiar to verbs. There are some features which mark verbs out from the prototypical nouns. Whereas verbs can be accompanied by TAM markers, nouns cannot. This TAM morphology is exemplified in (12), where the verb *kap³*'bite' is followed by the TAM marker *kaa¹*'GO', indicating past tense.

12) ဗ၁ၵ ၮၣ ဟၵ ၌ၠ ဗ၁ၵ ၸ ၮၯ ၮၣ ၮၵၭ ၵၭၝ

mən² aa³ paak¹ nɔi¹ mən² pai² kap³ kaa¹ kɔn³ nɯ³
3Sg open mouth PRT 3Sg go bite GO piece meat
"He opened his mouth and ate that piece of meat."

(Aiton Story, *Story of the Crow and the Fox*, No. (56), told by Ong Cham).

No noun can be followed by a TAM morpheme in this way, but adjectives in adjectival predicates can, as shown in example (13):

13) ဝ၁ၵ ၼ ၮၜၮ ဎၵၮ ၮၯၮ ၝ၁ၵ ၵၝၵ ၮၣ ။

wan² lɯŋ² cau³ phaa³ sik¹ kyaa² kam¹ pa¹ laa¹ mən² dɯt¹ kaa¹
day one king Sikkya throne 3Sg hot GO
"One day, the throne of Lord Sikkya was very hot."

(Aiton text, *History of the Aiton*, No. (5) told by Nang Wimala Thoumoung)

Syntactically, a verb is the head of a predication (for the syntax of predications, see below 8.5). In (14) where the structure is Subject - Verb, the subject nominal is *naaŋ² dɛ¹ wii¹ nai³*, and the predication is made up of a verb (*tɯn¹*) and its modifier, a TAM morpheme:

14) ၼၮ ၮၯၝ ၼ ၮၝၵ ၮၣ ။

naaŋ² dɛ¹ wii¹ nai³ tɯn¹ maa²
lady Devi this wake COME
'Queen Devi woke up.'

(Aiton Story, *The twelve questions*, No (7), told by Sa Cham Thoumoung)

In the case of both the morphological and syntactic arguments advanced in favour of a word being a verb, it could be argued that these are properties of the phrase in which the word appears, rather than of the word itself. So, for example, it could be argued that *kaa¹* in examples (12) and (13) is a marker of the whole predication, rather than of the verb.

This is akin to the key idea of construction grammar, in which the full specification of a word is not apparent until it is in a construction. Goldberg (1996:32) sums up the program of construction grammar:

"In the theory of Construction Grammar, constructions play a central role in that grammar itself is claimed to consist of a structured inventory of constructions."

Following this argumentation, a word like *dɯt¹* 'hot' in (13) only achieves its syntactic status as a predication when it is in a construction like (13). If it were in a different construction, it would have a different status.

Let us consider the word *kaaŋ³/kāŋ³* 'wide'. This is often found as a modifier of a noun, as in example (15):

15) ၮၝၵ ၼ ၮၵ

toŋ⁵ nā² kāŋ³
area ricefield wide
'a wide paddy-field'
(Phake sentence, Banchob 1987:9)

Example (15) demonstrates *kaaŋ³/kāŋ³* in one of the prototypical functions of an adjective, that of an attribute of a noun. In example (16), on the other hand, *kaaŋ³/kāŋ³* appears as a noun:

16) ၮၵ ဗ၁ၵ ၮၜၮၮ ၮၝၵ ၮၝ ဗ၁ၵ ၮၜၮၮ ၮၝၵ

kaaŋ³ man² yuu¹ caa¹ naa¹ lɯŋ² ʒaau² man² yuu¹ caa¹ naa¹ lɯŋ²
wide 3Sg league one long 3Sg league one
"Its width was one league/yojana, its length was one league."

(Aiton Story, *The twelve questions*, No (118), told by Sa Cham Thoumoung)

Statistically, by a very large margin, the commonest use of ၮၵ *kaaŋ³* is as an adjective. Nevertheless, sentences of the form of (16) do occur, where it is the head of a noun phrase, as shown by the fact that it is modified by another nominal, ဗ၁ၵ *man²* 'third person singular pronoun' (see 8.3.2.1).

Had he been writing this grammar, Noss might have argued that there are two separate morphemes, *kaaŋ³* 'an adjective (verbal) meaning wide, to be wide' and *kaaŋ³* a noun meaning 'width'.

On the other hand, the native speakers of these languages certainly regard *kaaŋ³* as a single word and for this reason it will be treated in this thesis as such.

The problem of how to categorise word classes in the Tai languages is not a new one. Grierson was well aware of the problem, stating that he could not "divide the vocabulary into parts of speech" (1904:73), and adding that:

> "As to what function each word performs, that is determined partly by custom. Although, theoretically, every word may perform the function of any part of speech; in practice, such is not the case."

Grierson's statement that the function is "determined partly by custom" might be a forerunner of the approach of construction grammar, taking the view that a word does not acquire its full meaning until it is in a syntactic construction.

In this section of the thesis, words will be divided into classes on the basis of loosely semantic categories. We will discuss nominals (see 8.2.2), verbals (see 8.2.3), adjectives (see 8.2.4), prepositions (see 8.2.5), numbers and quantifiers (see 8.2.6), demonstratives (see 8.2.7), conjunctions (see 8.2.8) and isolatives (see 8.2.9). This is then followed by some remarks on morphemes which cannot stand by themselves (see 8.2.10). This section is then concluded by a discussion of borrowed words (see 8.2.11).

8.2.2 Nominals

The prototypical nominal is a noun and nouns are typically animate creatures or material objects. The class of nominals treated here includes nouns (see 8.2.2.1), and three groups of words that can stand in the place of nouns, namely pronouns (see

8.2.2.2), interrogative words (see 8.2.2.3), classifiers (see 8.2.2.4) and noun class markers (see 8.2.2.5).

8.2.2.1 Nouns

The class of nouns is probably the largest class of words in Tai. It is certainly an open class, with new words, usually either compounds or borrowed words, being added constantly.

Nouns could be divided into several classes, of which the most numerous are the common nouns (see 8.2.2.1.1) but there are also proper nouns (see 8.2.2.1.2).

8.2.2.1.1 Common Nouns

8.2.2.1.1.1 Simple monosyllabic nouns

Most nouns in Tai are monosyllabic words, many of which can be traced back to Proto-Southwestern Tai. Some examples are given in Table (1):

Tai	Phake (Banchob)	Aiton	Ahom	Standard Thai	Proto South-western tai	Meaning
က�ာ	kã²	kaa²	�naᶦ	กา	kaa A2	crow
ကိင်	kiŋ¹	kiŋ¹		กิ่ง	kiŋ B2	branch
ကိပ်	kip¹			กีบ	kiip DL2	hoof
ကုန်	kon³	kun³		กน	kon C2	bottom
ကပ်	kop¹	kup¹		กบ	kop DS2	frog
ကောင်	kauŋ²	kɔŋ²	ᶦnᶦ	กลอง	kɔɔŋ A2	drum
ကောင်	kauŋ³			กลอง	kɔɔŋ C2	pipe

Table 1: Some common nouns in Tai

8.2.2.1.1.2 Compound nouns

Compound nouns are made up of two or more elements, the most typical being a two morpheme compound, consisting of two monosyllabic words, often of proto Tai origin.

Aimya Khang Gohain (1995:53) listed a number of types of compounds. An example of each type is reproduced in Table (2):

Elements	1st word		gloss	2nd word		gloss	compound gloss
N + N	လိင်	laiŋ⁶	vehicle	လုံ	lom²	wind	motor-vehicle
N + Adj	နင်	nin²	earth	နမ်	nam²	black	coal
N + V	ေစ	cē³	paper	မိန်	men²	fly	kite
V + N	ပ	pāi¹	strike	ဖိ	phai²	fire	safety match
V + V	စင်	cāŋ⁵	know	ဖန်	phan⁴	massage	masseur

Table 2: Compounding processes after Aimya Khang[58] (1995)

Unlike in standard Thai, where the word-class of the head usually decides the word class of the whole compound, in the examples given by Aimya Khang, there are compounds headed by verbs which are nevertheless nouns, although, as discussed below in 8.2.3.1, there are also compounds headed by verbs which are verbs.

Compounding is one of the two most productive ways of forming new nouns, the other being borrowing (see 8.2.11). Compounding may also occur between a compound word and a simple monosyllabic word, as in Phake *khun⁶ hɔ⁶ kham²*, 'king' which is a compound of *khun⁶* 'prince' and *hɔ⁶ kham²*, 'golden palace', itself a compound.

---

[58] The transcriptions in this table follow Banchob (1987), rather than Aimya Khang's system.

8.2.2.1.2 Proper Nouns

Proper names, mostly the names of people and places, are usually polysyllabic, as the examples listed in Table (3), from the *Treaty between the Aitons and the Turungs*, indicates:

ဢၢႆ့ၶွင်	aai³ khɔŋ¹	name of a person
ၺီ့ၵျင်ꩦင်	ŋii² khaaŋ³ tɔŋ¹	name of a person
သၢမ်ꩆꩳၶၢင်	saam¹ ruu¹ khaaŋ³	name of a person
လၢႆꩦꩳႁꩪ်	laai¹ ruu¹ huk¹	name of a person
ပေႃ့ပေꩩꩪꩩꩳ	bɔr¹ pɔ¹ thaat¹	name of a town
ၵႃꩮꩳ	kaa¹ sɔ¹	name of a river
တꩪꩳ့	ta¹ kum¹	name of a tribe
တꩳꩦ	ta¹ rai¹	name of a tribe
ပꩴ့ꩦꩳꩪ်	pɤk¹ san³	name of a year
ꩁꩴ့ꩧꩳ	huŋ³ mut¹	name of a year

Table 3: Proper names from the *Treaty between the Aitons and the Turungs*.

The first three examples in Table (3) are native Tai names, instantly recognisable by the presence of the prefixes *aai³* 'first son', *ŋii²* 'second son' and *saam¹* 'third son'. These are arguably compounds. The fourth personal name may be that of a non-Tai person, and as with many borrowed words, the first tone is generally preferred (see 6.3.4.4).

Table (3) also includes some place names, names of other tribes and the names of the years. The names of the years are native Tai words, although the correct pronunciation of them is no longer known to most Aiton speakers. The tones given here for the names of the years were established by comparison with the List of Phake years elicited from Sam Thun Wingkyen.

The names of the various Tai groups in Assam are all polysyllabic, as *phā⁴ ke⁵*, the Phake term for themselves.

8.2.2.2 Pronouns

The Tai languages of Northeast India have a basic six pronoun system, as shown in Table (4):

Word	Phake	Aiton	Gloss (mostly from Banchob 1987)	ref in Banchob
ᵕᵕ	kau²	kau²	I	(1987: 18)
ᵕ	maü²	maɯ²	you	(1987: 311)
ᵕᵕ	mün²	man²	he, she, it	(1987: 325)
ᵕ	hau²	hau²	we	(1987: 410)
ᵕ	sü⁶	suu¹	you all	(1987: 130)
ᵕ	khau⁶	khau¹	they	(1987: 60)

Table 4: Tai pronouns

Two other first person non-singular pronouns are occasionally used. The word *tuu²/to²* is most often found as a classifier for animals, but it can be used as an exclusive first person pronoun, as in example (17):

17)

    lv⁴ pv⁵ nai⁴ sii⁶ to⁵ taŋ¹ loŋ⁶ cep¹ na¹ huɯ⁶
    therefore PRT body all   hurt how

    khan⁵ naa⁶ ta¹ au⁵ pɔk⁴
    direction PRT.QN WILL take return

'So we are distressed, and do not know how we will get (our traditions) back.'

(Khamyang text, *The future of the Khamyang language*, No. (40), spoken by Deben Chowlik.

The origin of the use of the word *tuu²/to²* to refer to the first person exclusive is not known. It may arise from a desire to be deferential to the addressee and wish to

emphasise the low status of the speaker. It probably can be used as a singular pronoun as well.

The pronoun *hā²*, the meaning of which in Shan is 'we two', also survives in some texts, as in example (18):

18)

    ā¹ sam⁶ hau² hā² yuŋ³ se⁶ ü¹
    Assam 1Pl 1Dl disordered PRT PRT

'This Assam is very disturbed at present.'

(Tai Phake Song, *Words of gladness - in honour of Stephen Morey*, No. (8)), sung by Am Saeu Khyo)

Banchob (1987:44) also points out that in Phake, the word ᵕᵕ *khɔ⁶* 'both' can be added to the singular pronouns to form dual pronouns, as in ᵕᵕ *hau² khɔ⁶* 'both of us'. The marking of duality is not obligatory and is only used if the speaker or writer has a very specific reason for wishing to indicate it.

Aimya Khang (1995:52) discusses the use of the respectful particle ᵕ *cau³*, glossed as 'RESP' to mark a respectful pronoun in Phake. This is common to all the Tai varieties, and is exemplified in (19):

19)

    kau² khaa¹ nɔ¹ bin³ tɛm³ laai² maai¹ sii¹
    1Sg am Nabin write letter   PRT

    suŋ¹ hɔt³ maa² tii² maɯ² cau³ ʒau³
    send arrive come at 2Sg RESP FINISHED

'I, Nabin, am sending a letter to you.'

(Aiton letter, No (2), written by Nabin Shyam Phalung, 1999)

The third person pronoun *khau¹/khau⁶* can also be used, generally in final position in a noun phrase, to indicate the plurality of a noun or noun phrase, as in (20):

20)   ကဲ   ၵၚ်   ၀ိုၼ်   တိုင်   ကွၼ်   ပၠဲ   ၵံ

     khai⁵   khām²   wün⁴   tüŋ¹   kon²   pā¹ wāi⁶   khau⁶
     tell   word   PRT   with   people   Pawaimukh   3Pl
     'Spoke with the people of Pawaimukh.'
       (Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to the Tai Phakes*,. No (75.1), read by Ai Che Let Hailung)

8.2.2.2.1 Reflexive pronouns

In both Phake and Aiton, a reflexive ပၠၚ် *pa¹ cau³* is formed by the combination of the respectful particle *cau³* with *pa¹*, the noun class marker for humans, exemplified in (21):

21)   ကပ်တၢၼ်   က   ၸၚ်   ပၠၚ်   ယိုဝ်   ကၠ

     kap tan   nɛ⁴   haŋ²   pa¹ cau³   yū²   kā¹
     captain   ex   (to)   himself   shoot   past
     'The captain shot himself.'
     (Phake sentence, from Kingcom 1992:310, No. 489)[59]

8.2.2.2.2 Reciprocals

In addition to the reflexive *pa¹ cau³*, a reciprocal marker *kan²*(or *kün²*) is a very common morpheme in all Tai varieties. It is glossed 'RECIP' and exemplified in (22):

22)   ဝၼ်   လိုၚ်   တၼ်   ၸေႃ   ထဝ်   ၸေ   ထဝ်   ၻ   ကၼ် ॥

     wan²   luŋ²   taŋ¹   pɔ²   thau³   me²   thau³   daa¹   kan²
     day   one   with   father old   mother old   quarrel RECIP
     'One day the old couple quarrelled.'
     (Aiton Story, *Story of the forest ghosts and the opium pipes*, No. (4), told by Mohendra Shyam).

In example (22) we can see the reciprocal marker in what would be the Object position in a transitive sentence. The literal meaning of *daa¹* is 'curse, abuse with words' and in the reciprocal form is 'curse each other', here translated as 'quarrel'.

Occasionally the reflexive and reciprocal forms are used together, as in (23):

---

23)   မိၚ်   ႁဝ်   ႁဝ်   ပၠၚ်   တ   သၠလဲ   ကိၼ်

     mvŋ²   hau²   hau²   pa¹ cau³   ta¹   sɔ¹ lai¹   kin⁵
     country   1Pl   1Pl   self   WILL   rule   RECIP
     'We will rule our country ourselves.'
     (Khamyang Text, *The second world war and its aftermath*, No. (79), told by Sa Myat Chowlik.)

The Khamyang pronunciation of the reciprocal particle differs from that of the Aiton and Phake, but Aiton and Phake informants agreed that the syntax in (23) would be acceptable to them. The implication of (23) is that we ourselves will rule each other in our own country. Once again, as with (22), the English translation does not require a overt reciprocal.

Sometimes *kan²* is not a reciprocal, as in (24):

24)   က   မိုၼ်   ကဲ   တိ   ထၼ်   တိုက်   ကၠ   တိုက်   ကၼ် ॥

     nɛ⁴   mün²   khai¹   ti⁵   than⁴   taik¹   kā¹   taik¹   kan²
     DEF   3Sg   lay egg   to   there   break   GO   break   RECIP
     'She laid her eggs there and they shattered and broke.'
     (Phake Story, *Story of the kum⁴ bird*, No. (5), told by Ee Nyan Khet)

Example (24) refers to young chicks breaking out of their shells. No reciprocality is implied; it is not being suggested that the young chicks help break each other's eggs. This pluralising use of *kan²* is also found in Standard Thai, where the cognate *kan* was defined by Haas (1964:25) as a reciprocal pronoun with the meaning "1. each other, one another, mutually, together. 2. severally (imparting a distributive sense to the verb). Often not rendered explicitly in English."

8.2.2.3 Interrogative Words

Interrogative words in Tai behave syntactically like noun phrases, and they are therefore included in the class of nominals. This is shown in examples (25.1) and (25.2) where the object of the verb *hɔm⁶* 'gather' is *ka⁻¹ saŋ⁶* or *saŋ⁶* 'what' in the question (25.1) and a noun phrase *kɔ²* 'salt' in the answer (25.2).

25 1)

ဃုၚ်	လယ	ကျိုဝ်	မ္ဍိ	တၢ	ဟ္ဍိ
yuŋ²	yɛ⁴	ka³ saŋ⁶	maü²	ta¹	hɔm⁶
mosquito	PRT	what	2Sg	will	gather

မ္ဍိ	တၢ	ကၢ	ကၢ	ဝၚ် ॥
maü²	ta¹	kā¹	kā⁴	saŋ⁶
2Sg	will	go	trade	what

"'Mosquito, what will you gather, what will you trade?'"

25.2)

ဃော	ကၟ	တၢ	ဟ္ဍိ	ဝဲ	ကၟဝ် ॥
ɔ²	kau²	ta¹	hɔm⁶	wai⁴	kə²
YES	1Sg	will	gather	KEEP	salt

"'Oh, I will gather salt.'"

(Phake Story, *The Dolphin, the Crow and the Mosquito*, Nos (15 & 16), told by Aithown Che Chakap)

Table (5) lists interrogative words and the kind of phrases that can be used to answer them:

Word	Phake	Aiton	Gloss	Answering phrase
ကျိုဝ်	ka³ saŋ⁶	ka¹ saŋ¹	what?	Noun phrase or sentence
ကၟ	kī¹	kii¹	how many?	Numeral or quantifier
ကၟၑ	kha¹ laü¹	kha¹ daɯ¹	how much?	Quantifier Phrase
ၑၟ	thaü⁶	thaɯ¹	where?	Locational Phrase or Place name
ၑၟ	naü⁶	daɯ¹, naɯ¹	which?	Demonstrative or sentence
ၑၟ၊ၑၟဝ်	nüŋ¹ hü⁶	naŋ¹ hɯ¹	how?	Sentence
ၑၟ	phaü⁶	phaɯ¹	who?	Proper Name
မ္ဍဝ်ၑၟ	mə⁵ naü⁶	mɯ² daɯ¹, mɯ² naɯ¹	when?	Temporal Phrase

Table 5: Interrogative words in Aiton and Phake

8.2.2.4 Classifiers

In the languages of the Tai family, as in many of the languages of mainland Southeast Asia, many nouns cannot be directly counted. The term 'classifier' has long been used (for example Haas 1964) to refer to the morpheme which is counted. In example (26), the noun is *muu¹*'pig', but the word *tuu²*'animal classifier' is required for pigs to be counted:

26)

ၑ	ၑၟ	တၢ
muu¹	haa³	tuu²
pig	five	CLF

'Five Pigs'

(Aiton sentence)

Aikhenvald (2000:1) discusses the various grammatical means for the categorisation of nouns and nominals, and uses the term 'classifier' as "an umbrella label for a wide range of noun categorization devices." The particular device exemplified in (26) is a numeral classifier in Aikhenvald's typology, so called to distinguish them from other classifiers. For the purpose of this study, the term classifier will continue to be used to refer to these numeral classifiers, and they will be glossed as 'CLF'. The syntax of these classifiers and the phrases in which they occur is discussed at 8.3.2.3.

Not all nouns require classifiers to be counted. Some abstract concepts, as we would expect, are never counted at all, but in addition there are some nouns which refer to material objects that do not need a classifier in order to be counted. For example, body parts are usually directly counted, as in (27), where the word *niu⁴*'fingers' is counted without a classifier.

27)

ၑၟဝ်	ၑၟ	ၑၟဝ်
mü²	ha³	niu⁴
hand	five	fingers

'A hand with five fingers'

(Tai Phake *Riddles*, No. (5), read by Ee Nyan Khet)

Large objects, such as countries, are also directly counted, as *muŋ²*'country' in (28):

28) မိုင် နာန် နောင် ဟာက် ကျာပ္ပ ထောဗော်
mɯɯ² nan³ nɔŋ³ hak³ kɔ¹ saa¹ pa¹ thɛ¹ myat¹
time that brother love Kassapa holy

မုန် ဟုင် ဖုင် ကော ချောတ် လဒ် မိုင်
mun² huŋ¹ phuŋ² kɔ³ khyɔt³ laai¹ mɯɯŋ²
holy famous power LINK preach many country

'Then the beloved brother Kassapa, the holy and powerful, preaching in many countries.'

(Tai Khamyang text, *The Parinibbana*, No. 25, read by Chaw Cha Seng)

Where classifiers are to be used, there is a choice between specific classifiers and general classifiers. The general classifier *an²* is used for most objects, but never for persons, for whom the classifier is generally *kɔ³/kɔ⁴*, and only rarely for animals, for which the classifier is *tuu²/to²*.

There are a large number of specific classifiers, the use of most of which is quite rare. The full list of classifiers for Phake identified by Banchob (1987), together with some additional classifiers found in the texts, are listed in Table (6). These are compared with a list of Aiton classifiers elicited from Bidya Thoumoung. Where there is an empty square in either the Aiton or Phake columns of Table (6), it indicates that the particular classifier has not yet been identified in the language concerned. It does not indicate that it does not occur.

Word	Phake	Aiton	Gloss
ကုန်	kon²		a classifer for human beings
ကုပ်	kup¹		a classifier for a piece of wood which is cracked open
ကော	kɔ⁴	kɔ³	a classifier for persons
ကမ်	kam²		a classifier for a spoken word
ကုတ်	kɔt¹		a classifier for the circumferance of the trunk of a tree
ကုပ်	kɔp¹		a handful, used as a classifier

		khan¹	a classifier for valuable and expensive things
	khɛp¹		a classifier for thin and flat things, mostly round
	ŋāi²	ŋaai²	a classifier for meals
		cɔŋ	a classifier for questions, points of dispute
	sīk⁴		a classifier for a torn thing, as a banana leave or a banana stem
	sī³		a classifier for a tooth or the similar things
	tāŋ⁴		a classifier for shelves
	tāp¹		a classifier for a flat and thin thing
	to²	tuu²	a classifier for animals
	təp¹		a classifier for a flat and round thing
	tɔi²		a volume of book
		thaa¹	a classifier of time
	nāŋ²		a classifier for a prince (now obsolete)
	nɛp¹		a classifier for packets
	nut¹		a classifier for cigarettes
	non³		a classifier for pieces or sticks of cut firewood
	nɔ⁶	nɔ¹	a classifier for persons (poetic)
	pāp⁴	paap³	a classifier for 'lik⁴' (a scroll of book)
	pā²	paa²	a classifier for respectable persons, as monks.
		puŋ²	a classifier for holy things
	phɛn¹		a classifier for thin things
		phuu¹, phaa¹	a classifier for persons
	phət¹		a classifier for skeins of thread, wool, or the like
	phün⁶	phun¹	a classifier for clothes and cloths
	māŋ²		a classifier for thin things, as paper, leaves of a tree, or the like
	muŋ⁴		a classifier for a brood

မုၚ်	muŋ⁵		a classifier for a distance within eye reach
မဲ	maü²	maɯ²/baɯ²	a classifier for knives
လၚ်	laŋ⁶	laŋ¹	a classifier for houses
လၚ်		laŋ¹	a classifier for drums
လၚ်	lăŋ⁶		a classifier for many kinds of things, also for human beings
လာန်	lan¹	lan¹	a classifier for leaves of books
လဳမ်	lɛm³		a classifier for small and long things, as boats, pencils, etc.
လုက်		luk³	a classifier for worlds and heavens
လူ	lū¹		a classifier for flowers
လော	lɛ²		a classifier for some amount of gunpowder used at one shot
လမ်	lam²	lam²	a classifier for long things, as bamboo
လော	lau²		a classifier for guns
ဝါန်	wān¹	waan¹	a classifier for an amount a bowl contains
ဝ	wī		a classifier for the above hands (of bananas);
ဟပ်	hăp¹		a classifier for 2 bundles of baskets of things carried by a long stick on the shoulder
ဟာ	hă¹	haa¹	a classifier for a fall of rain
ဟောဲ	hoi¹	hui¹	a classifier for round or spherical things
ဟဲ	hə²		a classifier for bunches of bananas or betel nuts
ဟောဲ	hoi⁴		a classifier for strings of things
အာန်	an²	an²	a classifier for many sorts of things, both big and small

Table 6: Classifiers in Phake and Aiton

These classifiers appear to fall into the following groups:

- The general classifier *an²*,
- Specific classifiers for classes of nouns according to animacy, such as *kɔ³/kɔ⁴* 'classifier for humans' and *tuu²/to²* 'classifier for animals',
- Special classifiers for important or respected humans and religious objects, such as *paa²/pă²*,
- Specific classifiers for classes of objects according to size and shape, such as *lam²*,
- Specific classifiers for specific objects, such as *baɯ²/maü²* 'classifier for knives'

Table 7: Types of classifiers in Tai

The grouping of classifiers in Table (7) is related to the frequency of their use. The general classifier *an²*, which appears at the top of Table (7), is much more frequent in the texts that any of the other groups of classifiers. The next most frequent classifiers are *kɔ³/kɔ⁴* 'classifier for humans' and *kɔ³/kɔ⁴* 'classifier for animals.'

Furthermore, as will be seen in examples (30) and (31) below, it is probable that among younger speakers classifiers from the groupings listed at the bottom of Table (7) are being replaced by the more general classifiers from groupings at the top of Table (7).

Semantically, most of the classifiers are classifiers of things on the basis of their shape, such as *lam²* 'classifier for long things'. Some of the classifiers have developed from nouns, like *tuu²/to²*, the original meaning of which was 'body', and which still has that meaning in some contexts. Others of the classifiers are of unknown origin.

One special classifier, recorded for Aiton, is *phuu³* 'classifier for persons' is worthy of further comment. It is found in example (29), where it is modified by the number *luŋ²* 'one', the only numeral which follows the classifier:

29) ခွတ် လုက် ဝဲ   ဖူ   လုၚ်   ဝ   ၼီဲ။

    khɔt¹ luk³ wai³ phuu³ luŋ² caai² niu²/diu²

    leave son keep CLF one male only one

    'He left one son,'

    (Aiton manuscript, *History from the time of the ancestor Chaw Tai Lung up until Sukapha*, No. (58)

This classifier is cognate with Standard Thai *phǔu*. In Standard Thai it cannot be counted by numbers, as it can in Aiton, but can be followed by a demonstrative, as in Standard Thai *phǔu nán* with the meaning 'that person'. It may be that the noun prefix *pa¹* (see 8.2.2.1.3 above) is a reduced form of this classifier.

Given the large number of specific classifiers, it would not be surprising to find speakers getting confused as to their use. In (30) below, Sa Cham Thoumoung makes a false start after the noun *mit³*, starting to say *an²* 'general classifier' and then correcting himself with the specific classifier *baɯ²*.

30) ပု   ၃၁   ၮၢ   ၯ   ၯ   ၯ   ၮ..   ၯ   ၯ ॥
     puɯ¹   naai²   a¹   haɯ³   maa²   mit³   a..   baɯ²   luŋ²
     grandfather maternal HESIT give   come   knife   *   CLF   one
     'His maternal grandfather gave him a knife.'
     (Aiton text, *History of the Tai*, No. (94) told by Sa Cham Thoumoung)

When Bidya Thoumoung, a much younger speaker, made a transcription of this story, he wrote ၮၢႆ *an²*. Among younger speakers, this use of the general classifier in place of the specific classifier seems to be quite widespread, as example (31) shows. In (31) Chaw En Lai Phalung uses the general classifier *an²* in place of the animal classifier *tuu²*.

31) ၀ၢႆ   ၯ   ၮၢ   ၮ   ၯ   ၯ   ၮ   ၯ
     wan²   luŋ²   phan²   aai³ tɔn² khau¹ luk³   aai³ tɔn² khau¹
     day   one   tribe   Aiton   3Pl   young/child Aiton   3Pl

     ၯ   ၯ   ၯ   ၮၢႆ   ၯ
     ʒɯ²   maa²   mɛn³   an²   luŋ²
     shoot   come   porcupine CLF   one
     'One day one of the young men of the Aiton shot a porcupine.'
     (Aiton Story, *How the Aitons and the Khamtis quarelled*, No. (3), told by Chaw En Lai).

8.2.2.5 Noun class markers

In addition to the numeral classifiers discussed in the previous section, there is a series of noun class markers listed in Table (8). These noun class markers precede the noun, unlike the numeral classifiers which occur postnominally.

Word	Phake	Aiton	Gloss (mostly from Banchob 1987)	ref in Banchob
ၮ	ca¹		a word which prefixes the names of insects and small animals	(1987: 94)
ၯ	pa¹	pa¹	a particle prefixed to a noun denoting a person	(1987: 244)
ၯ	pa¹		a particle prefixed to the names of fish	(1987: 245)
ၯ		pa¹	a particle prefixed to the names of vegetables	
ၯ	ma¹	ma¹	a prefix denoting fruits	(1987: 302)
ၯႆ	māk¹	māk¹	used as a prefix to the names of some kinds of fruit or something similar to that	(1987: 312)
ၯ	maiŋ²	mɛŋ²	some kinds of flying insects with wings	(1987: 327)
ၯ	la⁴	la³	a particle prefixed to some names of animals	(1987: 358)

Table 8: Noun class markers.

The syntax of these noun class markers differs. The prefix *la³/la⁴*, for example, is only found in combination with the names of animals and never stands alone as a separate morpheme. The word *mɛŋ²/maiŋ²*, on the other hand, is a full noun which can stand by itself, or be combined with a noun referring to the specific insect, as in Aiton *mɛŋ² sap¹* 'cockroach', or combined with a word that does not refer to a specific insect and may not even be a noun, as in Aiton *mɛŋ² dɛŋ²* 'a tick, literally insect-red'.

The noun class marker *pa¹* combines with adjectives and verbs to form a noun referring to a person. Example (32) demonstrates a number of these.

32) ꩠ꩞ꩻꩻ ꩻ ꩻꩻ ꩻꩻ ꩻ

lai⁴ kɔ⁴ pə⁵ nai⁴ kau²   cɔn² nüŋ⁵   tauŋ² pauŋ²   haü³
therefore   1Sg a little   remember   GIVE

ꩻꩻ ꩻꩻ ꩻ ꩻ ꩻ ꩻ ꩻ ꩻꩻ ꩻ

khɛn² kon² pa¹   hü⁴   pa¹   yaü¹ pa¹   cāŋ⁵ khau⁶
please person CL-human know CL-human big CL-human know 3Pl
'Therefore I remember the old men and the knowledgeable men.'

(Tai Phake Speech,. No (35), given by Ai Chanta)

When discussing the translation of (32), Ai Chan Ta indicated that pa¹ yaü¹ meant 'an old man', and pa¹ cāŋ⁵ 'a clever man'.

8.2.2.6 Nominalisation

Nominalisation occurs in several ways. Sometimes verbs or adjectives occur in a construction where they are the head of a noun phrase, as with kaaŋ³'wide' in example (16) above.

Another method of nominalisation is exemplified in (33), where the word tāŋ²'way' is prefixed to the word for 'to die', to create the abstract nominal tāŋ² khɔi³'death':

33) ꩻ ꩻ ꩻ ꩻꩻ ꩻ ꩻꩻ ‖

phaü⁶   yaŋ⁶   wɛ⁵   tāŋ² khɔi³ püŋ   hü⁶
who   have.NEG avoid   death   way   how
'There is no way for anybody to avoid death.'

(Phake Manuscript, Chaw Mahosatha meets Nang Pingya, No. (22), read by Ee Nyan Khet)

The general classifier an²can also be used in this way to nominalise a verbal, as in (34).

34) ꩻꩻ ꩻꩻ ꩻꩻ ꩻ ꩻꩻ ꩻ ꩻꩻ ꩻ ‖

hā¹ nai⁴   tauŋ⁴ mai³   an²   ta¹   kin²   ma¹   yaŋ⁶   yɛ⁴
at that time hungry   CLF   will eat   NEG   have.NEG PRT
'At that time they were hungry, but there was nothing to eat.'

(Phake Story, The two brothers, No. (4), told by Ee Nyan Khet)

The phrase an² ta¹ kin²is a noun phrase because in sentences of the form X ma¹ yaŋ⁶. the element X is a noun phrase.

8.2.3 Verbals

Verbals will be dicussed under the headings of verbs (see 8.2.3.1), auxiliary verbs (see 8.2.3.2) and completive verbs (see 8.2.3.3).

8.2.3.1 Verbs

8.2.3.1.1 Monosyllabic verbs

Most verbs in the Tai languages are monosyllabic, as those exemplified in Table (9):

Tai	Phake (Banchob)	Aiton	Ahom	Standard Thai	Proto South-western tai	Meaning
ꩻ	phat¹	phat¹	ꩻ			to read
ꩻ	phat¹			ꩻ	fat DS1	to winnow
ꩻ	phat⁴	phat³	ꩻ	ꩻ	fat DS4	to whip
ꩻ	phan²	phan²	ꩻ	ꩻ	fan A4	to cut
ꩻ	phan⁴					to massage
ꩻ	phaan¹		ꩻ	ꩻ	faan A1	to peel, slice

Table 9: Some simple verbs in Tai

The semantics of verbs in Tai would be the subject for a thesis in its own right. One example will be presented to raise this issue. Nabin Shyam Phalung explained that in Aiton different words for 'cut' would be used, according to the implement being used. This is detailed in Table (10):

Aiton Verb 'cut'		Cutting implement		
ꩻ	ŋɯt³	ꩻ	daap¹	'sword'
ꩻ	pat¹	ꩻ	phaa³	'cleaver'
ꩻ	tɛp¹	ꩻ	mit³	'knife'
ꩻ	sɔi²	ꩻ	phaan¹	'slicer'

Table 10: Terms for cutting in Aiton

8.2.3.1.2 Compound verbs

Verbs compounds are generally formed by a process that appears to be noun incorporation, as in Phake *tɔ² huk¹* 'to weave', which is literally 'weave-loom'. English speaking Phakes consistently translated *tɔ² huk¹* as the single English verb 'to weave'.

When Aithown Che was asked how to negate this verb, however, he offered example (35), which suggests that *tɔ² huk¹* is not a true compound and not a case of noun incorporation.

35)   မိုၣ် ႁုၵ် မ ေတာ့
     mün² huk¹ ma¹ tɔ⁶
     3Sg loom NEG weave.NEG
     'She didn't/doesn't weave.'
     Audio Link for this sentence (\wave files\phake\extracts\Aithown Che\1.mp3)

However, Aithown Che did accept that (36) was a possible, although not preferred, way of speaking:

36) မိုၣ် မ ေတာ့ႁုၵ်
     mün² ma¹ tɔ² huk¹
     3Sg NEG weave
     'She didn't weave!'
     Audio Link for this sentence (\wave files\phake\extracts\Aithown Che\2.mp3)

In example (36) *tɔ² huk¹* is arguably a compound, because the tone of *tɔ²* does not change after the negative particle *ma¹*, as it did in (35) (see above 6.2.4.5.2 for a discussion of the Phake negative tone). In (35), on the other hand, the noun is clearly not fully incorporated.

Several of these verb-noun combinations in Aiton are exemplified in Table (11):

Elements	1st word		gloss	2nd word		gloss	compound gloss
V + N	ကိပ်	kip¹	take	လၢၚ်	laai²	pattern	to make a design on cloth
V + N	ကၢပ်	kɔp³	cling	ယႃ	ɟaa³	medicine	to bind a medicinal poltice
V + N	ကိꩫ်	kin²	eat	ၡဝ်	khau³	rice	to dine

Table 11: Some V+N combinations in Aiton

Another process akin to compounding is the use of multi-verb sequences or verb serialisation (see below 8.5.8). Sometimes these serial verbs are translated in English by a single verb, but they are nor single words in the Tai languages, as in (37), where *khup³* 'kneel' and *paai³* 'pray' can be translated as a single word 'pray', but syntactically this is a case of serialisation.

37) မုꩱ်ကၢꩫ် ꩦ လၢ ꩦꩫ်ꩦꩭ လုꩠ် ꩦꩲ ꩤ ꩦ
     muŋ² kāŋ² ai³ lai³ cau³ phaa³ luŋ¹ khup³ paai³ ai³
     the earth PRT thus king big kneel pray PRT
     'So the great king (and his people) prayed.'
     (Aiton text, *History of the Aiton*, No. (4) told by Nang Wimala Thoumoung)

8.2.3.2 Auxiliary Verbs

True auxiliary verbs are those verbals that cannot stand alone as a predication but always have to be followed by another verb. In example (38), we see the auxiliary *khaü³* 'want', a verb which does not occur unless followed by another verb.

38) ဗ မ ၡဝ် �word ၡဝ် ꩩိꩫ် ။
     baa³ maa² khaɯ³ tɔŋ² khaɯ³ dai³ ŋin²
     crazy come want know want listen
     'She was crazy and wanted to know ...
     (Aiton Story, *The twelve questions*, No (8), told by Sa Cham Thoumoung)

The syntactic frame for an auxiliary verb is expressed in (39)

39) NP   Auxiliary   Verb   (TAM)   (NP)   (TAM)

In the data collected so far on the Tai languages, the TAM morpheme is never attached to the auxiliary verb, but only to the main verb of the predication.

Auxiliary verbs differ from verbs in a multi-verb or serial construction (see 8.5.8), in that the first verb in a multi-verb construction can be a main verb in other circumstances, whereas the auxiliary verb is never found as a main verb.

Banchob (1987) listed a number of Phake verbs which we can consider as auxiliaries:

Word	Phake	Gloss
ꩪꩨ	küŋ[1]	should
꧒	khaü[3]	to want
ꩫꩯ	khãn[4]	to be unwilling or not care to do, to dislike, to be lazy
ꩫꩴ	cãŋ[5]	may
ꩪꩨ	saiŋ[1]	should

Table 12: Auxiliary Verbs in Tai Phake (after Banchob 1987)

Banchob (1987) exemplified all of these, as in (40):

40)
maü[2] küŋ[1] tãn[3]
2Sg should speak
'You should speak' (Banchob 1987: 24)

These auxiliaries are also found in the text, as in (41):

41)
phãn[6] pheu[6] cãŋ[5] pen[2] ŋü[5]
poor too much may become foolish
'Poverty may cause you to become foolish.'
(Tai Phake Text, *Grandfather teaches Grandchildren*, Proverb No. (5), read by Yehom Buragohain)

The Aiton auxiliary *saɯ[3]* 'should, let', which is not found in Phake, is exemplified in (42):

42)
liŋ[3] niŋ[3] wai[3] saɯ[3] liŋ[3] niŋ[3] maa[2] ta[1] waa[2]
feed tea keep let feed tea COME will say
'Feeding people tea, giving them tea, we say.'
(Aiton text, *Conversation with Nabin Shyam*, No. (80), told by Nabin Shyam.)

Another auxiliary like morpheme is *laa[2]/lā[2]* 'should', which is exemplified in (43):

43)
an[2] khãm[2] tai[2] haü[3] pɔ[2] cãŋ[5] ɛ[1]
CLF word Tai GIVE PRT know PRT

lā[2] lɛ[1] mɔŋ[2] nam[6] nam[6]
should roam country many
'In order to know the Tai language, (you) should go to many places.'
(Tai Phake Song, *Poem in the khe[2] khyāŋ[2] style: How Stephen Morey came to the Tai Phakes*, No (16.1), read by Ai Che Let Hailung)

The auxiliary *laa[2]/lā[2]* can sometimes stand on its own. Tai informants will often say *ma[1] laa[2]* 'no need to' in response to a suggestion that something be done. None of the other auxiliaries discussed in this section have been recorded without a following main verb.

8.2.3.3 Completive Verbs

Since the Tai languages permit verb serialisations (see 8.5.8), the situation can arise in which one verb completes the action of another, as in (44) where the verb *mai[3]* 'burn' syntactically completes the action of main verbs *sɔt[1]* and *phau[1]* 'to put to the fire'.

44)
phai[2] phii[1] phai[2] kun[2] sɔt[1] phau[1] khiŋ[2] mau[1] mai[3]
fire spirit fire person burn burn body NEG burn
'The fire (lit by) the Gods and men would not burn his body.'

(Tai Khamyang text, *The Parinibbana*, No. (19), read by Chaw Cha Seng)

Since *mai³* 'burn' in (44) is negated, the action of the main is not in fact completed. Had the negation not been expressed in (44), the meaning would have been 'The fire (lit by) the Gods and men burned his body up.'

The syntactic pattern of sentences like (44) is expressed in (45):

45)   NP   V   (TAM)   (NP)   (NEG)   COMPLETIVE   (TAM)

Some verbs are only able to occupy the completive position, unlike *mai³* 'burn', which can also be a main verb. A very common general completive is *pɛ³/pɛ⁴* 'can,able', found in sentences like (46):

46)   ဃဉ် ဈ္ ကာဟ် ဧၞ ဃ ကၚ် တၞ ကိဟ် ၼၞ ဈ္ ॥
      süŋ⁶ maü² ān¹ ma¹ pɛ⁴ kau² ta¹ kin² cǎ⁵ maü²
      if   2Sg   count NEG can 1Sg will eat   to   2Sg
      "'If you cannot count them, I will eat you.'"
      (Phake Story, *Story of Deception*, No. (16), told by Ec Nyan Khet)

In Aiton, this completive verb *pɛ³* usually occurs before the main verb, as in (47)

47)   ကၚ် ကၞိဟ် ဃဃ ၐ္ဈ ဃသ် ၓ   ॥
      kau² uŋ¹ pɛ³ hauṣ³ saam¹ pii²
      1Sg   NEG can give   three years
      'I cannot give them three years.'
      (Aiton Story, *The twelve questions*, No (20), told by Sa Cham Thoumoung)

This divergence in constituent order between Aiton and Phake is one of the syntactic distinguishing features between the two varieties. It will be discussed further in 8.3.3.1.

Further research is needed to identify all the verbs which can act as completives and in what circumstances.

8.2.4 Adjectives

The primary function of adjectives is to describe or mark an attribute of a noun, as in example (48), where the noun has the attribute *sɔp¹* 'beautiful':

48)   ၸၞဉ် ဃ္ဈဃ်
      pa¹ saau¹ sɔp¹
      girl   beautiful
      '(a) beautiful girl' or '(a) girl is beautiful', 'girls are beautiful'
      (Aiton Sentence)

In the Tai languages, adjectives can be both modifiers within a noun phrase (see below 8.3.2) or the heads of intransitive predications; both analyses are available for example (48). In actual speech, however, such ambiguity is rare. If (48) were part of sentence like (49), the noun phrase translation of '(a) beautiful girl' would be the only one available:

49)   ၸၞဉ် ဃ္ဈဃ် ၼၞဉ် ၼၞ ဉ္ဈ ၓ္ဈ
      pa¹ saau¹ sɔp¹ nan³ uu¹ tii² dauu¹
      girl   beautiful that live at   where
      'Where does that beautiful girl live?'
      (Aiton Sentence)

If, on the other hand, there was no other possible predication in the sentence, then the second reading of (48) becomes possible, as in (50):

50)   ၸၞဉ် ၼၞဉ် ဃ္ဈဃ် ၮၚ်
      pa¹ saau¹ nan³ sɔp¹ luŋ¹
      girl   that   beautiful big
      'That girl is very beautiful'
      (Aiton Sentence)

In the texts, if the adjective is the head of an intransitive predicate, the noun phrase is usually marked with some kind of discourse particle, as in (51):

51) ကၟ ၠ ၵႃ ဃၨၵ် �290 ဟ်
    tu²   maɯ²   ɛ¹   sɛn³   naŋ¹   kham²
    body   2Sg   PRT   beautiful   as   gold
    'Your body is as beautiful as gold.'
    (Aiton Story, *Story of the Crow and the Fox*, No. (33), told by Ong Cham).

At first glance, it may appear that adjectives are indistinguishable from verbs, and they are categorised as verbs or verbals by both Noss (1964) and Vichin Panupong (1970) (see above 8.2.1). However in this study adjectives are recognised as a separate class because they can be both intransitive predicates (which nouns cannot) and can modify nouns (which verbs cannot). Furthermore, adjectives can be part of comparative constructions (see below 8.3.2.2.1).

One of the chief arguments in favour of including adjectives with verbals would be that they can be modified by TAM markers. However, adjectives are only modified by TAM markers when they are an intransitive predicate, and not when they are modifying a noun. In example (52), the adjective *kaaŋ³* 'wide' is followed by the TAM marker *kaa¹* 'GO', which is ingressive in combination with adjectives:

52) ကႃလုၵ် ပၢႃႃ် ၮ ကၨ ကႃ လ�ဒ ကႃ ॥
    a¹ luk³   pa¹ nan³   khau¹   kaaŋ³   kaa¹   laai¹   kaa¹
    from   there   3Pl   wide   GO   many   GO
    'From this point they increased in number.'
    (Aiton text, *History of the Aiton*, No. (13) told by Nang Wimala Thoumoung)

In translating (52), Nang Wimala stated that ကၨကႃလၒကႃ *kaaŋ³ kaa¹ laai¹ kaa¹* meant 'increase in numbers', that is literally 'get wider and get bigger'. The ingressive meaning conveyed by *kaa¹* is quite different from its meaning in combination in verbs, where it suggests past tense, as discussed below in 8.5.7.5.

Some monosyllabic adjectives are listed in Table (13):

Tai	Phake (Banchob)	Aiton	Ahom	Proto South-western tai	Meaning
ကၢင်	kāŋ³	kaaŋ³	မၢ	kwaaŋ C2	'wide'
ယၢ	yāu²	yaau²	ဟ	yaau A4	'long'
ၼၢင်	naiŋ²	dɛŋ²	ၼၢ	dɛɛŋ A3	'red'
လၢင်	ləŋ⁶	luŋ¹	ဟၵၢ	luaŋ A1	'yellow'
ၠ	maü¹	maɯ¹	ဟ	maɯ B1	'new'
ၵ	thau³	thau³	ၸ	thau C1	'old'
ၼ	nī²	dii²	ၸ	dii A3	'good'

Table 13: Some monosyllabic adjectives in Tai

The syntax of adjectives is further discussed below in 8.3.2.2.

8.2.4.1 Adjectival modifiers

Another point of divergence between adjectives and verbs are the adjectival post-modifiers. Adjectives can be followed by a modifier which intensifies the meaning, as in (53):

53) ၼ လုင် ၼ လုင်
    dii²   luŋ¹   dii²   luŋ¹
    good   big   good   big
    'Very good, very good'
    (Aiton sentence)

The word *luŋ¹* is itself an adjective, but in (53) acts as an intensifier/modifier. It can be termed a general modifier, since it seems to be able to be used with many adjectives.

Many adjectives have specific modifiers, which also intensify the meaning and follow the adjective, in the same position as *luŋ¹* in (53). Example (54), from Banchob, shows the use of such a modifier. Specific modifiers are very often either fully or partially reduplicated:

54) ယ ယ္ခုင္ဆင္လာက္

yam² yük⁴ yāk⁴
wet   modifier
'very wet'
(Phake sentence, Banchob 1987:350)

These specific modifiers are also found in the texts, as in (55):

55) ဟ္ခုင္လ္ဍ ၍ (ကလင္) ဒင္ဆင္ ကူင္ ॥

khum² kɔn¹ waa² aa¹ saŋ¹ dap¹ siŋ¹ kui² kui²
again say HESIT   dark   very dark
'Again I say, it is very dark'
(Aiton Story, *The twelve questions*, No (85), told by Sa Cham Thoumoung)

The full list of such adjectival modifiers can be found in the English–Phake word finder based on Banchob (1987), under the entry 'modifier'.

8.2.5 Prepositions

Prepositions in the Tai languages share a number of characteristics with nominals. Many are historically nouns, such as *tii²/tĩ⁵*'at', which was originally and sometimes still behaves like a noun meaning 'place'. Prepositions can head phrases which are core constituents of a predication, and behave like noun phrases (for the syntax of these core prepositional phrases, see below 8.3.3). Noss (1964:146) regards prepositions as bound lexemes, and therefore neither verbal nor nominal.

Banchob (1987) identified a number of prepositions in Phake, and these are presented in Table (14)

Phake preposition			Ref. in Banchob
ကၟ	cā⁵	for, to, with	(1987: 95)
ကၟ	cū²	to be near, to go near (i.e. towards)	(1987: 96)
တၟ	tĩ⁵	at	(1987: 165)
တၟ	taü³	under, beneath	(1987: 170)

တၟ	tai³	near	(1987: 179)
နၟ	naü²	in, inside	(1987: 220)
လာက္	yān⁵	about, around (time)	(1987: 345)
ကာက္လင္	kan³ laŋ⁶	behind	(1987: 364)
လုက္	luk⁴	from	(1987: 374)
လုင္လင္	lop¹ laŋ⁶	behind	
လၟ	loi³	with	(1987: 384)
ဝင္	wāŋ⁵	between, during, within	(1987: 390)
ၡင္	woŋ⁵	during	(1987: 397)
ၡက္	him²	near	(1987: 411)
ကၟဝင္	ā¹ tāŋ¹ (twāŋ¹)	during	(1987: 425)

Table 14: Phake prepositions found in Banchob (1987)

8.2.6 Numbers and quantifiers

Numbers, and certain quantifiers like *kū⁵*'all, every', do not behave syntactically like adjectives, and are therefore treated separately in this analysis. Numerals cannot appear as the head of a predication as adjectives can (see 8.2.4 above). Furthermore, whereas adjectives post-modifier nouns, numbers and quantifiers premodify the classifier, or, if it is a noun that can be counted, the noun.

Therefore (56) is grammatical in Tai and (57) is not:

56) ဆၟင္ ဝၟက္
    sɔŋ² wan²
    two   day
    'Two days.'

57) * ဝၟက္ ဆၟင္
      wan² sɔŋ²
      day   two

The only exception to this is *luŋ²/nüŋ⁵* 'one', which together with the demonstratives postmodifies the classifier or noun.

8.2.6.1 Cardinal Numbers:

Table (15) lists the numbers used by the Tai

Number	Phake	Aiton		Number	Phake	Aiton	
၀, ⵏ	nüŋ⁵	luŋ²	one	၁၀	sip¹	sip¹	ten
J	sauŋ²	sɔŋ¹	two	၁၁	sip¹ et¹	sip¹ it¹	eleven
၃	sãm⁶	saam¹	three	J၀	sãu²	saau²	twenty
၄	sĩ¹	sii¹	four	၁၀၀	pãk¹	paak¹	hundred
၅	hã³	haa³	five	၁၀၀၀	heŋ⁶	hiŋ¹	a thousand
၆	hok¹	huk¹	six	၁၀၀၀၀	mün¹	mun¹	ten thousand
၇	cɛ̃t¹	cit¹	seven				
၀	pɛt¹	pɛt¹	eight				
၉	kau³	kau³	nine				

Table 15: Tai numbers

For the higher numbers, borrowed words are used in a combination with words of Tai origin. Both Aiton and Phake informants use *sɛn¹/sɛn⁶* '100,000' to mean 'many', but the Indian English term *lakh* is also used for '100,000', along with *laan³/laan⁴*, the latter of which is also used to mean 'million' as in Shan and Standard Thai.

For the higher numbers, *kuk⁴ te¹* '10 million' and *saŋ⁴ khe¹* '100 million' are borrowed from Burmese, and *crore* '10 million' from Indian English.

8.2.6.2 Ordinal Numbers

There are no ordinal numbers in Tai. Aimya Khang Gohain, in the *Elementary Tai Primer*, written for students of Ahom, but using Phake grammar, stated that: 'To mean

the order of place, the word *ⵏ* is placed before the numerals. Nowadays a Pali word is also used:" (1997:63).

This is exemplified in Table (16):

Tai		Pali		Assamese	English
အ⁰ⵏⵏ	tĩ luŋ	ບ;ထ;ⵏ	pa tha maa	প্ৰথম	first
အ⁰ⵏ	tĩ sɔŋ	ⵏⵏ⁰ⵏ	tu tii yaa	দ্বিতীয়া	second

Table 16: Ordinal numbers in Tai (after Aimya Khang 1997)

Aimya Khang's 'Tai method' of expressing ordinal numbers in Table (16) is that used in Standard Thai, and is not accepted by most by speakers of the Tai languages in Northeast India. The Pali system, however, is seen in many manuscripts, as in example (58):

58)

kau²	sɛ̃⁶	pan¹	khɔi⁵	tã² ti¹ yã⁴	tãu⁵	khɔp⁴
1Sg	PRT	whirl	meet again	third	turn	complete a circle

ho⁶	hauk¹	khiŋ²	yɔm²
head	grey hair	body	be decreased

'I have completed the third stage of my life, my hair is grey and my body is decreasing.'
(Tai Phake Text, *Grandfather teaches Grandchildren*, Intro No. (5), read by Yehom Buragohain)

8.2.6.3 Other quantifiers

A number of other quantifiers are found whose syntax is similar to the numbers. One of these is *kuu²/kũ⁵* 'every', as in example (59).

59)

kũ⁵	wan²	kũ⁵	wan²	kə⁴	khun⁶ ho⁶ kham²	nɛ⁴
every	day	every	day	LINK	king	DEF

ကၟ ကိⁿꩻ ꩻꩻꩻ ꩻꩻꩻꩻ ꩻꩻꩻ ॥
khaü³ kin² nə⁴ nok⁴ tü³ wət⁴
want   eat   flesh dove   PRT
"Every day the king wanted to eat the meat of dove."
(Phake Story, *Story of the foolish king*, No. (3), told by Ee Nyan Khet)

Like the numerals, these quantifiers pre-modify the noun rather than postmodifying it like adjectives (see 8.2.4). A fuller study of non-numeral quantifiers remains to be done.

8.2.7 Demonstratives

Demonstratives in Tai have certain similarities to both nominals and verbals. They behave like verbals in postmodifying nouns, as in example (60):

60)   ꩻꩻꩻ ꩻꩻꩻꩻꩻ ꩻꩻꩻ ꩻꩻ ꩻꩻꩻ ꩻꩻꩻ ꩻꩻꩻ ꩻꩻꩻ
pun³   luk³ ɛn¹   nan³ aa¹ luk³ kun² phaan¹ nan¹
yonder child   that HESIT child person poor   that
'Yonder, that child, that poor child.'
(Aiton Story, *The twelve questions*, No (56), told by Sa Cham Thoumoung)

There are some features of demonstratives which are closer to nominals. In (61) the demonstrative follows the word *nüŋ¹* 'like, as', a position which is normally occupied by a nominal:

61)   ꩻꩻ ꩻꩻ ꩻꩻ ꩻꩻꩻ ꩻꩻꩻꩻ ꩻꩻꩻ
phī⁶ nam³ kā¹ het¹ pha¹ nüŋ¹ nan⁴
spirit ancestor go   do   like   that

ꩻꩻ ꩻꩻꩻ ꩻꩻ ॥
wā⁵ nüŋ¹ nai⁴
say   like   this
'"The spirit ancestors will do it in that way", She spoke like that'
(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (30), told by Ee Nyan Khet)

There are three positions of deixis in Tai demonstratives: 'proximate', 'distant' and 'far distant'. The first two are arranged in pairs of /-ai³/⁴/ 'proximate' and /-an³/⁴/ 'distant'. These are listed in Table (17):

Tai word	Phake	Aiton	gloss
ꩻꩻꩻ	than³	than³	there
ꩻꩻ	thai³	thai³	here
ꩻꩻꩻ	nan⁴	nan³	that
ꩻꩻ	nai⁴	nai³	this
ꩻꩻꩻ	lan⁴		that
ꩻꩻ	lai⁴		this

Table 17: Pairs of demonstrative words in Tai Aiton and Tai Phake

In addition there is a word with far distant deixis, namely *pun³/pon⁴* 'over there'. Its syntax is somewhat limited in comparison to *nai³/⁴* and *nan³/⁴*, in that it does not appear in combination with *naŋ¹/nüŋ¹* 'like'.

The definite particle, *nɛ³* in Aiton and *nɛ⁴* in Phake, is among the commonest morphemes in the language and appears to be derived from the demonstratives. Its use is exemplified in (62):

62)   ꩻꩻ ꩻꩻꩻ ꩻꩻꩻꩻ ꩻꩻ ꩻꩻ ॥
mə⁵ nan⁴ a¹ māt¹ nɛ⁴ wā⁵
time that   minister DEF   say
"Then the minister said."
(Phake Story, *Story of the foolish king*, No. (26), told by Ee Nyan Khet)

When translating this, I first wrote ꩻꩻ *nai⁴* 'this', but Ee Nyan Khet insisted on ꩻꩻ *nɛ⁴*, stating that "ꩻꩻꩻꩻꩻꩻꩻꩻꩻ *a¹ māt¹ nɛ⁴ ta¹ pen²* 'It should be *a¹ māt¹ nɛ⁴*.'"

The late Aimya Khang explained that whereas *nai⁴* and *nan⁴* differed because they were respectively proximate and distant, *nɛ⁴* was neither of these and rather placed

stress on this being a particular person or object. He said that *luk⁴ nɛ⁴* meant 'this particular child.' I have therefore glossed *nɛ⁴* as 'DEF', indicating that it marks definiteness. Foley and Van Valin (1985:287) characterised definiteness as "Speaker assumes the hearer can identify referent." In the case of example (63), *nɛ⁴* marks the word *mɔŋ²*, which the speakers knows the hearer can identify.

63) ကျုံဏ် ယၑ ဈၑ ဋိၑ် ၐ ဈ္ ကျုံဏ်

kyun²	yɛ⁴	haŋ²	mɔŋ²	nɛ⁴	wā⁵	kyun²
CLF	PRT	to	country	DEF	say	CLF

'The word *kyun²* is used for countries.'

(Tai Phake Manuscript *Explaining the meaning of words to children*, No. (4.24), read by Ee Nyan Khet)

The word *nɛ⁴* is also used to mark a noun phrase which has been preposed, as in (64), where the object *nɔ⁴ mün²* has been preposed and it, along with the subject, are both marked with *nɛ⁴*.

64)

mɔ⁴	nan⁴	pū¹	nɛ⁴	nɔ⁴	mün²	nɛ⁴
time	that	old man	PRT	meat	his	PRT

maiŋ¹	haü³	kā¹	haŋ²	khau⁶	pī⁵	nauŋ⁴	se⁶
dist.	GIVE	GO	to	3Sg	elder	younger	PRT

'Then the old man gave his meat to those two brothers.'

(Phake Story, *The two brothers*, No. (14), told by Ee Nyan Khet)

8.2.8 Conjunctions

Conjunctions in Tai are of two types, those which link constituents within a noun phrase, and those which link two predications.

8.2.8.1 Linking two nouns in a noun phrase

Two nouns of equal status can be linked with *khaa¹/khā⁶* or *ta¹ khā⁶*, as in example (65):

65)

wan²	sāu²	sauŋ⁶		mɛ⁵	ta¹ khā⁶	tī³	ta¹ khā⁶	lɛ¹	kɔ yā
day	22	February		ther and	father	and	travel Gaya

"My mother and father will [to Budhagaya on the 22nd of February."

(Phake letter, Nc (2), written by Peim Thi Gohain, 1997)

8.2.8.2 Linking two predications

When linking two predications, *khaa¹/khā⁶* is never used. Several morphemes such as *sii¹/sī⁶/se⁶*, *yɛ³/yɛ⁴* or *nɔi³/nɔi⁴* mark the end of the first predication in a complex sentence, when the second predication refers to events that occur later, and perhaps as a result of the first. This is exemplified in (66):

66)

mün²	khaü³	caü²	se⁶
3Sg	think	PRT	

mün²	yā³	an²	nüŋ⁵	au²	mā²	se⁶
3Sg	medicine	CLF	one	take	COME	PRT

mā²	haü³	haŋ²	cau³	ma¹ hɔ¹sa¹ thā¹
come	give	to	RESP	Mahosatha

mɔ²	tī¹	tī⁵	naü²	tauŋ⁴	nɛ⁴
time	stay	at	in	stomach	DEF

'When he had thought of this, he took a sprig of medicine and came and gave it to Chaw Mahosatha, who was at that time in his mother's womb.'

(Phake Story, *The birth and early life of Chaw Mahosatha*, No. (6), told by Ee Nyan Khet)

Grierson (1904) and following him, Gogoi (1994) have stated that *sii¹/se⁶* "suffixed to the verb" forms a participle, for which Gogoi gave the example in (67)[60]:

---

[60] The script and translation are after Gogoi; the phonemicisation and interlinear gloss have been added.

67) ဗက် ကျ ဆေ ပွက် မ

mün² kā¹ se⁶ pauk⁴ mā²
3Sg go PRT return come
'Having gone, he came back.'
(Phake sentences from Gogoi 1994:56)

The analysis in Grierson and Gogoi is akin to treating *sii¹/se⁶* as a clause final marker, regarding the clause that it marks as an anterior dependent or subordinate clause. If this were the case, we might expect that there would be restricted use of TAM markers in clauses that are marked by *sii¹/se⁶*. In example (68.1), however, *se⁶* co-occurs with the TAM marker *kā¹*.

68.1) ကျ တဲ ခွဲ ဆေ ကျ ॥

kā¹ tai¹ kho⁶ se⁶ kā¹
go walk bridge PRT GO
'Went to walk on the bridge.'

68.2) ကျ ထွဲ ခွဲ ကာ ခွဲ ဆေ ယယ ယွ တွဲ

kā¹ thüŋ⁶ khüŋ⁵ kāŋ² kho⁶ ye⁴ pe⁴ yā³ sauŋ⁶ to²
go reach middle middle bridge PRT goat two CLF

န ကွဲ ॥

nā¹ kün²
quarrel RECIP
'When they reached the middle of the bridge, those two goats quarrelled.'
(Phake Story, *The story of the two goats*, Nos. (4) and (5), told by Yehom Buragohain)

When asked the meaning of *sii¹/se⁶*, Tai informants sometimes answered that it is a *khām² haü³ sɔp¹* 'a word to make the language sound beautiful'. This is one of ways in which sentence final particles (see below 8.2.10.2) are characterised by the Tai informants. In this thesis, *sii¹/se⁶* is treated as a kind of conjunction and glossed as 'PRT', a term used for function morphemes whose precise function is not always clear. Further examination of the texts would be needed to adequately categorise this morpheme. The richness of the data presented here will allow this in the future.

Another very common linking word in the Tai languages is *kɔ³/kɔ⁴* which is glossed

throughout this thesis as 'LINK'. This word has cognates in all the languages of the Southwestern Tai group and has a number of functions. It can be translated as 'also', as in (69), where the father *kum⁴* bird joins all the other birds in fleeing a fire.

69) ư ခွဲ ခွဲ ကာ ခွဲ ဒ ယကာ ॥

ư³ nok⁴ kum⁴ nε⁴ kɔ⁴ pai² sa¹ nε⁴
father bird kum⁴ DEF LINK go PRT
'The father *kum⁴* bird also flew away.'
(Phake Story, *Story of the kum⁴ bird*, No. (8), told by Ee Nyan Khet)

In Phake this particle is often realised as [kə⁴], as in (69) and is now often spelled ခွဲ by the Phakes as a result. An alternative analysis of (69) is that *kɔ³/kɔ⁴* marks a topicalised phrase, signalling that the core sentence is following. This analysis would explain why in many cases it simply cannot be translated at all, as in example (59) above.

In many cases *kɔ³/kɔ⁴* marks both of the clauses which it is linking, as in example (70):

70) ကွဲ ထွဲ အာ နာ ခွဲ ကွဲ

kun² thau³ an² nan³ khuun¹ kɔn¹
person old CLF that again

ဟွဲ ကော ဆွဲ ဒ ကာ ဖာ နဲ ဆဲ ॥

hau² kɔ³ saŋ¹ daa¹ kan² phan² nai³ sii¹
1Pl LINK if quarrel RECIP like this PRT

ဟွဲ ကော တ မက် ကဲ ဒ ဆဲ ॥

hau² kɔ³ ta¹ maak¹ kai³ kai³ wai³ sii¹
1Pl LINK will rich quickly KEEP PRT
'That neighbour said, "If we quarrel like this we too will become rich very soon.'
(Aiton Story, *Story of the forest ghosts and the opium pipes*, No. (25), told by Mohendra Shyam).

In combination with interrogative words (see above 8.2.2.3), *kɔ³/kɔ⁴* has indefinite meaning, such as Phake *phaii⁶ kɔ⁴* 'whoever' or *ka³ saŋ⁶ kɔ⁴* 'whatever'. When the verb is negated, the English translation will be 'nobody' or 'nothing', as in (56):

56) ᨶᩖ ᨠᩬ ᨾᩤ ᨴᩬ.. ᨣᩢ᪶ ᨴᩬᩢ ॥

phaɯ¹ kɔ³ ma¹ tɔ.. uɯ¹ tɔŋ²
who   LINK NEG kn.. NEG know
'Nobody could understand.'

(Aiton Story, *The twelve questions*, No (15), told by Sa Cham Thoumoung)

8.2.9 Isolatives

A further class of words are those which Noss called "isolatives" in Standard Thai. This is a grouping of miscellaneous forms found throughout the languages of the world. Noss (1964:81) defined them as follows:

> "Isolatives typically occur as sole lexemic constituents of entire phonemic phrases which precede, follow or interrupt the large syntactic constructions that are their co-constituents."

In the Tai languages of Assam these are often found in phrases, for which the term isolative phrase is used. The syntax of isolative phrases is discussed below in 8.3.5.

The following sub-classes of isolatives are discussed below: Interjections and Exclamations (see 8.2.9.1), Responses (see 8.2.9.2), Vocatives (see 8.2.9.3) and Imitatives (see 8.2.9.4).

8.2.9.1 Interjections and Exclamations

Banchob (1987) reported a number of interjections and exclamations, which are listed in Table (18):

ᩈᩮ	cɛ²	Fie! an exclamation expressing disapprobation and disgust (used by both men and women)
ᩈᩮ	cɛ⁶	an exclamation expressing displeased feeling (used by women only)
ᩈᩪ	sɛ¹	exclamation expressing disgust
ᨶᩥᨧᩣ	nik⁴ ca⁴	exclamation expressing pity or sorrow (from Pali <u>anicca</u>)
ᨾᩮ	me¹	exclamation expressing surprise (used by women)
ᩉᩪᨯ	hüt¹	exclamation expressing disgust
ᩋᨠᩣᩃᩱᨿ	a¹ kā¹ lɔi¹	exclamation expressing pain
ᩋᩃᩱᨿ	a¹ lɔi¹	exclamation expressing tiredness
ᩋᩣ	ā⁶	exclamation for ease and comfort before starting speaking
ᩒ	ɔ⁶	Oh!
ᨶᩥᨷ	eu²	exclamation expressing insult

Table 18: Phake interjections and exclamations found in Banchob (1987)

In the *Phake-English-Thai Dictionary* (1987), Banchob exemplified many of these, such as (57):

57) ᩈᩮ ᨶᩱ ᨸᩥ ᨷᩮᨴ

cɛ²   nai⁴ pi¹   het¹
Fie! PRT don't do
'Don't do that'
(Phake sentence, Banchob:1987: 98)

Most of the interjections and exclamations reported by Banchob have not been encountered in the present study, nor have the shades of meaning found by her been exemplified in the texts. Exclamations are found frequently in spoken texts, as in (58):

58) ᩒ ᨧᩦ ᨾᩣ ᨧᩢᨣᩢᩣ ᨠᩮᩢ ॥

ɔ²   cii³ maa² cɔk¹ khaa¹ kau²
EXCL point come basket   1Pl
'"Point out my basket!"'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (9), told by Sa Cham Thoumoung)

8.2.9.2 Responses

Another subgroup of isolatives are called "responses" by Noss. In the Tai languages, the most frequent of these is that which is translated into English as 'yes', as in (59):

59) ꯀꯣꯛ ꯀꯥꯂꯥꯏ ꯂꯣꯛꯍꯨꯛ ꯀꯣ ꯀꯤ ꯀꯤꯑꯣꯛ ꯂꯣꯛ ။
caɯ² aa¹ saŋ¹ luk³ khui¹ maɯ² phii¹ thɯm¹ lau²
YES HESIT son in law 2Sg spirit forest tell
'"Yes, your son-in-law is a forest spirit", she said.'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (39), told by Sa Cham Thoumoung)

Three such responses are listed in Banchob (1987) and are presented here as Table (19):

ꯀꯣ	caɯ⁵	yes
ꯀꯤ	ɔ²	yes
ꯀꯣ	ɔ²	yes, sir, used to express affirmation in answer to the monks and the chief's questions

Table 19: Responses found in Banchob (1987)

8.2.9.3 Vocatives

In addition to interjections, which tend to precede the utterance with which they are connected, there are vocatives, which can both precede or follow the utterance, as in the following examples. The vocative word ɯi²/ɔi² follows the name of the person being called.

60) ꯃꯦ ꯊꯧ ꯀꯤ ꯃꯧ ꯌꯥ ꯋꯤꯟ ꯅꯝ ꯇꯥꯏ ။
me² thau³ ɯi² maɯ² yaa¹ wiŋ¹ nam³ taai²
mother old VOC you don't jump water to die
'Old mother, don't jump into the water and die.'

(Aiton Story, *Story of the old woman*, No. (17), told by Nabin Shyam).

Table (20) lists the vocatives found in Phake by Banchob (1987):

ꯀ	hɔi²	a vocative particle, used for calling the inferior persons or animals, such as dogs, bulls or buffaloes
ꯀ	ɔi²	a vocative particle, following the name or the personal pronoun

Table 20: Phake vocatives found in Banchob 1987

ꯀꯣꯛ ɯi²/ɔi² which is also written as a special symbol, variously ꯀ or ꯀ, has a very special use in manuscripts. At the end of important sections of manuscripts, a three or four word phrase often commencing with cau³ 'RESP' and concluding with ɯi²/ɔi². One of its functions is to draw back the attention of the listener, remembering that in former times these manuscripts would have been intoned. It is exemplified in (61):

61) ꯅꯨꯝ ꯍꯨꯛ ꯋꯟ ꯀꯧ ꯃꯥ ꯀꯧ ꯇꯨꯟ ꯍꯧ ꯍꯣ
num² huk¹ wan² khau¹ maa² khau³ tun² hau² hɔ¹
month six day they come enter caste, race our PRT

ꯀꯧ ꯀꯧ ꯀꯤ ။
cau³ khau¹ ɯi²
RESP they PRT
'In the sixth month, (was) the day that they came and entered our caste, Oh everyone!'

(Aiton manuscript, *The treaty between the Aitons and the Turungs*, No. (18)

The phrase *cau khau ɯi* has developed an almost sacred meaning in revived Ahom.

8.2.9.4 Expressives or Imitatives

Noss (1964:88) defined a fourth sub-class of isolatives which are morphologically reduplications. He exemplified them as in (62):

62) kʰau dây- yin siaŋ , iɔt-iɔt
he hear sound creak
'He heard something go creak-creak'

Noss's argument that *išt-išt* is an isolative is based on the fact that in this example there is a pause (marked by a comma), which separates the imitative from the rest of the sentence.

Noss's imitatives are more commonly called 'expressives' in the languages of the world. Banchob has collected many of these for Aiton, and exemplified some of them, such as (63), in which Banchob's version of the Aiton text is given in her original Thai. The Aiton script and phonemic versions are supplied by me.

63) ကြက် ခ ကြက် မ
     ကြုက်ၣ ‌ၣ ‌ၣ ‌ၣ
     krak[1] krak[1]   mai[3]   rak[1]   maa[2]
     ONOM.   tree   break   come
     'Krak krak, the wood / tree breaks'
     (Aiton sentence after Banchob 1977)

Table (21) lists a number of these from among the large number recorded in the *Aiton–English–Thai Dictionary* (Banchob ms):

krük[10] krük[10]	the sound of cows walking
krɔk[3] krɔk[3]	the cackling of hen
krɔŋ[8] krɔŋ[8]	the sound of striking a pot
krɔt[1] krɔt[1]	the sound of pulling timber
krɔp[1] krɔp[1]	the sound of beating the wooden clappers

Table 21: Aiton imitatives in Banchob ms.

8.2.10 Bound Lexemes

Noss argues that there are two types of lexemes in Thai, free and bound. For Noss (1964:133), bound lexemes include any lexeme that cannot stand alone, such as prepositions or conjunctions. In this thesis, two types of words will be treated as bound, TAM words (see 8.2.10.1), and sentence particles (see 8.2.10.2) Neither of these are bound in the sense of bound affixes in agglutinating languages.

8.2.10.1 Tense / Aspect Words

Tense/Aspect/Modality (TAM) words are of two types, those which precede the verb, and those which follow it.

The three types of TAM words are those which immediately precede the verb., as with *ta[1]* and *ti[1]* 'WILL'; those which follow the verb and can be immediately attached to it, as *kaa[1]* 'GO'; and those which follow the verb but always appears in utterance final position, as *yau[4]*. The TAM particles are discussed in detail at 8.5.7.

8.2.10.2 Sentence Particles

Sentence particles occur at the end of sentences and usually cannot be translated. For Phake, Banchob has recorded a number of these and has categorised them as in Table (22):

Phake		Gloss	Ref in Banchob
ကၣ	ka[1] sa[1]	a final particle denoting a familiar request	(1987: 1)
ကၣ�070ၣ	ka[1] sa[1]	a final particle denoting uncertainty	(1987: 1)
ကၢၣ	kɔ̃[2]	a final particle denoting certainty	(1987: 5)
ကၢ	kɛ[4]	a particle for emphasizing (mostly come along with sɛ[6])	(1987: 6)
ကၢၣ	kin[2]	a final particle for emphasizing	(1987: 20)
ကၢၣ	kün[2]	a final particle denoting an indefinite sense (= it seems)	(1987: 25)
ကၢ	kɔi[4]	a final particle for emphasizing	(1987: 38)
ဃၣ	ta[4]	a final particle denoting a familiar request, used among equals in rank, interchangeable with ka[1]	(1987: 163)
ကၢၣ	tɔ[6]	a final particle denoting familiarity in request	(1987: 167)
ၣ	nã[2]	a final particle denoting asking for permission or consent	(1987: 218)
ကၢ	nɔ[6]	a final interrogative particle sometimes used in a persuasive sense	(1987: 221)
ဃၣ	ya[2]	a final particle denoting the present continuous tense	(1987: 336)

| ဃ | yā³ | a final particle denoting certainty | (1987: 338) |
| ခြွ်ခ် | əi² | a terminal particle, used at the close of poetry | (1987: 436) |

Table 22: Phake Sentence particles in Banchob (1987)

Of the particles listed in Table (22), only *ta⁴*, the hortative particle (see below 8.6.4), *nɔ⁶*, the questioning particle (see below 8.6.2.1) and *əi²* are found in the texts collected for this thesis.

Several other particles do occur in the texts, such as *kɔn¹* in (64):

64)
ကိုင်	၌ပ်ဝ	ပဋ္ဌ	ဝဌ်	ကဋ်	ကွၣ်ဂ် ॥
ɯ²	maɯ²	caɯ³ saɯ³	sau²	au²	kɔn¹
EXCL	you	let	rest	take	PRT

'You should take rest.'

(Aiton Story, *The twelve questions*, No (75), told by Sa Cham Thoumoung)

This particle is also found in Shan, where it was defined as 'verbal sign of the future often with the idea that an action or condition has still to take place, or that a condition is more intensive than what one has thought to be.'

In Phake narratives, the particle *wət⁴* is frequently found. It is thought to be a discourse particle which was originally *wā⁵ nüŋ¹ nai⁴* 'say like this'. It is exemplified in in (65):

65)
ကျၣ်ၥ	လှုၥ်လာင်	ဃ	ၝ	ၡ	ဧၥ်	ၡ	၌ပ်ဝ ॥
hā¹ nai⁴	lop¹ laŋ⁶	yɛ⁴	pī⁵	cāi²	mün²	wā⁵	wət⁴
at that time	afterwards	PRT	elder	male	3Sg	say	PRT

'After that, the elder brother said, '

(Phake Story, *The two brothers*, No. (11), told by Ee Nyan Khet)

Ee Nyan Khet explained the meaning of *wət⁴* in (66):

66)
ဃင်	၌	ၨၥ်	ၔၥ်	၌ပ်ဝ	ကာ	ကၣ	ဃင် ॥
sün¹	khai⁵	haŋ²	pɔn⁵	wət⁴	nɛ⁴	lā²	yaŋ²
if	tell	to	other	wət⁴	DEF	should have

If telling the story to others, the word *wət⁴* should be there.

(Phake sentence, spoken by Ee Nyan Khet)

Table (23) lists some of the other particles which have been recorded. A full examination of these is beyond the scope of this thesis.

Particle	Phake	Aiton	
ဃ		saa¹	
ဃကာ	sa¹ nɛ⁴	sa¹ nɛ³	
ၰာ	yɛ⁴	yɛ³	
ၵာ	lɛ⁴	lɛ³	
၌		ai³	possibly a reduced form of *wai³* 'keep'

Table 23: Sentence particles in Tai

8.2.11 Borrowed words

The Tai languages also have a great number of borrowed words. The principal sources for borrowing are Burmese (see 8.2.11.1), Pali & Sanskrit (see 8.2.11.2), Assamese (see 8.2.11.3) and English (see 8.2.11.4). In this thesis only a very cursory examination of borrowings has been possible.

A full discussion of the assignment of tone in borrowed words, discussed briefly above in 6.2.4.6 (for Phake) and 6.3.4.4 (for Aiton), is also beyond the scope of this thesis.

8.2.11.1 Burmese loans

The Burmese loan words are often monosyllabic, and cannot be immediately identified as non-Tai words. Table (24) lists all the words which Banchob (1987) identified as Burmese loan words beginning with /s/. David Bradley (pers. comm.) provided the Burmese transliterations. He indicated that there seemed to be a very

inconsistent representation of the Burmese vowels, and tones. This may suggest that the words were borrowed at different stages of the history of contact between Tai and Burmese.

Tai	Phake (Banchob)	Burmese	Burmese transliteration	Meaning
ဝဘာ	sāk[4]	သက်	hsa?55	to pay tribute to the king
ဝန်	sān[2]	သန်	than22	to be healthy
ဘိင်	siŋ[2]			to attach
ဝုင်	suŋ[2]	သုန်း	thoun42	to use, spend
ဝု	sū[2]			reward, present
ဝဲ	sɔm[2]	ဆွမ်	hsun42	food offered to Buddhist monks
ဘဲတ်	süt[4]	ဆွဲ	hswe44	to move backward, forward or sideways
ဝဲန်	sɔn[2]	သွန်	thun42	to pour upon
ဝဲ	soi[2]			to rub on a slab
ဝဲ	sɔi[1]	သေး	the42	narrower, more slender

Table 24: Burmese loan words in Phake

Compounds can also be formed between Burmese borrowings, such as *maan[1]/mān[1]* 'glass' and Tai words, such as *taa[2]/tā[2]* 'eye', as in *maan[1] taa[2]/mān[1] tā[2]* 'spectacles'.

8.2.11.2 Pali Loans

Table (25) lists some Pali loans which have come through Burmese:

Tai	Phake (Banchob)	Burmese	Pali	Meaning
ဝက်ဝဲ	sat[1] ta[1] wā[2]	သတ္တဝါ	satvā	creature
ဝုဘော	sa[1] the[2]	သေဌေး	seṭṭhi	millionaire
ဘိကျ	si[1] kyā[2]	သိကြား	sakka	Indra
ဘဲန်	sün[5]	သီလ	sīla	precept, morality

Table 25: Pali loans through Burmese

It is probable that many Pali words were introduced into Tai when the Tais were converted to Buddhism, and that this influence has come through Burmese.

8.2.11.3 Assamese Loans

There are fewer Assamese loans than Pali/Sanskrit or Burmese loans listed in Banchob (1987). This does not indicate that the number of Assamese loan words is small. Rather it may indicate that Banchob's Tai informants did not wish these loan words to be included in the Dictionary, or that they regarded the use of Assamese words as cases of code-switching rather than borrowing.

Some of the Assamese loan words listed in Banchob are in Table (26):

Tai	Phake (Banchob)	Assamese	Assamese Transcription	Meaning
ဝဲ	toi[1]	দৈ	dɔi	curd, sour milk
ဘိန်	si[5] ni[2]	চনি	sɛ ni	sugar
ဝဝုလ	sɛ[5] sā[1]	চিচা	si sa	bottle
ဗုလ	poi[1] sā[1]	পইচা	pɔi sa	money in general

Table 26: Assamese loan words in Tai

Assamese loan words appear to be even more common in Aiton than amongst the Phakes. The Assamese words for the days of the week, for example, are now generally in use amongst the Aiton, whereas the Phake use the Burmese names for the seven days of the week. The original Tai system of numbering the 60 days in a cycle seems to be completely lost from everyday usage.

8.2.11.4 English Loans

Loan words from English are also gradually entering the language. They can even be used in poetry, as example (67) indicates:

67) က္ၚစ္ဝေသ္ *s00ĵ* ဝၚၵ္ ဎ္ဟ္ ဝၚၵ္ ၊

    kɔm[1] piu[1] tāt[1]   cak[1] kyā[1] tauk[1] pen[2] to[2]

    computer   machine   print be   letter

    'The computer printed the books.'

      (Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to the Tai Phakes*. No (92), read by Ai Che Let Hailung)

8.3 Constituents

8.3.1 Preliminary theoretical considerations

A sentence in the Tai languages of Assam consists of several constituents, and those constituents may themselves be made up of other constituents. The sentence constituents are:

- Noun phrases, which may include noun modifiers such as:
  - Possessor phrases,
  - Adjective phrases,
  - Classifier phrases,
  - Demonstratives
  - Relative clauses,
  - and Prepositional phrases;
- Core Prepositional phrases;
- Non core elements such as:
  - Locational phrases and
  - Temporal phrases;
- Isolative phrases;
- Predications, which may include
  - TAM marking - which may be discontinuous from the verb
  - Auxiliary verbs
  - Completive verbs - which may be discontinuous from the verb
  - Negation
- Sentence final particles

It is assumed by many scholars that in every language verbs and their objects form a single constituent, called a verb phrase. There are arguments both for and against a verb phrase being present in the Tai languages of Assam. This is discussed below in 8.4.1.2.

In the following section, noun phrases will be discussed first (8.3.2), then Core Prepositional Phrases (8.3.3), Non core elements (8.3.4) and Isolative Phrases (8.3.5). Sentence final particles were discussed above in 8.2.7.2. Predications will be discussed in a later section in 8.5.

8.3.2 The Noun Phrase:

The noun phrase in the Tai languages is a strongly head initial structure. The head of a noun phrase, which is either a noun, a pronoun or an interrogative word, is always the first element in the phrase, as in (68), where the Subject noun phrase is headed by the pronoun *kau²*, and the direct object noun phrase, *pāp⁴ lik⁴ nai⁴* 'this book', is headed by the noun *pāp⁴ lik⁴* 'book'.

68) ကၟ �002 ဝၸၵၵ္ ၵ ၵ ၵ္ၚၵ္ ၵ ၊

    kau² haü³ pāp⁴ lik⁴ nai⁴ haŋ² pa⁴ sāu⁶ nai⁴

    1Sg give   book   this   to   young woman this

    'I gave this book to the girl.'

    (Phake sentence elicited from the late Aimya Khang Gohain)

The basic structure of the Tai Noun Phrase is given in (69):

69)   Noun   Modifier

The modifier can be:

- An adjective phrase, consisting either of a single adjective or an adjective and its modifier,
- A quantifier phrase, consisting of a classifier, preceded by a number or quantifier, or followed by a demonstrative
- A possessor phrase

- A relative clause
- A demonstrative

If more than one of these modifiers is present, the most unmarked order appears to be that given in (70):

70)   Noun   Adjective   Possessor   Classifier   Relative Clause   Demonstrative

In cases of special emphasis, the order of the elements Possessor, Adjective Phrase, Classifier Phrase and Demonstratives can be altered. However any such reordering is usually marked by prosodic features such as slight pauses.

Examples (71) - (74) below demonstrate some of the possible noun phrases:

71)   Head Noun   Adjective(s)   Possessor

    mɑ̄u¹   khɑ̄u⁶   ɔn¹   sɔ² pɔ²   hau²
    young man   white   little   beautiful   1Pl
    'Our beautiful little white young man,'
    (Phake Lullaby *The little rooster*, No. (6), sung by Ee Nyan Khet)

72)   Head Noun   Classifier Phrase

    pī⁵   nauŋ⁴   sauŋ⁶   kɔ⁴
    elder   younger   two   CLF
    'Two brothers'
    (Phake Story, *The two brothers*, No. (1), told by Ee Nyan Khet)

73)   Head noun   Classifier Phrase   Demonstrative

    pa¹ sa¹ naa¹   sip¹   sɔŋ¹   cɔŋ³   nai³
    question   ten   two   CLF   this
    'The 12 questions.'
    (Aiton Story, *The twelve questions*, No (1), told by Sa Cham Thoumoung)

74)   Head Noun   Possessor   Relative Clause

khau³   suu¹   khɔŋ¹   suu¹   kaa²   an²   kun²   hau²   kin²   kaa¹
rice   you   things   you   equal   CLF   people   we   eat   GO
'Your rice and goods, which we have eaten, ...'

(Aiton manuscript, *The treaty between the Aitons and the Turungs*, No. (72)

In example (74), the head noun and possessor are combined into a reduplicated four syllable expression, or elaborate expression (see below 8.3.6). This type of expression is preferred by the Tai for aesthetic reasons. The head noun of the phrase is a compound *khau³ khɔŋ¹* 'goods', literally 'rice and things', and the possessor is *suu¹*.

As will be seen below in 8.3.2.3, sometimes the quantifier is separated from the rest of the noun phrase by other constituents. This is often called 'Quantifier Float', although as will be discussed below it may be the case that it is the head of the noun phrase that has moved and not the quantifier. Verbs can also be separated from their TAM markers by other constituents. Writing of Standard Thai, Vichin Panupong (1970) used the term discontinuous constituents for this phenomenon (see below 8.5.7.5).

The different types of modifiers within a noun phrase are discussed below, namely the possessor phrase (see 8.3.2.1), the adjective phrase (see 8.3.2.2), the classifier phrase (see 8.3.2.3), demonstratives (see 8.3.2.4) and the relative clause (see 8.3.2.5).

8.3.2.1 Possessor Phrase

In Tai, the placing of two noun phrases in apposition usually means that the second noun phrase is the possessor of the first. This does not apply when two noun phrases follow a ditransitive verb, in which case the first noun phrase will be interpreted as the theme and the second as the beneficiary (see 8.5.6).

In example (75), the noun phrase *daap¹ cau³* consists of two nouns, the first of which is the possessed, *daap¹*, 'the sword', and the second the possessor, *cau³* 'RESP', which here refers to the King.

75) ⟨Tai script⟩ ‖

pɔ²   mau¹   kɛ³   nai² dai³ daap¹ cau³ saa¹
if   NEG   answer this   get   sword RESP PRT

'If you cannot answer, you will die by the King's sword.'

(Aiton Story, *The twelve questions*, No (12), told by Sa Cham Thoumoung)

Possessive phrases are full noun phrases and can have all the constituents of a noun phrase. and can be headed by pronouns as well as nouns.

As with the other modifiers of a head noun, the possessive phrase can be separated from the head noun as in (76).

76) ⟨Tai script⟩ ‖

tɛp¹ huu¹ khaat¹   kaa¹ caa² huu¹ suuk¹ aa¹ hɔm¹ ŋii² raa² ŋii² ruɯ³ ruɯ³
cut   head separate   GO   to   commander Ahom   Ngi Ra Ngi Reu Reu

'He cut off the head of Ahom Army commander, Ngi Ra Ngi Reu Reu.'

(Aiton text, *History of the Aiton*, No. (59.1) told by Nang Wimala Thoumoung)

In example (76) the possessor phrase *huu¹ suuk¹ aa¹ hɔm¹ ŋii² raa² ŋii² ruɯ³ ruɯ³* 'the Ahom commander Ngi Ra Ngi Reu Reu', has been postposed. The head of this possessor phrase is the compound noun *huu¹ suuk¹* 'commander', itself literally 'the head of the army'. The head of the whole noun phrase is *huu¹* 'head'. It may have been to avoid having the head of the noun phrase and the head of its modifying possessor phrase (both of which are the word *huu¹* 'head') being directly adjacent, that the modifying possessor phrase is separated from its head noun[61].

It appears to be the head noun *huu¹*, rather than its possessive modifier that has moved. Example (76) is from a text for which more than one version exists, and the alternate version is given in (77).

---
[61] In view of this discussion, and the fact that the commander has been separated from his head, one is tempted to ask whether (76) a case of iconicity in syntax.

77) ⟨Tai script⟩ ‖

tɛp¹ khaat¹   kaa¹ huu¹ ŋii² raa² ŋii² ruɯ³ ruɯ³
cut   separate   GO   head Ngi Ra Ngi Reu Reu

'He cut off the head of Ngi Ra Ngi Reu Reu.'

(Aiton text, *History of the Aiton*, No. (59) told by Nang Wimala Thoumoung)

The noun phrase in (77) is *huu¹ ŋii² raa² ŋii² ruɯ³ ruɯ³* 'the head of Ngi Ra Ngi Reu Reu' and it is found after the second verb *khaat¹* 'separate' in the position where the possessor phrase was found in (76). Since the noun phrase in (77) is much shorter that that in (76) and since it does not contain a potentially problematic repetition of the word *huu¹* head, the noun phrase can comfortably remain as a single unit.

There is one major exception to the rule that where two nouns are in apposition, the second is the possessor of the first. This relates to nouns referring to minerals and other objects such as *kham²* 'gold'. When this is placed in apposition to a head noun, it is interpreted as an adjective with the meaning 'golden', as in (78):

78) ⟨Tai script⟩

phun¹ ŋum² phun¹ kham² au²   tuɯ³   maa²
rain   silver rain   gold   take fall down come

'Made silver and golden rain fall.'

(Aiton Story, *The twelve questions*, No (150), told by Sa Cham Thoumoung)

Example (79) demonstrates a possessor phrase within a possessor phrase.

79) ⟨Tai script⟩ ‖

luk⁴ me²   pān² ɔn²   mün² kɔ⁴   sip¹ pī²
child wife time lead 3Sg   LINK ten   year

'The child of his previous wife was also ten.'

(Phake Story, *The widow*, No (24), told by Aithown Che Chakap)

The noun phrase in (79) can be schematised as (80)

80)   (child ((previous wife) of him))

8.3.2.2 Adjective Phrases

The simplest Adjective phrase is just a single adjective, as in (81), where *thau³*'old' is
an attribute of *me²*'mother', and *phaan¹*'poor' is an attribute of *kun²*'person':

81)   ငယ်   ..
         me²   thau³ nan³ kun² phaan¹ ...
         mother old   that   person poor   ...
         "That old woman was a poor person...

         (Aiton Story, *Story of the old woman*, No. (2), told by Nabin Shyam).

Sometimes adjectives are themselves modified by either a general or a specific
modifier. These were discussed above in 8.2.4.1.

As shown in example (71) above, there can be more than one adjective phrase
modifying the head noun of a noun phrase.

Most attributive adjectives are capable of being the heads of intransitive predicates, as
discussed in 8.2.4 above. There are a small set of adjectives, many of them related to
states of health, which can follow the copula verb *pin²/pen²*, as in Aiton *pin² bat¹*'to
have a cold', and *pin² khai³*'to have a fever'.

8.3.2.2.1 Comparison and similarity

Comparison is expressed by an adjective phrase which is headed by a comparison
word, such as *khen²*in Phake, which precedes its adjective, or *me¹*in Aiton which
often follows its adjective. Example (82) demonstrates the use of *khen²*.

82)
     pa¹ sau⁶   nai⁴   haŋ²   luk⁴ ɔn¹   nai⁴   khɛn² suŋ⁶
     girl   this   to   child   this   COMP tall
     'The girl is taller than the boy.'
     (Phake sentence elicited from Aimya Khang Gohain)

In (82), the adjective *suŋ⁶*is analysed as the head of an intransitive predicate and the
entire predicate is *haŋ² luk⁴ ɔn¹ nai⁴ khɛn² suŋ⁶*'taller than the boy'.

This type of adjective phrase is also found with increase, as in (83):

83)
     lop¹   laŋ⁶   pī²   nüŋ⁵ maü² ta¹   khɛn² suŋ⁶ tə²   mā²
     come after back year one   2Sg will COMP tall   increase come
     'After one year, you will grow taller.'
     (Phake sentence elicited from Aimya Khang Gohain)

In Aiton, the comparative word *me¹*follows the adjective, as in *nam¹ me¹*'much-
COMP', as in example (84):

84)
     taai² naw² phaa¹ taa¹ kap¹ an²   lum² an²   taai² pɔ¹   taai² pɔ¹
     die   in   world   CLF one CLF die   born die   born

     nam³ taa² man² nai³ saŋ¹   hɔm²   wai³ haŋ² nam³
     tears   3Sg   this if   together keep to   water

     saa¹ muk¹ traa¹ kɔ³   ta¹   me¹ ta¹   nam¹ me¹ lau²
     ocean   LINK will more will much COMP say
     'In this world, if all the tears of one person through all the cycles of birth and
     death were gathered together, they would be more than the all the waters of
     the oceans, it is said.'
     (Aiton Story, *The twelve questions*, No (178), told by Sa Cham Thoumoung)

The comparative word is not necessary for the comparison to be grammatical, as in
(85):

85)
     wan² nai³ khyaam² saa¹ caa² mɯ² ŋaa²
     today   good   to   yesterday
     'Today I am better than yesterday.'
     (Aiton sentence, spoken by Nabin Shyam Phalung)

Another adjective phrase construction with syntax analogous to that of the comparatives is the expression of similarity, as in (86):

86) ကၞ�226 ပ်ဟ253 ၊၊ ရှိင် ၊ လ၊လ8င် ပ၊ဝ�’က် ॥

kɯn²	paai²	can²	naa³	heŋ³	naŋ¹	la³ liŋ²	hat¹ hat¹ kai² kai²
person	south	face	dry	like	monkey	very angry	

'The faces of the people of the south were dry like angry monkeys.'

(Aiton text, *Why Buddha was born in this world*, No. (8), told by Sa Cham Thoumoung.)

The frame for (86) is given in (87):

87)

Head noun	Adjective	Phrase modifying adjective
naa³	heŋ³	naŋ¹ la³ liŋ² hat¹ hat¹ kai² kai²
face	dry	like very angry monkeys

8.3.2.3 Quantifier Phrases:

The prototypical quantifier phrase consists of a numeral followed by a classifier that is specifically associated with the head noun, as in (88):

88) လ၃က် ၈၃ ဃ၃ ၊ကၞ ॥

luk³	caai²	saam¹	kɔ³
child	male	three	CLF

'(his) three sons,'
(Phake sentence)

The numeral *luŋ²/nɯŋ⁵* 'one' has a special syntax and can only follow the classifier, as in (89):

89) ပု လ၃င် ဝ၈ ၼ၃5 ၊ကၞ ၼ ॥

puu¹	luŋ¹	cau³	baan³	kɔ³	lɯŋ²
grandfather	big	respected	village	CLF	one

"A respected village leader,"

(Aiton manuscript, *The treaty between the Aitons and the Turungs*, No. (33)

Quantifier phrases can stand in the place of noun phrases in the syntax of Tai languages as is shown in (90):

90) ၂၆ယ်၃၃5 ဃ၃ ၼၼ5 ၃၃5 ၆ ထ၅5 ၊ကၞ

mɯ²	ʒaam²	nan³	sii¹	an²	nan³	maɯ²	thɔt¹	kaa¹
then			four	CLF	these	2Sg	take out	GO

'Then, four of these have been explained,

(Aiton Story, *The twelve questions*, No (145), told by Sa Cham Thoumoung)

In example (90), the quantifier phrase *sii¹ an²* is the object of the verb *thɔt¹* 'take out'. The sentence literally means 'At that time, you have taken out these four (arrows)', where the understood *lɛm³* 'arrow' is a metaphor for the questions that are being answered. This dropping of the head noun is only possible because it is understood from the context.[62]

Example (90) is also interesting because it contains a quantifier phrase with what appears to be the maximum number of modifiers. The full phrase is *sii¹ an² nan³* 'these four' and is of the form shown in (91), with both modifiers in place.

91) Classifier Phrase → (Numeral) Classifier (Demonstrative)

Literary manuscripts sometimes use multiple classifiers, as in (92):

92) လ၃လ၈င် ၼ၅၈ ဝ၃င် ဝၞင် ဝ၈ဝ

lai⁴	cɯŋ¹	au²	tüŋ¹	khauŋ⁶	tü²
so	take	all	thing	use	

ၼၞင် လ၃ ၊ကၞၼ ၼၝဝ ၼ ॥

cauŋ⁵	lai⁶	an²	nɔp⁴	haü³
CLF	many	CLF	offer	give

'Bring all the many utensils which were given'

(Tai Khamyang Manuscript, *The book of calling back the Khon*, No. (107), read by Sa Myat Chowlik)

---

[62] It is not known whether *maɯ² thɔt¹ kaa¹ sii¹ an²*, literally '2Sg-take out-GO-four-CLF' would be grammatical. This is a question which only occurred to me long after the last field trip for this project.

In example (92), the quantifier phrase is *cauŋ⁵ lãi⁶ an²*, which is translated as 'the many'. The head of this phrase appears to be the classifier *cauŋ⁵*, but the general classifier *an²* is also present. Interestingly the head noun of the whole NP, *khauŋ⁶ tü²* 'utensils' is premodified by *tüŋ¹* 'all'.

As mentioned above in 8.3.2, the quantifier phrase is sometimes separated from the head noun, as in (93):

93)　 သ　 မ်ရ်　ရ္ဂ　လ္က်ရ် က　က်ဉ်　တ္ဂ　ရ်ဉ်　ယ္ဂင်　တ္ဂ်　॥
　　mɛ²　mün²　wä⁵　luk⁴　nai⁴　kau²　ta¹　au²　sauŋ⁶　to²
　　wife　3sg　say　child　that　1Sg　will　take　two　CLF
　　'The wife said: "I'll take two of the children".'
（Phake Story, *Story of dogs dividing up their children*, No. (7), told by Sam Thun Wingkyen)

The separated quantifier is sometimes called 'quantifier float', but this may not be a case of floating quantifier at all. In (93) the noun phrase with the quantifier is *luk⁴ nai⁴ sauŋ⁶ to²* 'these two children (puppies)' is discontinuous. The head of the object noun phrase (*luk⁴* 'child;') and its determiner (*nai⁴* that) occur before the subject of the sentence. Assuming the unmarked word order is SVO, we could analyse this as a case of the single movement (topicalisation) of the head of the object noun phrase to clause initial position, the quantifier remaining in the unmarked postverbal position.

If we were to argue that the quantifier in examples like (93) was 'floated' or postposed, this would assume that the unmarked constituent order is SOV.

8.3.2.4 The syntax of demonstratives

Demonstratives are one of the constituents of a noun phrase identified above in 8.3.1. The demonstrative may post modify a noun phrase, or, as in (94), a classifier:

94)　ကွ်ဉ်　ှိဉ်　လ်ဉ်ဉ်　ၛ
　　kɔn³　hin¹　an²　nai³
　　lump　stone　CLF　this
　　'The lump of stone'
（Aiton sentence spoken by Sui Khong Thoumoung)
Audio Link for this sentence example (\wave files\aiton\extracts\2-1-1-3 \suikhong.mp3)

In (94) both the noun (*kɔn³ hin¹* 'stone') and the classifier (*an²*) are postmodified by a demonstrative. This is quite rare in the texts. More often, the noun may be dropped because it is understood from context, as in (95):

95)　လ်ဉ်ဉ်　ၛ　ၐ　က္ဂ　ဝ္ဉ်　လ်ဉ်ဉ်　ယ္ဂ်ဉ်　က္ဂ
　　an²　nai³　pai²　kaa¹　sii¹　an²　sut¹　kaa¹
　　CLF　this　go　GO　four　CLF　finish　GO
　　'These four questions are finished.'
（Aiton Story, *The twelve questions*, No (142), told by Sa Cham Thoumoung)

Example (95) is literally 'These are gone, these four (questions) are finished.' The head noun, *pa¹ sa¹ naa¹* 'questions' is understood.

8.3.2.5 Relative Clauses

Another of the constituents that can modify a noun phrase is a sentence. Sentences of this type will here be described as relative clauses. There does not appear to be any difference in the syntax between restrictive and non-restrictive relative clauses (see Comrie 1989:138), although this is an area that requires further research. Most of the examples presented here are non-restrictive relative clauses.

Relative clauses are often introduced by the general classifier *an²*, here acting as a relative pronoun. In (96), the head noun is *luk¹ tï⁵* 'the realm under ground', and the relative clause is at least *an² yü¹ yäk⁴*, if not the whole of the rest of the sentence

96)

woi¹	haü³	pāk¹ sāu² pɛt¹	luk¹		tï⁵	an²	yū¹
bless	give	128		realm under ground	place	CLF	stay

yāk⁴	khau⁶	khām¹	wai⁴	nət¹	mai³	peu²	au²	ŋa² lāi²
difficult	they	suffer	keep	hot	burn	flame	take	hell

'Blessings to the 128 realms under the earth, where those who live in there in difficulty suffering the heat and burning flames of hell.'

(Phake Prayer, *Prayer of Blessings*, No. (25), told by Sam Thun Wingkyen)

It is not necessary for relative clauses to be headed by a relative pronoun, as shown in (97). Here the head noun *phï⁶* 'spirit' and the adjective *hoŋ⁶ haŋ²* 'famous' have been combined into a semi-reduplicated four syllable phrase *phï⁶ hoŋ⁶ phï⁶ hāŋ²*. The remainder of the sentence can be read as a relative clause modifying this head noun and adjective.

97)

woi¹	cēm²	phï⁶	hoŋ⁶	phï⁶	hāŋ²	paü⁴	tüŋ¹	cɔi¹
bless	give	ghost	famous	ghost	EUPH	take care	all	eye tooth

sain⁶	phum⁶	kham²	thāt³		phrā²	kɛn¹ cau³
diamond	hair	gold	element, relic		Buddha	lord

'Blessings to the great and fine ghosts who take care of the diamond tooth and golden hair (of the Buddha), the relics of the lord Buddha.'

(Phake Prayer, *Prayer of Blessings*, No. (13), told by Sam Thun Wingkyen)

Spoken language rarely provides us with relative clause structures of the complexity of (96) or (97). In speech, sentences such as (98) are common:

98)

luk⁴ ɔn¹	an⁵	khuɯn³	maa²	nai⁴	khau⁶
child	CLF	increase	COME	this	3Pl

'These children who are growing up.'

(Khamyang Story, *Advice to children*, No. (31), told by Sa Myat Chowlik.)

The structure in example (98) is laid out in (99):

99)

Head noun	Relative clause	Demonstrative
luk⁴ ɔn¹	an⁵ khuɯn³ maa²	nai⁴ khau⁶
child	who are growing	these

It might be argued that the pluraliser *khau⁶* is a floated quantifier in (99). The analysis followed here is that it is modifying the demonstrative as *nai⁴ khau⁶* 'these'. This combination is found very frequently, as does *nan⁴ khau⁶* 'those'.

If the noun phrase is the object of a sentence, and is also the object of the embedded sentence which is the relative clause, it is still possible to relativise it. In (100), the relative clause is *tï⁵ khun⁶ hɔ⁶ kham² khau⁶ kā¹ khai⁵* '(to) which the king and others had told' and it is postposed and separated from its head noun *mai⁴* 'tree':

100)

kā¹	han⁶	mai⁴	loŋ⁶	nan⁴	ne⁴	koŋ²	haü³	wai⁴	yɛ⁴
go	see	tree	big	that	PRT	hollow	give	keep	PRT

tï⁵	khun⁶ hɔ⁶ kham²	khau⁶	kā¹	khai⁵
place	king	3Pl	go	tell

'They went to see the big tree with the hollow in it, to which the king (and minister) had told the story.'

(Phake Story, *Story of the foolish king*, No. (75), told by Ee Nyan Khet)

In example (100), the relative clause is headed by the noun *tï⁵*, a word which is most frequently found as a preposition meaning 'to' or 'at'. It is used here because the item being relativised (*mai⁴ loŋ⁶* 'the big tree') is regarded as a place. Example (100) is presented as Ee Nyan Khet spoke it. Interestingly, when transcribing and translating this text with her, in place of the particle *yɛ⁴* at the end of the first line, she wrote *naü² mïn²* (in-3Sg), where *mïn²* refers back to the item being relativised.

Another interesting feature of (100) is the presence of the inclusory construction *khun⁶ hɔ⁶ kham² khau⁶*, literally 'king-3Pl', which means the king and others.

8.3.3 Core Prepositional Phrases

One very interesting aspect of the Tai languages is the use of prepositional phrases for core arguments of the verb, either for the patient in transitive sentences, which will be investigated in 8.3.3.1, or for the beneficiary in three participant events 8.3.3.2 or for the experiencer in certain circumstances 8.3.3.3

8.3.3.1 Object as prepositional phrase

Examples (101) elicited from a Phake speaker, shows the object, the semantic patient, marked in a prepositional phrase, something which is not found in most other languages of the Tai family:

101) မ္ တ တာ္ ကၢ �070 �won တ
mā⁶ nɛ⁴ kap⁴ kā¹ haŋ² miu¹ nɛ⁴
dog this bite GO to cat DEF
'The dog bit the cat.'
(Phake sentence elicited from Aithown Che Chakap)

This marking of the object by *haŋ²* also occur in texts:

102) မ္ င္ ၢၢ္ မ္ တ ကၢ္ တ္ ကၢ္ ၢ င္
süŋ⁶ maü² ān¹ ma¹ pɛ⁴ kau² ta¹ kin² cā⁵ maü²
if 2Sg count NEG can 1Sg will eat to 2Sg
"'If you cannot count them, I will eat you.'"
(Phake Story, *Story of Deception*, No. (16), told by Ee Nyan Khet)

In both Tai Phake and Tai Aiton, *haŋ²* is one of two prepositions used in this way. The other, *caa²/cā⁵* is also found marking the object, as in (103):

103) တ ၢၢ္ တတ္ ၢၢ္ တ ကၢၢၢၢ
pai² au² lap¹ caa² mɛ² pɔ¹ haa¹ rii¹
go TAKE afraid to woman businessman

    တ ကၢၢၢၢ တတ္ တ
    mɛ² pɔ¹ haa¹ rii¹ lap¹ sii¹
    woman businessman afraid PRT
'They frightened the woman businessman. She was afraid'
(Aiton Story, *Story of the monkey and the fox*, No. (4), told by Bidya Thoumoung).

In this example, the verb *au²* acts as a causative, raising the number of participants from one to two for *lap¹* 'afraid'. Diller (1992:24) pointed out that this construction was often "case marked" by the preposition *haŋ²*. It may be that in this type of construction the prepositional phrase is always necessary. Further research would be needed to establish this claim.

Verbs which allow the patient to be expressed in a prepositional phrase are not confined to verbs such as 'bite' or 'frighten', but also include verbs such as 'leave', in example (104)

104) မ္ တ္တ္ မ္ တ ၢ္ တ ၢ္ တ တ ၢၢ္
mɯ² cip¹ mau¹ pɛ³ haa³ sii¹ khɔt¹ pai² caa² la³ liŋ²
hand hurt NEG can tolerate PRT leave go to monkey
'His hand hurt and he could not tolerate it so he left the monkey.'
(Aiton Story, *Story of the monkey and the fox*, No. (14), told by Bidya Thoumoung).

The marking of the object in a such a prepositional phrase is not obligatory. In *History of the Aiton*, the verb *sɛu³* 'catch' is found five times, but the object of the verb is only marked in a prepositional phrase once, in (105):

105) မ္ တ တ ၢၢ္ တ္
maa² sɛu³ caa² aai³ tɔn² khau¹
come catch to Aiton 3Pl
'Came and caught the Aitons.'
(Aiton text, *History of the Aiton*, No. (78) told by Nang Wimala Thoumoung)

There is one set of verbs for which the object is always marked in a prepositional phrase, namely verbs of speaking and asking, as in (106):

106) တ တ ၢ္ တ္ တၢ္
ta¹ thām⁶ haŋ² nāŋ² piŋ¹ ñā¹
WILL ask to lady Pingya
'(He) wanted to ask Miss Pingya ...
(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (9), told by Ee Nyan Khet)

Furthermore, there are verbs, such as *haŋ⁶* 'to see' for which neither participant is generally marked in a prepositional phrase, as in (107) below:

107) ... ...

cau³	ma¹ hɔ¹	han⁶	nāŋ²	piŋ¹ ñã¹	yɛ⁴
RESP	Mahosatha	see	lady	Ping Ya	PRT

'Chaw Mahosatha saw Miss Pingya.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (8), told by Ee Nyan Khet)

It appears that the marking of core arguments of the sentence in prepositional phrases headed by *haŋ²* or *caa²/cā⁵* is determined by two factors, namely the animacy of the phrase, and the agency. The agent of a sentence is prototypically animate; it is never marked in such a prepositional phrase. Moreover, the marking is only used for the patient when, as in the preceding examples, the patient is also animate. In example (108), the inanimate patient (*cauk¹* 'basket'), is not in a prepositional phrase.

108) ... ...

nāŋ²	piŋ¹ ñã¹	yɛ⁴	pã²	cauk¹	khau³	se⁶
lady	Pingya	PRT	carry/bring	basket	rice	PRT

'Miss Pingya was carrying a basket of rice.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (4), told by Ee Nyan Khet)

The fact that animacy is a key factor is confirmed by example (109):

109) ... ...

tɛp¹	huu¹	khaat¹	kaa¹	caa²	huu¹ suuk¹	aa¹ hom¹ ŋii² raa² ŋii² ruɯ³ ruɯ³
cut	head	separate	GO	to	commander	Ahom Ngi Ra Ngi Reu Reu

'He cut off the head of Ahom Army commander, Ngi Ra Ngi Reu Reu.'

(Aiton text, *History of the Aiton*, No. (59.1) told by Nang Wimala Thoumoung)

In (109), although it is the head of Ngi Raa Ngi Reu Reu that is cut off, it is the person, rather than his body part, which is marked in the prepositional phrase. In this sentence (also discussed above in (76)), the possessor of the noun *huu¹* 'head' is the phrase *huu¹ suuk¹ aa¹ hom¹ ŋii² raa² ŋii² ruɯ³ ruɯ³* 'the Ahom commander Ngi Ra Ngi Reu Reu'.

There are, however, sentences in which there are two non-animate noun phrases, one of which is marked in a prepositional phrase. These are definitional sentences (see 8.5.3), of which (110) is an example.

110) ... ...

haŋ²	pək¹	khau³	nɛ⁴	wā⁵	kyɛp⁴
to	skin	rice	PRT	say	chaff

'The skin of rice is called 'Chaff''.

(Phake sentence, uttered by Ee Nyan Khet)

A final example of one participant in a two participant event being marked in a prepositional phrase is (111), in which Sui Khong is explaining the meaning of a line in the manuscript relating to the new year festival of Sangkyen[63].

111) ... ...

haŋ²	saaŋ² kyɛn¹	ɛ¹	mum²	paw³	mum²	nuu²	mum²	lɛm²
to	Sangkyen	PRT	3Sg	watch	3Sg	look	3Sg	look at

'She (will) watch over and take care over Sangkyen.'

(Aiton manuscript, *Calendar for the year 2001/2002*, No. (24.2), told by Sui Khong Thoumoung.)

This example is interesting in that the other participant in this event, here expressed by the pronoun *mum²* '3Sg' is animate, whereas the participant marked in the prepositional phrase is inanimate. In the manuscript of which example (111) is an explanation, the prepositional phrase is not found, as in (112):

112) ... ...

kiŋ¹ ñaa¹	tɛ¹ wii¹	naaŋ²	phii¹	cam³	nuŋ²	khuɯŋ²	an²	khaau¹
Kinnya	Devi	lady	spirit	PRT	wear	thing	CLF	white

an²	kiŋ³	an²	luŋ¹	an²	sɔm³	an²	muu²
CLF	green	CLF	yellow	CLF	orange	CLF	purple

---

[63] Held on 14th April each year.

ꩮ ꩯ ꧞ ꧑ ꩦꩩꩫ ꩧ ။

haa³ pban² sii¹ paur³ saaŋ² kyɛn¹ yau³
five type PRT watch Sangkyen FINISHED
'As for Kingya Devi, who wears the five colours - white, green, yellow, orange and purple, she will watch over Sangkyen.'
(Aiton manuscript, *Calendar for the year 2001/2002*, No. (24), told by Sui Khong Thoumoung.)

The fact that the prepositional phrase is found only in the spoken explanation of the manuscript and not in the written form, may suggest that this widespread usage of prepositions is more frequent in spoken language than in written language. If so, given that manuscripts are generally written in very archaic style, it would suggest that this use of prepositional phrases is an innovation.

Perhaps this prepositional marking has developed because of the relatively free constituent order, and is used to distinguish one participant (usually that which is more like the patient) from the other. Where one of the participants is inanimate, it will not be the agent, and consequently no such marking is necessary - except in those special cases where both participants are inanimate - such as (110) above.

The form of this prepositional marking is possibly borrowed from the typical three participant construction (see below 8.3.3.2), where the animate beneficiary, that is to say the animate non-agent, is usually marked in this type of prepositional phrase. The early recorders of the Tai languages, using Latinate terms for describing a case system, indicated that the dative case[64] was marked by haŋ², and perhaps this indicates that the prototypical usage of this preposition is to mark the beneficiary. If so, then the sentences examined above may be an extension of the syntax used for three participant events to two participant events.

8.3.3.2 Prepositions in three participant events

The adjunct strategy for three partipant events (see 8.5.6), is the most common and therefore probably unmarked strategy for such sentences. Example (113) is an elicited

---

[64] The word *haŋ²* is never used for the allative.

sentence showing the marking of the beneficiary in the prepositional phrase headed by *haŋ²*.

113) ꩫ ꩮ ꩩꩦꩫ ꧑ ꧞ ꩧꩩ ꧑ ။
kau² haü³ pāp⁴ lik⁴ nai⁴ haŋ² pa⁴ sāu⁶ nai⁴
1Sg   give   book   this   to   young woman this
'I gave this book to the girl.'
(Phake sentence elicited from the late Aimya Khang Gohain)

The typical three participants of such events and there animacy status and status of marking in prepositional clauses is indicated below in Table (27):

Participant	Animacy status	Status of marking in prepositional phrase
Agent	animate	unmarked
Theme	inanimate	unmarked
Beneficiary	animate	marked

Table 27: Relationship between animacy, agency and the marking of core arguments in prepositional phrases.

In actual discourse all three participants are rarely expressed, although they are in (114).

114) ꩦ ꩫ ꩩ ꩮ ꩦ ꩫ ꩮ
mɔ⁴ nan⁴ pū¹ nɛ⁴ nɔ⁴ mün² nɛ⁴
time that old man PRT meat his PRT

ꩦ ꩮ ꩫ ꧞ ꩧ ꩥ ꩦ ꩦ
maiŋ¹ haü³ kɛ¹ haŋ² khau⁶ pī⁵ nauŋ⁴ se⁶
dist.   GIVE GO   to   3Sg   elder younger PRT
'Then the old man gave his meat to those two brothers.'
(Phake Story, *The two brothers*, No. (14), told by Ee Nyan Khet)

Usually, at least one of the human or animate participants is dropped because it is understood. Example (115) is a set of three sentences where the agent/subject is 'the old man', the beneficiary/indirect object is 'us', or 'me' and the theme/direct object is 'food'.

115.1) ꩬꩬꩬ ꩬ ꩬꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ

əi²	pū¹	nə⁴	maü²	nai⁴	haü³	mã²	cã⁵	kau²
VOC	grandfather	meat	2Sg	this	give	come	for/to	1Sg

"Oh! granddad, give me your meat!"

115.2) ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ

main¹	mã²	cã⁵	hau²	wã⁵	nüŋ¹	nai⁴
distribute	come	to	us	say	like	this

"'Distribute it to us!", he spoke like this'

115.3) ꩬ ꩬ ꩬ ꩬ

pū¹	nai⁴	mã⁵	haü³
grandfather	that	NEG	give

'(But), the old man didn't give (them) (the food).'

(Phake Story, *The two brothers*, Nos. (7) & (8), told by Ee Nyan Khet)

In this example, none of the three sentences fully spells out all the arguments as NPs
of the main verb, because in (115.1), the agent is a vocative. The syntax of these three
sentences is:

116) 115.1   Voc (Agent)   O (Theme)   V   cã⁵   I (Beneficiary)
      115.2   V   cã⁵   I (Beneficiary)
      115.3   S (Agent)   V

In Tai, an expression can be regarded as having three participants even if not all of
them are expressed. Example (115.2) is quite grammatical by itself and it forms part
of a three place expression. The other two places are understood, having been
introduced in (115.1).

Sometimes, there are three animate participants in a situation. Example (117) comes
from a story of a forest spirit and a mother. The spirit had stolen the mother's spade,
with which she dug for root vegetables. To get this back, she offers the spirit her
daughter, saying:

117) ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ॥

kau²	ta¹	haɯ³	luk³	saau¹	lau²
1Sg	WILL	give	child	female	say

"'I will give you my daughter", she said.'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (20), told
by Sa Cham Thoumoung)

Here there are three animate participants, the mother (agent), the daughter (theme)
and the forest spirit (recipient), the last of which is unexpressed. When discussing this
example with Bidya Thoumoung, I asked him whether (118) could have been said
here and he replied that it could.

118) ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ॥

kau²	ta¹	haɯ³	haŋ²	luk³	saau¹
1Sg	WILL	GIVE	to	child	female

'I will give you my daughter.'

It therefore appears that *haŋ²* can be used to mark an animate participant, even when it
is not a recipient/beneficiary. Example (118) could mean 'I gave it to my daughter', in
other contexts, where the theme was unexpressed, but here the context would make it
clear that the 'daughter' is the theme and it is the beneficiary that is unexpressed.

We can then perhaps categorise this use of *haŋ²* as being available to mark a non-
agent animate participant, regardless of the semantic role of that participant.

When both the theme and recipient/beneficiary are animate and both are expressed, it
is possible for both to be marked in adpositional phrases, as in (119), where the King
of Varanasi is deciding a custody dispute between the parents of young *kum⁴* chick.

119) ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ ꩬ

hã¹	nɛ⁴	yɛ⁴	haŋ²	luk⁴	nok⁴	kum⁴	nɛ⁴	haü³	wɛn⁵
time	DEF	PRT	to	child	bird	kum⁴	DEF	GIVE	discard

ꩬ ꩬ ꩬ ꩬ ॥

haŋ²	ü³	mün²	nɛ⁴
to	father	3Sg	DEF

"'The *kum⁴* chick will be given to the father".'

(Phake Story, *Story of the kum⁴ bird*, No. (64), told by Ee Nyan Khet)

When discussing the translation of this story with Ee Nyan Khet, she told me that all three participants could have been expressed as in (120):

120) ဟွက်ၚၞဟ်   ကၠ   ၜၚ်   လွက်   ဟက်   ၛၟ   ၜၚ်   ကၟ   ၝၟၝ် ။
    khun⁶ hɔ⁶ kham²   nɛ⁴   haŋ²   luk⁴   nan⁴   haü³   haŋ²   ü³   mün²
    king   DEF   to   child   that   give   to   father   3Sg
    'The king gave the chick to the father.'

The syntactic frame of (120) is given in (121):

121) Subject   Direct Object/Theme Verb   Indirect Object/Beneficiary
     khun⁶ hɔ⁶ kham² nɛ⁴   haŋ² luk⁴ nan⁴   haü³ haŋ² ü³ mün²
     the king   (to) that chick   give to its father

There appear to be several ways in which the theme and recipient are disambiguated in (120). The first is that it could never be the case that a chick could be given its father. The second is contextual, it is clear from what has gone before that it the chick is the theme. The third is syntactic; it does appear that in verbs such as 'give', the beneficiary is a more salient argument of the verb and so it is the one most likely to appear in the canonical post-verbal position. The constituent order of Tai sentences is discussed in more detail in 8.4 below.

It is not only with verbs of giving that this structure is found. The word *caa²* is used in (122), where the verb means 'to beg' and the second animate participant is not a beneficiary but a source:

122) ၝၟၠဝၵ်ကၠၚ်ၜၚ်   ၒ   ဃၟၚ်   ကၟ   ၝၟတ်   ၓတ်   ၝၟၚ်
    muɯ² ʒaam² an² nan³   pai²   ʒɔn²   kaa¹   mit³   daap¹   mən²
    then   go   beg   GO   knife   sword   3Sg

    ဝၟ   ဝၟၚ်ကၟ   ၜကၟလၚ် ॥
    caa²   cau³ ph...   ...ŋ³
    to   king   ...
    'And then s... w... and beg... for her knife from the king of Ko Lang.'
    (Aiton tex... ..., No. (62.1) told by Nang Wimala Thoumoung)

8.3.3.3 Experiencer as Adjunct:

There is another class of animate participants which are sometimes marked with *haŋ²* or *caa²/cā⁵*. Example (123) was reported by Kingcom (1992).

123) ၝၟၚ်ၜ   ၜၚ်   ၟ   ဃၟၚ်ၜ   ကၟၚ်ၜၚ်   ကၠ
    cüŋ¹ nai³   haŋ²   maü²   naŋ¹ hü⁶   an² pen²   nɛ³
    now   to   you   how   feel   PRT
    'How do you feel now?'
    (Kingcom 1992:56)[65]

A better translation might be 'How does it feel for you now?' In February 2002 a number of sentences of this type were elicited from Yehom Buragohain, including example (124), although she added that this type of sentence is rarely found in everyday speech.

124) ၜၚ်   ကၟ   ကၟတ်   လၚ် ॥
    haŋ²   kau²   kat¹   loŋ⁶
    to   1Sg   cold   big
    'I feel very cold.'
    (Phake sentence, elicited from Yehom Buragohain)

It was argued above that marking of animate participants in prepositional phrases was necessary to mark the non-agent animate participant, and this analysis is not disturbed by this example, even though there is no agent in this sentence.

A further example of a non-agent animate participant being marked in prepositional phrases is presented as (125). This is a sentence from the Khamyang speaker, Sa Myat Chowlik, whose speech often mixes Phake and Khamyang forms. This sentence has been phonemicised using Khamyang phonology:

125) ကၟၝ   ကၟၚ်   ၜ   ၟ   ကၠ   ၜ   ၛၟ   ၝၟတ်
    aa¹ muu²   an⁵   ma¹   nii⁵   nɛ⁴   ma¹   haü³   het¹
    work   CLF   NEG   good   this   not   GIVE   do

---

[65] The script and interlinear gloss have been added by me, and the phonemic transcription altered to accord with other Phake transcriptions. The original example and English translation are Kingcom's.

ꩪꩮ အꩮꩫ/ꩫꩭꩫ ꩫꩫꩫ ꩫꩫꩫ ꩫ ꩫ ꩫꩫ
haŋ² ɔn¹ an⁵ khuɯ³ maa² nai⁴ khau⁶
to children CLF increase COME these 3Pl
'The growing generation of children should not be allowed to do these wrong deeds.'

(Khamyang Story, *Advice to children*, No. (31), told by Sa Myat Chowlik.)

In (125), there is an unstated agent, perhaps 'parents' or 'the community' which may explain the use of the prepositional phrase.

In the light of these examples, it will be worthwhile to review Pongsri Lekawatana's (1970) case grammar explanation. In a study of Standard Thai using Fillmore's Case Grammar as its theoretic base, she (1970:114) defines what she calls the dative case in standard Thai as the argument which answers these three questions:

126) a) Who experiences something?
      b) To whom does something happen?
      c) Who is affected by something?

These three were exemplified by Lekawatana (1970:114), restated as (127):

127) a) sùda (D)   chɔ̂ɔp   năŋsɨ̀   lêm   nii (O)
          Suda   like   book   cl   this
          Suda likes this book.

      b) dæɛŋ (A)   hây   năŋsɨ̀ (O)   kɛ̀ɛ   sùda (D)
          Dang   give   book   to   Suda
          Dang gave a book to Suda

      c) dæɛŋ (A)   khâa   sùda (D)
          Dang   kill   Suda
          Dang killed Suda

In the Tai varities of Northeast India, we have seen the marking in prepositional phrases of the participants which Lekawatana identifies as Dative, (D) in example (127), although this marking is not obligatory.

The agent is never marked in a prepositional phrase in the Tai languages. In sentences like (123) above, the animate participant is not considered an agent and is analysed by Lekawatana as Dative, and so can carry the prepositional phrase marking.

However, although Lekawatana's explanation in (127) very neatly corresponds with the three situations in which the marking with prepositional phrases has been observed, it does not explain why such marking appears to be used more often with some verbs than with others.

If we consider again example (107), we can see a sentence in which there are two animate participants, neither of which is marked in a prepositional phrase. The first participant, *cau³ ma¹ hɔ¹* 'Chaw Mahosatha', is the experiencer, and as we have seen in (123), experiencers can be marked in a prepositional phrase; however this would certainly lead to some confusion in that the second participant, which is arguably the least agent-like, might also be able to marked in a prepositional phrase.

We are lead to the conclusion that where there are two animate participants, neither of which is an agent, the marking with a prepositional phrase does not occur.

8.3.4 Non core elements

The two non-core elements discussed below are both in the form of prepositional phrases, headed by any of the large range of prepositions, which were discussed in 8.2.5 above.

8.3.4.1 Temporal Phrases

Temporal phrases appear to be permitted both before and after the core sentence, although in story telling they are used as links between different sections and so usually occur as the first element in the sentence, as in (128):

128)

မ္	ႁ	႑	က	႑	႑	က
mə⁴	nan⁴	pū¹	nɛ⁴	nɔ⁴	mün²	nɛ⁴
time	that	old man	DEF	meat	his	PRT

မ	ႁ	က	ႁ	ႁ	ႁ	ႁ	ႁ
maiɲ¹	haü³	kā¹	baŋ²	khau⁶	pī⁵	nauŋ⁴	se⁶
dist.	give	GO	to	3Sg	elder	younger	PRT

'Then the old man gave his meat to those two brothers.'

(Phake Story, *The two brothers*, No. (14), told by Ee Nyan Khet)

When another portion of the sentence is being emphasised, as in (129), the temporal phrase may appear at the end:

129)

မ္	ႁ	က	ႁ	ႁ	ႁ	ႁ
maü²	ta¹	kā¹	kāŋ²	naü⁶	kāŋ²	kham⁵
2Sg	will	go	middle	morning	middle	evening

'Will you go in the morning or in the evening?'

(Phake sentence, elicited from the late Aimya Khang Gohain)

8.3.4.2 Locational Phrases

As with temporal phrases, locational phrases are found as adjuncts either in front of, or after, the core sentence, as in (130):

130)

က	ႁ	ႁ	က	ႁ	ႁ	ႁ	ႁ
nɛ⁴	tī⁵	khun⁴ tāŋ²	nai⁴	kyauŋ² mū²	an²	nüŋ⁵	yaŋ²
PRT	at	road	PRT	pagoda	CLF	one	is

'Along the road there is a pagoda.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (53), told by Ee Nyan Khet)

Sometimes the locational phrase ends up in the middle of the sentence, because another constituent has been topicalised, as in (131), in which the subject noun phrase, 'the king together with all the people of the country', has been topicalised and appears before the non-core locational phrase *tii² dɔi² tau¹ san³ nai³* 'at Mount Tau San'.

131)

ႁ	ႁ	ႁ	ႁ	ႁ
lai³ pɯ² nai³	kaa¹ cɯ²	cau³ phaa³	cau³	phɔi³
therefore	all	king	resp	EUPH

ႁ	ႁ	ႁ	ႁ	ႁ
kup¹ taŋ¹	kum²	suŋ³	kum²	muŋ²
with	person	country	person	city

ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ
tii²	dɔi²	tau¹ san³	nai³	ta¹	nai³	pai²	khup³	paai³	wai³
at	hill	Tau San	this	will	GET	GO	kneel	pray	pray

ႁ	ႁ	ႁ	ႁ	ႁ	ႁ
caa²	phii¹	luŋ¹	phii¹	taŋ³	kam¹ phaa³
to	ghost	big	ghost	set up	world

'The king, together with all the people of the country, went to Mount Tau San in order to pray to the great spirit who had made the world.'

(Aiton text, *History of the Aiton*, No. (3) told by Nang Wimala Thoumoung)

Often, the locational phrase is a full argument of the verb, as in (132), where it is the argument of *uu¹* 'live at':

132)

ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ	ႁ
mɯ²	cau³ phaa³	aai³ tɔn¹	uu¹	tii²	cɛ²	rɔŋ	nɔi² / dɔi²	suŋ¹
when	King	Aiton	live at	at	city	range	mountain	high

ႁ	ႁ	ႁ	ႁ	ႁ	ႁ
uu¹	tii²	him²	nam³	kaa¹ sɔ¹	nam¹ naa¹
live	at	near	river	Kaso	PRT

'Then, the King of the Aitons lived at the City of Long Doi Sung (high mountain range), near the bank of the Kaso river.'

(Aiton manuscript, *The treaty between the Aitons and the Turungs*, No. (22)

Locational phrases are not always headed by prepositions, as with *mān³ pā¹ wāi⁶* in (133). The verb *kā¹* 'go" does not require a prepositional phrase as its object, although in this case the locational phrase is not a non-core element.

133) [Tai script]
   məˀ⁵ naü⁶ maü² ta¹ kã¹ mãn³ pã¹ wãi⁶
   when 2Sg will go village Pawai
   'When will you go to Pawoimukh?'
   (sentence elicited from Aithown Che Chakap)

8.3.5 Isolative Phrases:

As discussed above in 8.2.9, the various exclamations, reponsives and vocatives are categorised as isolatives. Sometimes they are only a single word, as in (134):

134) [Tai script]
   ə² ñãn¹ nɛ⁴ hüŋ² piŋ¹ ñã¹
   YES knowledge DEF to knowledge
   'Yes, ñãn² means knowledge.'

   (Tai Phake Song, *Poem in the khe² khyãŋ² style: How Stephen Morey came to the Tai Phakes,.* No (10.3), read by Ai Che Let Hailung)

This is a definitional sentence where the verb is not stated; the thing being defined is the first noun phrase (ñãn¹ nɛ⁴) and its definition is in an adpositional phrase (piŋ¹ ñã¹). The other element of the sentence, the reponse ə²'yes' is not part of the core syntax, hence the use of the term isolative. Initial isolatives are always in absolute initial position (unless preceded by a hesitation).

Sometimes an isolative is a phrase, as in (135), which was spoken by an Aiton informant when she picked up my small back-pack and found it to be much heavier than she expected. The whole utterance could have been regarded as an exclamation phrase, had it been followed by something like 'This is really heavy'.

135) [Tai script]
   cau³ phrã² pin² cau³
   RESP Buddha be RESP
   'Oh! God!'
   (Aiton Sentence)

Example (135) is the closest to swearing that I have been able to record in the Tai languages. I have not been able to record such a phrase in combination with a predication.

In (136), there are arguably two isolative phrases, the first being an exclamation (ã²) and the second a vocative phrase pū¹ cau³ əi².

136) [Tai script]
   ã² pū¹ cau³ əi² nə⁴ maü² nɛ⁴ khɛn²
   EXCL grandfather RESP VOC meat your this please

   maiŋ¹ haü³ mã² cã⁵ hau² hau² kə⁴ ‖
   dist give COME to us us LINK
   'Oh, respected grandfather, please, give your meat to us.'
   (Phake Story, *The two brothers*, No. (12), told by Ee Nyan Khet)

This syntax of (136) is:

137)

EXCL	VOC	Object NP (topicalised)
ã²	pū¹ cau⁴ əi²	nə⁴ maü² nɛ⁴
Oh!	respected grandfather	your meat

Verb complex	Indirect Object NP	Indirect Object NP (postposed)
khɛn² maiŋ¹ haü³ mã²	cã⁵ hau²	hau² kə⁴
please give	to us	us

It is not necessary for the Isolative Phrase to be headed by an isolative, such as a vocative. The name of the person being called, or their kinship relationship, is sufficient, as in (138):

138) [Tai script]
   pə² kau² wan² nai³ pai² sən¹ lik³ mau¹ pai² ‖
   father me today go study NEG go
   'Father, I am not going to school today.'

   (Aiton Story, *The boy who wouldn't go to school*, No. (2), told by Bidya Thoumoung).

Sometimes, the Isolative Phrase is the first element in a sentence which is itself embedded, as the Isolative Phrase *mɛ²*'mother' in (139). There is a pause after the vocative, and this makes it clear that he is not asking, 'Who are my parents?'. If there were no pause, the two nouns *mɛ²* and *pɔ²* would be assumed to form a single noun phrase

139) ꩠꩠꩠ   ꩬ   ꩬꩠꩠ ꩠ   ꩬ   ꩠ   ꩠ

a¹ nan³   ʒau³   luk³   caai²   ʒaɯ¹   maa²   lau²
there   FINISHED   child   male   big   come   tell

ꩬ   ꩬꩠ   ꩠ   ꩠ   ꩠ   ॥

mɛ²   pɔ²   hau²   phaɯ¹   ne³
mother father   1Pl   who   PRT
'Then the eldest son came and asked, 'Mother, who is our father?''
(Aiton text, *History of the Tai*, No. (88) told by Sa Cham Thoumoung)

As discussed above in 8.2.9.4, for Standard Thai, Noss included a category of isolatives that he named imitatives, but which are better known as expressives. Banchob (1977) collected many of these for Aiton, such as (140) below. In this example, Banchob's version of the Aiton text is given in her original Thai. The Aiton script and phonemic versions are added by me.

140) ꩬꩠꩬꩠ   ꩠꩠ   ꩬꩠ   ꩠ   ꩠꩠꩠ

ละ² ลิ้ง²   หัน¹   สือ   ไห้   กรๅก ๆ
la³ liŋ²   han¹   sɯ¹   hai³   kraak¹ kraak¹
monkey   see   tiger   cry   ONOM
'The monkey saw the tiger calling *kraak¹ kraak¹*.'

The imitative *kraak¹ kraak¹* is arguably a core element of the sentence in (140). A fuller examination of these in the Tai languages is beyond the scope of this thesis.

8.3.6 Elaborate expressions

In example (74) above, we saw a four syllable expression, or elaborate expression. For standard Thai, Haas (1964:xvii) pointed out that most of these expressions have four parts and are semi-repeated.

The examples that will be presented here are all from a single text, a blessing given to the writer by Nang Pe of Namphakey village, who is particularly skilled in forming them. This should not be taken to imply that this type of expression is confined to texts such as blessings. In fact they are very common in all styles.

Sometimes these elaborate expressions are simply reduplicated compounds, as in (141), in which the elaborate expression *hum¹ hum¹ com² com²* 'to be glad' is a compound of two verbs.

141) ꩠ   ꩠ   ꩠ   ꩠ   ꩠ   ꩬꩠ   ॥

hã⁵   nai⁴   kau²   hum¹ hum¹   com² com²   se⁶
time   this   1Sg   glad   glad   PRT
'Now I am very appreciative.'
(Phake Blessing *Blessing*, No (3), intoned by Nang Pe)

When polysyllabic loan words are compounded, they can also form elaborate expressions, as with *puk⁴ cɔ¹ kan¹ tɔ²* 'worship' in (142):

142) ꩠ   ꩠ   ꩬꩠ   ꩠ   ꩠꩠꩠꩠ   ꩠꩠꩠꩠ   ꩠ

kau²   com²   se⁶   kau²   puk⁴ cɔ¹   kan¹ tɔ²   kã¹
1Sg   happy   PRT   1Sg   worship   homage   GO

ꩠꩠ   ꩠꩠ   ꩠꩠ   ॥

haŋ²   mün²   cau³
to   3Sg   RESP
'I pay my respects to him.'
(Phake Blessing *Blessing*, No (13), intoned by Nang Pe)

In example (142), the preferred four syllable structure is maintained and there is a rhyme on /ɔ/ which gives the expression a partly reduplicated feel.

As already mentioned in relation to example (74), sometimes these expressions conflate two constituents, as in (143), where the expression *thüŋ⁶ süŋ¹ thüŋ⁶ məŋ²* is translated as 'to (our) country'

143)

au²	thüŋ⁶	süŋ¹	thüŋ⁶	məŋ²	mā²	nai⁴
take	reach	country	reach	country	come	this

mā²	haü³	cā⁵	kau²	nāŋ² pe¹	ne⁴
come	give	to	1Sg	Nang Pe	DEF

'He has brought (them) here to (our) country, to give to me Nang Pe.'
(Phake Blessing *Blessing*, No (12), intoned by Nang Pe)

Often there is more than one elaborate expression in a sentence, as in example (144), which has three:

144)

khai³	wat¹	khai³	ai²	ho⁶	pan¹	ho⁶	mü²
fever	cold	fever	cough	head	giddy	head	dizzy

pi¹	haü³	nai³	mī²	nai³	yaŋ²
don't	GIVE	get	have	get	have

'May (you) not have colds and coughs, giddiness or dizziness.'
(Phake Blessing *Blessing*, No (16), intoned by Nang Pe)

The elaborate expressions in (144) are *khai³ wat¹ khai¹ ai²* 'have a cold and cough', *ho⁶ pan¹ ho⁶ mü²* 'giddy and dizzy' and *nai³ mī² nai³ yaŋ²* 'get to have'.

Finally, example (145) is presented, containing a series of elaborate expressions:

au²	mon²	phrā²	mon²	trā²	au²	mon²	cau³	saŋ¹ khā¹
take	glory	Buddha	glory	Dharma	take	glory	RESP	Sangha

haü³	hom¹	haü³	cā¹	haü³	ŋam²	kā¹
GIVE	cover	GIVE	cover	GIVE	cover	GO

sü⁵	naü⁶	tan²	naü⁶	sī⁶	kə⁴
straight	where	place	where	PRT	LINK

'By the grace of the Buddha, the Dharma and the respected monks, may you be protected and covered in every place.'
(Phake Blessing *Blessing*, No (17), intoned by Nang Pe)

In the first line of (145), the three gems of the Buddhist religion, *phrā²* 'the Buddha', *trā²* 'the Dharma (teaching of Buddha)' and *saŋ¹ khā¹* 'the Sangha (followers of Buddha)' are compounded into a long elaborate expression with the phrase *au² mon²* which means 'by the grace of'.

In the second line, *haü³* 'to give' is a causative or benefactive auxiliary combined with three verbs meaning 'to cover', and in the third line, *sü⁵ naü⁶ tan² naü⁶* is a semi reduplicated expression meaning 'every nook and cranny, everywhere.'

8.4 Constituent order

8.4.1 Preliminary theoretical considerations

Word order, or more accurately constituent order, is one of the typological features which distinguishes some languages from others. The most frequent constituent order of languages of the Tai family is Subject-Verb-Object (SVO) (Hudak 1990:40). However in Northeast India, Subject-Object-Verb (SOV) sentences have been reported by all researchers since Robinson (1849) (see 3.2.3 above), and SOV has been claimed to be the basic word order for at the Khamti language. This claim will be discussed in detail in 8.4.2 below.

For Standard Thai, Vichin Panupong (1970), has defined 'subject' as the noun or noun phrase which appears in the position in front of the verb and 'object' as that which

immediately follows the verb. She then posits that in certain circumstances one of these noun phrases may be topicalised.

For the Tai languages of Northeast India, a definition of subject and object based solely on the position of a noun phrase in the sentence is problematic. This is because of the following factors, some of which apply to Standard Thai also:

146.1) In any sentence any of the core participants (expressed as noun phrases) may be dropped if its meaning is otherwise understood,

146.2) Any of the noun phrases may be topicalised.

146.3) In some circumstances, particularly with quantifier phrases, the modifier within the noun phrase may be separated from its head.

146.4) Non agent noun phrases, if they are animate, are often marked in a prepositional phrase, and

146.5) The languages prefer that if there is only one noun phrase, such as the patient when the agent is dropped, it should stand before the verb, excepting where it is a case of (near-) noun incorporation.

Before proceeding to discuss constituent order in detail, each of these points need to be exemplified.

The dropping of core participants is exemplified in (147), which is the kind of sentence spoken by one Tai Aiton person to another when meeting along the main street of the village:

147) ကၟ ၌
kaa¹ dauɯ¹
go where?
'Where are (you) going?'
(Aiton sentence)

The subject noun phrase, 'you', is understood, even though it would not have been stated already in the conversation. The sentence 'Where am I going' would be a highly marked sentence and the subject would have to be stated.

A second circumstance in which a core argument NP can be dropped arises when the referent of the NP has already been mentioned and is clearly understood. Such an example is (148), where the subject, a¹ māt‘ 'minister', is understood because it was introduced earlier in the story.

148) ကၟ ၈၌ ၈၌ ၈ၗ၍ ၈၈ၗ၍ ၈၌ ၈ၕ ။
kā¹ thüŋ⁶ tï⁵ mān³ an² nüŋ⁵ yɛ⁴
go reach place village CLF one PRT
'He reached a village.'
(Phake Story, Story of the foolish king, No. (13), told by Ee Nyan Khet)

The topicalisation of NPs is demonstrated in example (149), where the object NP nɔŋ³ saau¹ nai¹ is the topic and the constituent order has become effectively OSV.

149) ၈၌ ၈ၗ ၌ ကၟ ၈၌ ၈၌၍ ၌ ၈ၕ ။
nɔŋ³ saau¹ nai³ (0.7) aa¹ suɯ¹ phuik¹ nai³ (0.35) au²
younger female this HESIT tiger white this take
'As for this younger daughter, the tiger took her (as a wife).'
(Aiton text, History of the Tai, No. (81) told by Sa Cham Thoumoung)

Tai speakers know which noun phrase is which constituent in the sentence on the basis of context and also prosody. To indicate the importance of prosodic features, the significant pauses in (149) have been marked with their length. The first pause is somewhat longer than would normally be expected and is followed by a hesitation.

The marking of the direct object or patient in prepositional phrases has been discussed above, in 8.3.3.

As stated above, it seems to be the case that noun phrases are sometimes topicalised to ensure that there is a noun phrase in front of the verb. In example (150), there are

three clauses, each with the main verb *kin²*'eat', in the causative, so that the subject of the causation is an unstated 'they'. The first clause is not associated with any other noun phrase, but in the second and third clauses, the other constituent, a prepositional phrase headed by *haŋ²*, occurs in front of the verb.

150)

hauɯ²	kin²	dii²	ʒau³	haŋ²	mɛ²	kɔ³	hauɯ³	kin²
GIVE	eat	good	FINISHED	to	mother	also	GIVE	eat

haŋ²	man²	kɔ³	hauɯ³	kin²
to	3Sg	LINK	GIVE	eat

'They gave good food to his mother and also to him...'

(Aiton Story, *The twelve questions*, No (77), told by Sa Cham Thoumoung)

There are numbers of exceptions to the principle expressed above in (146.5). One is when there is a particular wish to topicalise the verb, as in (151), where it is the forgetting of the Khamyang variety by younger people that Chaw Deben wants to comment upon, and so the verb appears before the subject of the sentence.

151)

naŋ¹	kɔ¹	luɯm²	caŋ¹	nai³	luk⁴	nɛ⁴	khau⁶
moreover	forget	now		child	PRT	3Sg	

luk⁴	ɔn¹	khau¹	ɣŋ¹	taan⁴	haai⁶	yaa⁴
child	3Pl	NEG	speak	disappear	FINISHED	

'Moreover, they've forgotten, those children now, those children can't speak, it's all gone.'

(Khamyang text, *The future of the Khamyang language*, No. (20), spoken by Deben Chowlik.

Another exception is object incorporation (see 8.2.3.1.2). Except in negative sentences, the object follows the verb, such as in *kin² khau³*'eat rice, dine', even if it is the only noun phrase in the utterance.

As a consequence of the various factors detailed in (146) above, in any sentence, whether it is a one, two or three participant event, one or more of the places may be vacant. Take for example the case of the verb 'give', which semantically requires an agent, a theme and a beneficiary. A sentence having 'give' as the main verb is always a three participant event, even, when, for example, the agent and theme are expressed but the beneficiary is not.

In example (152) (already discussed above as example (113) in 8.3.3.2), there are two sentences in each of which the main verb has three arguments, although in (152.1) the agent is a vocative and in (152.2) only one of the three arguments is expressed.

152.1).

ai²	pū¹	nɔ⁴	maü²	nai⁴	haü³	mā²	cā⁵	kau²
VOC	grandfather	meat	2Sg	this	give	come	to	1Sg

'"Oh! granddad, give me your meat!"'

152.2).

maiŋ¹	mā²	cā⁵	hau³	wā⁵	niŋ¹	nai⁴
distribute	come	to	us	say	like	this

'"Distribute it to us!", he spoke like this'

(Phake Story, *The two brothers*, No. (152), told by Ee Nyan Khet)

It is argued that (152.2) is a three place predicate, because the semantics of the verb requires it, and because following the principle in (146.1) above, the other two places are occupied by noun phrases that are understood.

8.4.1.1 Is there a basic constituent order in Tai?

Given the large variety of constituent orders that actually surface in the data for the Tai languages of Assam, it might be argued that there is no basic constituent order, and no need to postulate one. Sentences are uttered in a context, whether a written or spoken one. That context is made up of pragmatic factors and cultural factors as well as whatever has previously been stated in whatever text is being examined. These alone might be enough to disambiguate any sentence, and therefore a basic word order becomes unnecessary.

However, as we will see in 8.5.6, when a ditransitive verb has all its three arguments present, and none in a prepositional phrase, the order is fixed. From this, it is argued that the basic constituent order for one participant events is as in (153), for two participant events in (154) and for three participant events in (155), each subject to all the constraints and possibilities in (146) above.

153) S   V

154) S   V   O

155) S   V   O   I

Before discussing the findings of ... present study of Tai Aiton and Tai Phake, it will be necessary first to discuss the issue of verb phrases in the Tai languages (see 8.4.1.2) and then to review the literature relating to Tai Khamti (see 8.4.2).

8.4.1.2 Is there a verb phrase in Tai?

The concept of verb phrase is regarded by some linguists as a universal category. A verb phrase, consisting of the verb and its object or objects, is claimed to be present in all languages.

For the Tai languages of Assam, it will be necessary to offer evidence for a constituent verb phrase. On the surface, it seems counterintuitive to posit a verb phrase for the Tai languages when sentences like (156) occur:

156)
phün⁶ sɔn⁶ tǎŋ² nī² nai⁴ hau² tüŋ¹ hum¹ tüŋ¹ com² sī⁶
CLF teach way good this 1Pl all glad all glad PRT

cā⁵ maü² cau³ sa¹ tī¹ fɛn⁴ mo¹ lɛ⁴ pan² sü¹ haü³ mǎ² yau⁴
to 2Sg RESP Stephen Morey give PRT give COME FINISHED
'We gladly present this book *Teaching in a good way* to you, Mr Stephen Morey.'
(Tai Phake sentence, personal letter from Namphakey village, April 2001)

In the English translation of (156), the verb phrase is *present this book, 'Teaching in a good way', to you, Mr Stephen Morey.'*

The structure of (156) can be expressed as (157):

157) Object   Subject NP
      phün⁶ sɔn⁶ tǎŋ² nī² nai⁴   hau² tüŋ¹ hum¹ tüŋ¹ com² sī⁶
      this book of Teaching in a good way   we who are glad'

      Indirect Object PP   Verb complex
      cā⁵ maü² cau³ sa¹ tī¹ fɛn⁴ mo¹ lɛ⁴   pan² sü¹ haü³ mǎ² yau⁴
      to you, Stephen Morey   present[66]

If we accept that the unmarked constituent order for such sentences is that presented in (155) above, then we have to explain how the two objects can end up in different positions from those in (155), one (the direct object) in the topic position in front of the subject, and the other (the indirect object) between the subject and the verb.

A verb phrase analysis would allow a sentence like (156) to occur, provided that we argue that a noun phrase is still within the verb phrase if is occurs immediately before the verb, as well as immediately after it. We could then argue that in (156), the indirect object was still within the verb phrase and only the direct object had been topicalised. The subject remains in its canonical position given in (155).

If we do not accept that the indirect object in (156) is within a verb phrase, then we would have to posit three topicalisations. Firstly the indirect object is topicalised. Then the subject is topicalised in front of the indirect object, and finally the direct object is topicalised to the front of the utterance.

If multiple topicalisation were permitted, it would be confirmed if there were examples of both the direct object and indirect object occurring before the subject,

---

[66] Literally 'presenting so that it has come to you.'

that is to say the constituent orders OISV and IOSV. Neither of these has been recorded.

We will therefore conclude that an analysis positing a verb phrase can be defended for the Tai languages of Assam.

8.4.2 Constituent order in Khamti.

Greenberg (1990:67) listed the Khamti language, which is closely related to the Aiton and Phake, as having a basic constituent order of SOV. This is surprising considering the basic constituent orders posited in (154) and (155) above. Greenberg's conclusions appear to be based on Grierson (1904), which is itself based on Needham (1894). It is therefore necessary to review what Needham recorded.

8.4.2.1 Needham

Although many of Needham's sentences are in the SOV order, they are not always so, as in (158):

158) sü   kâp   ngô   yau
      ဃ၆   ကၵ   ၚ၆   ဃ၆
      tiger bite   cow   PAST
      '(A) tiger killed (a) cow.'
      (Needham 1894:11)

When we examine some of the texts provided by Needham, a more complex story emerges, as shown in (159):

159.1) ပၠလၚၵ်ၿၵ် ဃ၆ ကၤ ထၵ်က ကၤ တၜၵ် ၉ ၂ ထိၵ် ။
       Palungkhu   shoot PAST hit PAST foot ACC pig   forest
       'Palungkhu fired (and) hit the wild pig on (the) leg, '

159.2) ၂   တၜ ဃၜ   ၉ ကၤ ။
       NEG   die   go
       'but (the pig) not dying, fled.'
       (Needham 1894:127)

In (159.1) the verbs ဃ၆ *vü* 'shoot' has ပၠလၚၵ်ၿၵ် *Palangkhü* 'Palangkhu' as its subject. The second verb ထၵ်က *thuk* 'hit, touch' has တၜၵ် *tin* 'foot' as its object. This object is marked by the Khamti particle *mai*, which Needham glosses as 'accusative'. Example (159.1) certainly has SVO order.

Diller (1992:20) examined Needham's sentences in detail and listed the following sentence configurations:

160) S   O   V   [transitive]
      O+[MARKER]   S   V   [transitive]
      S+[MARKER]   V   [intransitive]
      O   S   DAT.+[MARKER] V   [transitive]
      S   V   O   [transitive]
      S   O+[MARKER]   V   [transitive]
      S   V   [intransitive]
      S   [(zero copula) (zero agent) V   O]PURPOSE

Diller's conclusion on the data from Needham is that "the general impression, especially from connected text, is of a very 'pragmatically' controlled configuration: issues such as topicality, specificness, and so forth, appear to play a crucial role on how sentence constituents are ordered." (1992:21). It is difficult to see how the contention that Khamti is a basically SOV language could be derived from this data.

8.4.2.2 Wilaiwan

After fieldwork in the Khamti lands, Wilaiwan (1986:178) concluded that "SOV is the dominant word order in Kamti while in other known Tai dialects SVO is the dominant one. We may conclude ... that Kamti has developed from SVO to SOV. It has also been shown that Kamti contains a set of object marking postpositions."

Wilaiwan (1986:175) exemplifies her claim with sentences like (161). The numbers above the transcription indicate tone.

         5   1   1   6   2   1   6
161) man   wan-phan ne? ?au-phee   yɔ?
      He   glass   the to break   past tense particle
      'He broke the glass.'

In (162) she exemplifies the object marker *mai*4:

2	5	4		3	1	3	6
162)	*kau*	*man+mai*		*khau-*	*san*	*hɔu*	*yɔ*ʔ
	I	he object marker		rice		to give	past tense particle

'I gave her rice.'

Although I have not been able to investigate Khamti to the extent that I would wish, it seems likely that the marker *mai*4 is a postposition similar in the range and meaning to the prepositions *haŋ²* and *caa⁵/cā²*, discussed above in 8.3.3.

Wilaiwan adds that "In the event that the speaker wants to emphasize the nominal object, it can be topicalized and moved to the front of the sentence." (1986:178). This is exemplified in (163), where the object in (161) has been topicalised:

1	1	6	5	2	1	6
161)	*wan-phan*	*neʔ*	*man*	*ʔau-phee*	*yɔ*ʔ
	glass	the	he	to break	past tense particle

'He broke the glass.'

Wilaiwan compares the Khamti data with the Phake, which she also investigated, and concludes that Phake "also has OSV as the alternative basic order, but its dominant order is SVO." (1986:176).

8.4.2.3 Chau Khouk Manpoong

A very brief survey of the Khamti primers published by Chau Khouk Manpoong in 1993 suggest that in Arunachal Pradesh Khamti, both SVO and SOV sentences are found, as in the following examples:

162). ဗက် န် ၮၢၵ �662မ် ြၵၢ် ယ၀ၱ

*man*⁷	*nai*⁶	*ñaan*³	*pen*⁴	*phraa*⁷	*yau*⁹/⁵
3Sg	get	wisdom	be	Buddha	FINISHED

'He got wisdom and became a Buddha.'
(Chau Khouk Manpoong 1993 II:15)

163). ဗက်ၵ် ယ၀တ န် ၮၢၵ် ၮ္ န် ၮၢꩮ ၮၵ ယ

*maa*³	*nin*⁴ *sit*³		*nai*⁵	*kɔn*⁶	*nɣ*⁵	*nai*⁵	*kaap*⁴	*au*⁴	*sii*³
fox	clever	this	lump	meat	this	bite	take	PRT

'The clever fox ate the lump of meat.'
(Chau Khouk Manpoong 1993 II:36)

8.4.2.4 Present investigation

In March 2000, I was fortunate to meet Chaw Khamoon Gohain, a traditional Khamti chief. I asked him how to say 'the tiger ate the deer' and he immediately gave (164.1) as the answer.

164.1) ယ၆ၵ် ၮၢꩮ ၮၢၵ်

*sɯ*³	*kaap*³	*phaan*⁷
tiger	bite	barking deer

164.2) ယ၆ၵ် ၮၢၵ် �e ၮၢꩮ

*sɯ*³	*phaan*⁷	*mai*⁶	*kaap*³
tiger	barking deer	to	bite

'The tiger ate the barking deer.'
Audio link for these sentences (\wave files\khamti\1.mp3)

After discussing the sentence, Chaw Khamoon Gohain added that (164.2) was equally acceptable to Khamti speakers. Among the small amount of text that I was able to record from him, several examples had the verb in final position, as in (165):

165) ၮၢၵ် ယ e ၮꩮၢ e ꩮ ꩮ ꩮ ယ ၵ

*kon*⁴	*yau*¹	*mai*⁶	*cau*⁶ *phaa*⁶	*mai*⁶	*hau*⁷	*khup*⁴	*khau*³	*sii*³	*paai*⁶
person	big	to	king	to	1Pl	kneel	knee	PRT	pray

'We kneel down and pay homage to important people and kings.'
Audio link for this sentence (\wave files\khamti\2.mp3)

On the basis of (164) and (165) above, we might tentatively conclude that SOV would be possible in Khamti if the O is marked by the word *mai*⁶. This suggests that SVO may be the more basic constituent order. Whilst it appears likely that Khamti constituent order is pragmatically controlled, as suggested by Diller above in 8.4.2.1, only further research, including the collection and analysis of substantial corpus of

Khamti texts, will be able to further illuminate this issue. In the meantime, scholars would be advised to treat the claim that Khamti is an SOV language with some caution.

8.4.3 Constituent order in the Tai languages of Assam

The late Aimya Khang insisted that SVO was the basic constituent order in Phake. Not long before he died, the *Tai Phake Primer* (Morey 1999c) was published, containing simple sentences for people to practise reading. Many of the sentences were supplied by Sam Thun Wingkyen, including (166)

166) ၵ် လၢင် မေၵ် ၵꩦ
    nău² ān¹ mau⁵ pɛ⁴
    star count NEG can
    'The number of stars cannot be counted.'

In his review of the Primer, Aimya Khang stated that this was an SOV order (with the subject unstated), and that (166) should be altered to (167):

167) လၢင် ၵ် မေၵ် ၵꩦ
    ān¹ nău² mau⁵ pɛ⁴
    count star NEG can
    'The number of stars cannot be counted.'

Writing of example (166), he said that:

> "Such way of writing syntactical sentence may lead others to think and
> practise corrupt sentence. Our aim and objectives should be to teach the
> learners the correct way." (Aimya Khang Gohain, pers. comm.)

Regardless of how we view Aimya Khang's aims and objectives, he was one of the most linguistically aware members of the Phake community, and his strong perception that SVO is the basic word order of the Phake language, and indeed of all the Tai varieties in Assam, must be given serious attention.

In this thesis, it is posited that there is an underlying constituent order in Tai Aiton and Tai Phake, which is presented in (168):

168) (Iso) (Non-core) (NP) V (NP) (NP) (Non-core) (Iso.),

In (168), *Iso* stands for an isolative phrase (see above 8.3.5), *Non-core* for the locational or temporal phrases (see above 8.3.4) and *NP* for the noun phrases or in some cases core prepositional phrases (see above 8.3.3). The grammatical roles of the noun phrases in (168) has already been set out in (155) above.

The frame in (168) is exemplified in (169), where all three noun phrases are present, although none of the non-core or isolative phrases are:

169) ၵ ၀ꩧ ၵꩧ ယꩦ ၵ် ၵꩦ နꩦ ‖
    kau² ta¹ kā¹ soŋ¹ khau³ phɩ⁶ nam³
    1Sg will go send rice spirit ancestor spirit
    'I am going to bring rice to the ancestor spirit.'
    (Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (18), told by Ee Nyan Khet)

Examples of all three noun phrases spelled out, none being in prepositional phrases, are rare. On the other hand, two participant events in which neither noun phrase is marked in a prepositional phrase occur frequently in the corpus, as example (170):

170) လꩨတꩫ် ၵ် ꩦၵ် ꩦၵ် နꩦၵ် မꩧ ‖
    aai³ tɔn² khau¹ sɛu³ suɯ¹ nip¹ /dip¹ maa²
    Aiton 3Pl catch tiger alive COME
    'The Aitons caught a live tiger.'
    (Aiton text, *History of the Aiton*, No. (79) told by Nang Wimala Thoumoung)

Non core elements such as locational phrases can be placed outside this core, at either end of the utterance. Examples (171) are two different versions of the same sentence, with the non-core locational phrase *tii² nan³* 'there' in both positions :

171.1) တီ့ နၢၣ် တဲ့ၣ် ကဲ့ ဆျိၣ် နွံ ကိ့ၣၢ မၢ ။

tii² nan³ tɛp¹ kai¹ sɛŋ¹ nɔi¹ kin² maa²
place that cut chicken diamond PRT eat COME

171.2) တဲ့ၣ် မၢ ကဲ့ ဆျိၣ် နွံ ကိ့ၣၢ မၢ တီ့ နၢၣ် ။

tɛp¹ maa² kai¹ sɛŋ¹ nɔi¹ kin² maa² tii² nan³
cut COME chicken diamond PRT eat COME place that

'At that place, (he) killed the diamond chicken and it was eaten.'

(Aiton text, *History of the Aiton*, Nos. (43) and (43.1), told by Nang Wimala Thoumoung)

Commonly, however, one of more participants, expressed either as a noun phrase, a classifier phrase or a prepositional phrase, is topicalised as in (172), where the subject noun phrase is in front of the temporal phrase.

172) ၢၣ်ၣ် ၤဲ့ နဲ ထိၣ် နႝ ၢႝတဲ့ၣ် တၢ နွံၣ် တၣ လ

ə² ŋo² nai⁴ thün⁶ naü² cet¹ wan² mün² ta¹ phɛ⁵
YES cow this reach in seven day 3Sg WILL bear young

'Yes, this cow of mine will bear young in seven days.'

(Phake Story, *Chaw Mahosatha and the raising of cattle*, No. (11), told by Ee Nyan Khet)

Example (172) was spoken immediately after (173), a sentence in which the canonical order proposed in (168) was observed. Since in the canonical structure, the subject occurs before the verb, subject topicalisation might only be apparent when there is a non-core element, in this case a temporal phrase, present.

173) နႝ ၢႝတဲ့ၣ် တၢ နဲ တႝ ၤဲ့ ကၢၣ် နဲ

naü² cet¹ wan² nai⁴ nɛ⁴ ŋo² kau² nai⁴
in seven day this DEF cow 1Sg this

တၣ လ ၤဲ့ လႝ ၤဲ့ ထိၣ် ၤဲ့ လႝ ။

ta¹ phɛ⁵ ŋo² loŋ⁶ ŋo² thük¹ ŋo² loŋ⁶
WILL bear young cow big cow male cow big

'Within these seven days my cow will bear large male young.'

(Phake Story, *Chaw Mahosatha and the raising of cattle*, No. (10), told by Ee Nyan Khet)

Topicalisation of the object is also frequently encountered, as in (174):

174) ပိၣ် ကဲ့ ၢၢၣ် ၣ ၡၢ တဲ့ၣ် နွံ န ၢႝၣ်ၣ်ၢၣ် ။

pik¹ kai¹ an² khau¹ kɔ³ tɛp¹ nɔi³ hau³ khum¹ kɔn¹
wing chicken CLF 3Pl LINK cut PRT give again

'So (she) cut the chicken wings off and gave it to her, as before.'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (62), told by Sa Cham Thoumoung)

In (174), the subject is not stated and the object is topicalised. This might suggest that object fronting is meeting the principle in (146.5) above. However in (175), the subject is stated and even so the object is topicalised.

175) လ ၤႝ ၣ လ့ၣ်ၣ် တ ၢၣ် တၣ ၢၣ် ဆျိၣ် တၣ် ။

mɛ² mün² wä⁵ luk⁴ nai⁴ kau² ta¹ au² sauŋ⁶ to²
wife 3sg say child that 1Sg will take two CLF

"The wife said: 'I'll take two of the children.'"

(Phake Story, *Story of dogs dividing up their children*, No. (7), told by Sam Thun Wingkyen)

Examples (175) was one of those sentences which Aimya Khang insisted should be altered so that SVO order be maintained, as in (176).

176) လ ၤႝ ၣ ၢၣ် တၣ ၢၣ် လ့ၣ်ၣ် တ ဆျိၣ် တၣ် ။

mɛ² mün² wä⁵ kau² ta¹ au² luk⁴ nai⁴ sauŋ⁶ to²
wife 3sg say 1Sg will take child that two CLF

Core participants can also occur at the end of a sentence, as in (177).

177) မၢ တၢ ၢၣ် တဲၣ် ။

ma¹ tü¹ kan² sɛŋ⁶
NEG same RECIP sound

'Their sound is not the same.'

(Phake sentence, spoken by Ee Nyan Khet)

Audio Link for this sentence (\wave files\extracts\1-10-2-9\seng.mp3)

In some particular circumstances, the subject can be postposed while the object is topicalised as in (178):

178)   လိုက် တမ် ကၞေ လဟတ် လ ကၟ ဿၯာတ်
    lik⁴ tɛm³ kə⁴ phat¹ pɛ⁴ kã¹ sã¹ hãp¹
    book write LINK read can GO Sahib
    'He can read the books when they are written, the Sahib.'

    (Tai Phake Song, *Poem in the khe² khyäŋ² style: How Stephen Morey came to the Toi Phakes,*. No (90.1), read by Ai Che Let Hailung)

Example (178) was spoken in explanation of (179), in which the object was topicalised:

179) လိုက် တမ် ဎိမ်လၟ လဟတ် လၟ ပေန် ၊
    lik⁴ tɛm³ cim¹ cɔm¹ phat¹ pɔ² pen²
    book write here and there read enough be
    'He can read the books a little.'

    (Tai Phake Song, *Poem in the khe² khyäŋ² style: How Stephen Morey came to the Tai Phakes,*. No (90), read by Ai Che Let Hailung)
    Audio Link for this sentence example.
    Audio Link for the *Poem in the khe² khyäŋ² style: How Stephen Morey came to the Tai Phakes.*

It appears that in (178), Ai Che Let is restating (179) in more everyday language, and then adding a resumptive subject. Thus on the surface the constituent order OSV appears, but this is a highly marked situation which cannot arise except in circumstances similar to those obtaining here. The roles of the various noun phrases are disambiguated by the context, and also semantically: obviously it cannot be the books that are doing the reading.

Constituent order variation is even wider where there are three core participants. Example (180) shows SOVI order, (181) SOIV and (182) OSIV. In each of these three sentences, topicalisation of one of the noun phrases has occurred, and at either the direct object or the indirect object or both has been moved to the position in front of the verb. This is argued to be still within the verb phrase, as discussed above in 8.4.1.2.

All of the noun phrases that are not in the canonical positions are marked either by demonstratives such as *nai⁴*, or the definite particles *nɛ⁴*.

180) မၟ �140 ပူ တ ၞၟ ဟေတ် တ
    mə⁴ nan⁴ pũ¹ nɛ⁴ nə⁴ mũn² nɛ⁴
    time that old man DEF meat his DEF

    မၟ ဟၟ ကၟ ဟၟ ခၟ ၂ ၼၟ ဿၟ
    maiŋ¹ haü³ kã¹ haŋ² khau⁶ pĩ⁵ nauŋ⁴ se⁶
    dist. GIVE GO to 3Sg elder younger PRT
    'Then the old man gave his meat to those two brothers.'

    (Phake Story, *The two brothers*, No. (14), told by Ee Nyan Khet)

181) ကၟ ပၟလိုက် ၼ ဟၟ ပၟဿၟ တ ဟၟ ကၟ
    kau² pãp⁴ lik⁴ nai⁴ haŋ² pa¹ sau⁶ nɛ⁴ haü³ kã¹
    1sg book this to girl PRT give GO
    'I gave this book to the girl.'
    (Phake sentence elicited from Aithown Che Chakap)

182) ဖၟ ဿၟ တၟ ၼ ၼ ဟၟ တၟ ဟၟ တၟ ၼၟ ဿ
    phün⁶ sɔn⁶ tãŋ² nĩ² nai⁴ hau² tüŋ¹ hum¹ tüŋ¹ com² sĩ⁶
    CLF teach way good this 1Pl all glad all glad PRT

    ၵၟ မၟ ၼၟ ဿၟတၟ ၼၟတၟ ဟၟ ၼ ၼၟ ၼ ၼၟ
    cã⁵ maü² cau³ sa¹ tĩ¹ fɛn⁴ mo¹ lɛ⁴ pan² sü¹ haü³ mã² yau⁴
    to 2Sg RESP Stephen Morey give PRT give COME FINISHED
    'We gladly present this book *Teaching in a good way* to you, Mr Stephen Morey.'
    (Phake sentence in personal letter from Namphakey village, April 2001)

In each of these three sentences, the semantic roles of each of the three noun phrases is clear both from the context, the syntax and semantics of the verb. The semantics ensures that the meat in (180) and the book in both (181) and (182) are clearly understood to be the theme of their respective sentences. Syntactically, in each of these sentences, the beneficiary/recipient is marked in a prepositional phrase. Context, at least in (180) and (182) will also make it clear who is transferring what to whom.

What then is the motivation for this considerable variation in constituent order? It has been suggested that these languages are moving towards verb final constituent order through the influence of the strongly verb final languages that surround them, namely Assamese and the languages of the Tibeto-Burman family. Alternatively the motivation may be largely discourse related, with items being topicalised for reasons of special emphasis as is the case in example (151) above. Furthermore, as will be argued in 8.5.7.5 below, there is an increasing tendency for the verb to be attached to TAM markers. Since the TAM markers we   orically utterance final, there may be an additional motivation for verb final structures.

However, the evidence presented in this section makes it clear that we cannot yet conclude that the Tai languages are synchronically verb final; the underlying order posited in (168) above remains the closest we can come to basic structure for these languages.

We can represent the entire frame as in (183), subject to all of principles in (146) above.

183) Sentence → (Iso P) (Non-core P) core (Non-core P) (Iso P)

   Core → S [(TAM$_1$) V (TAM$_2$)] O I (TAM$_3$)

8.4.3.1 Differences in constituent order between Tai Aiton and Tai Phake ·

There are a number of quite consistent differences in constituent order between Aiton and Phake. In a discussion with Nabin Shyam (Aiton) and Ee Nyan Khet (Phake), two particular constructions were identified as being different between the two varieties.

The first of these is the completive verb construction with pe^3/pe^4 'can', already exemplified above in examples (46) and (47) in section 8.2.3.3. In Aiton, the structure of a negative sentence with this completive is as in (184), in which the completive verb occurs before the main verb. The Phake structure is presented in (185), where the completive follows the main verb.

184) uŋ^1 pɛ^3 kin^2
      NEG   can   eat
      '(I) cannot eat (it).'
      (Aiton sentence spoken by Nabin Shyam)

185) kin^2 ma^1 pɛ^4
      eat   NEG   can
      '(I) cannot eat (it).'
      (Phake sentence spoken by Ee Nyan Khet)

A second difference between the two comes with the negation of verb compounds. Aiton compound verbs are generally negated without splitting the compound, as in (186), whereas, as discussed above in 8.2.3.1.2, in Phake the compound verb splits when negated, as in (187):

186) uŋ^1 kin^2 khau^3
      NEG   di..e
      '(I) haven't/won't eat.'
      (Aiton sentence spoken by Nabin Shyam)

187) khau^3 ma^1 kin^6
      rice   NEG   eat.NEG
      '(I) haven't/won't eat.'
      (Phake sentence spoken by Ee Nyan Khet)

8.4.4 A note on syntactic packaging

Many of the issues covered in the above discussions on constituent order have touched on issues of the way in which information is presented in a clause. Foley and Van Valin (1985:282) called this information packaging and discussed issues such as passivisation, topicalisation and left dislocation.

In the Tai languages, there is no passivisation in the sense that English *The dog was chased by the boy* is a passive sentence, because there is no 'by phrase' and there is no passive morphology. In the Tai languages, following the principles introduced in (146) above, if there is a desire to foreground one of the participants, it can be shifted to the front. If there is a desire to background one of the participants, it can be omitted. Thus we can translate (174) above as *The chicken wings were cut off and given to her as before*. The fronting of the patient argument is thus expressed in English as best as English can.

8.5 Predications

In this section on predications, three types of sentences in which the verbal constituent is sometimes dropped, copula (see 8.5.1), existential sentences (see 8.5.2) and definitional sentences (see 8.5.3) will be examined first. Following that, three types of sentences in which verbs are obligatory will be examined: one participant events or sentences with intransitive verbs (see 8.5.4), two participant events or sentences with transitive verbs (see 8.5.5), and three participant events or sentences with ditransitive verbs (see 8.5.6).

The verbal constituent or verbal complex is made up of, minimally, a verb. However other elements can be included in a verbal complex. The constituents of this verbal complex, in the order in which they appear, are as follows:

- Modal verbs,
- TAM markers that precede the main verb (TAM₁)
- The main verb
- TAM markers that immediately follow the main verb (TAM₂)
- Completive verbs that may be separated from the main verb and placed at the end of the core sentence, and
- TAM markers that may be separated from the main verb and placed at the end of the core sentence (TAM₃).

TAM markers and their syntax are discussed in detail in 8.5.7 and completive verbs in 8.5.9. There may be more than one main verb in a sentence. This verb serialisation is discussed in 8.5.8.

8.5.1 Copula

Copula sentences in Tai are expressed by two nominal constituents, either with or without the copula verb *pin²/pen²*. Dixon (2002) categorised the criteria for recognising a verb as a copula as follows:

- It can have two core arguments
- It has a relational rather than referential meaning
- It is used for, at least, the relations of identity/equation/naming and or attribution (2002:4)

Sentences of identity and equation are frequently expressed with the copula verb. In example (188), the two nominal constituents, the first argument *pa¹ ruk¹ nai³* 'tomorrow' (called the copula subject by Dixon) and the second, a classifier phrase *cit¹ wan²* 'seven days' (called the copula complement by Dixon) are linked with *pin²*.

188) ဗုၵ္ရုၵ် ၼိ တၢ ပ်ိၼ် ၵ်ိတ် ဝၼ် တ်ိ တၢႆ ဒီႝ အၢ ॥
pa¹ ruk¹   nai³   ta¹   pin² cit¹ wan² ti¹   taai² dii² ɔ¹
tomorrow   this   will be   7   day   WILL die   PRT
'Tomorrow it will be the seventh day, and we will die.'
(Aiton Story, *The twelve questions*, No (62), told by Sa Cham Thoumoung)

Frequently, however, the copula verb is omitted, as in (189):

189) မႄ ထဝ် ၼၢၼ် ၵုၼ် ၽၢၼ် ...
mɛ²   thau³ nan³   kun²   phaan¹ ...
mother old   that   person poor   ...
"That old woman was a poor person...
(Aiton Story, *Story of the old woman*, No. (2), told by Nabin Shyam).

In copula sentences, generally the first noun phrase or copula subject is marked by a demonstrative, such as *nan³*'that' in (189), or by some other particle. However, if both noun phrases have the same head noun, as in (190), this marking is not necessary.

190) ꩠꩠꩠꩠ   ꩠꩠ   ꩠꩠ   ꩠꩠ   ꩠꩠ

    cɛ² nai³   baan³   hau²   baan³   thɯm¹

    now   village   1Pl   village forest

    'Now our village is a forest village.'

    (Aiton text, *History of Barhula*, No. (36) told by Ong Thun

In (190), the two noun phrases which are being equated are both headed by *baan³* 'village'. The marking of a copula subject by a demonstrative or other particle is also unnecessary if the head nouns of both the copula subject and copula complement are from a similar class of words as in (191), where two kinship terms, *lo²*'sister in law' and *pī⁵*'elder', are linked.

191) ꩠꩠ   ꩠꩠ   ꩠꩠꩠꩠ   ꩠꩠ   ꩠ   ꩠꩠ   ꩠꩠ ॥

    ə²   lo²   ī¹ kiŋ¹ kham² khau⁶ pī⁵ sāu⁶ kau²

    Yes sister-in-law Ee King Kham 3Pl elder female 1Sg

    'Yes, Ee King Kham's sister in law is my elder sister.'

    (Phake Text, *Story of her youth*, No. (31), told by Ee Nyan Khet)

An alternative way of marking the copula relationship between two noun phrases that do not have the same or similar noun as their head is to use a pause, as in (192):

192) ꩠꩠꩠꩠꩠ   ꩠꩠ   ꩠꩠꩠꩠ   ꩠꩠ

    sa¹ tī¹ phɛn¹ (0.33) kon² pɔn¹ ɕit¹ loŋ⁶

    Stephen   person pandit   big

    'Stephen is a great pandit.'

    (Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to the Tai Phakes*,. No (7.1), read by Ai Che Let Hailung)

The third relation identified by Dixon (2002), that of naming, is also expressed with a copula sentence, as in (193), but in this case the copula verb *pin²/pen²* is never used.

193) ꩠꩠꩠ   ꩠꩠ   ꩠꩠꩠꩠ

    cü⁵   kau²   āi³ saiŋ⁶

    name 1Sg   Ai Seng

    'My name is Ai Seng.'

    (Phake sentence)

Here it is argued that *cü⁵ kau²* is the first argument or copula subject and *āi³ saiŋ⁶* is the second argument or copula complement. As noted above in example (3) in 3.2.4 above, there is some evidence for both arguments having being headed by *cü⁵* in an earlier phase of Aiton. Interestingly, in Standard Thai, the cognate of *cü⁵* heads the second argument (copula complement) and can be interpreted as both a noun and a verb.

In the Tai languages there is a second way of naming, using the verb *waa²/wā⁵* 'speak'. This is discussed below in 8.5.3.

Copula sentences, with or without the verb *pin²/pen²*, can have the meaning 'become' as well as 'be', as in (194):

194) ꩠꩠꩠ   ꩠꩠꩠꩠ   ꩠꩠ   ꩠꩠꩠ   ꩠꩠꩠꩠ ॥

    cau³ phaa³ kɔ¹ laaŋ³ ai³   cau³ phaa³ aai³ tɔn³

    king   Ko Lang   PRT   king   Aiton

    'The King of Ko Lang became the King of the Aitons.'

    (Aiton text, *History of the Aiton*, No. (55) told by Nang Wimala Thoumoung)

In some texts, copula sentences can have possessive meaning, as in example (195), which is a verbless copula sentence, or example (196), with the verb *pin²/pen²*.

195) ꩠꩠ   ꩠꩠ   ꩠ   ꩠꩠ   ꩠꩠ

    hɔn² loŋ⁶ sām⁶ pāk¹ hauŋ³

    house big   3   hundred rooms

    'The house is as big as three hundred rooms,'

    (Tai Phake *Riddles*, No. (4), read by Ee Nyan Khet)

196)   ဓ   ၈   ကၟင်   ၀ိတ်   ၈   မိၼ်   ၀ၼ်   ၏   ၀ိတ် ။

mai⁴	nai⁴	koŋ²	wət⁴	naü²	mün²	pen²	hü²	wət⁴
tree	this	hollow	PRT	in	3Sg	be	hole	PRT

'That tree had a hollow in it.'

(Phake Story, *Story of the foolish king*, No. (61), told by Ee Nyan Khet)

As mentioned above, the only circumstance in which the copula verb is used with adjectives is when the adjective relates to a state of health, as in (197):

197)   ကၟ   ၀ၼ်   ၈   ၀ၼ်   ၈

kau²	pen²	khai³	pen²	nāu⁶
1Sg	be	fever	be	cold

'I have a fever.'

(Phake Sentence)

The word *khai³* 'fever' is glossed by Banchob (1987) as 'to be ill', indicating that she regarded it as an adjective. The word *nāu⁶* 'cold' is clearly an adjective when used in a phrase like *nāu⁶ loŋ⁶* 'very cold' to refer to the weather[67]. One could not refer to the weather with a phrase like (198):

198)   * ၀ၼ်   ၈   လၟင်

pen²	nāu⁶	loŋ⁶
be	cold	big

However when *nāu⁶* refers to illness it can be used in combination with the copula verb as in (197) above. Another word referring to illness is *wat¹* 'have a cold', which can also be combined with the copula verb. This is used in combination with *khai³* 'fever' in (199):

199)   ၈   ၀၀တ်   ၈   ဢၟ   ၟၟ်   ပၼ်   ၟၟ်   မ

khai³	wat¹	khai³	ai²	ho⁶	pan¹	ho⁶	mü²
fever	cold	fever	cough	head	giddy	head	dizzy

---

[67] It can get very cold in Assam in December and January, particularly at night.

---

   ၈   ၈   ၈   ၈   ၈   ၀ၼ် ။

pi¹	haü³	nai³	mī²	nai³	yaŋ²
don't	GIVE	get	have	get	have

'May (you) not have colds and coughs, giddiness or dizziness.'

(Phake Blessing *Blessing*, No (16), intoned by Nang Pe)

The phrase *khai³ wat¹*, literally 'fever-cold' appears to be interchangeable with *pen² wat¹* 'be-cold'.

8.5.2 Existential sentences and possession

Existential sentences are a special kind of one participant event with the verb *yaŋ²* 'have'. In existential sentences, because there is only one other constituent, the verb is usually in final position. Example (200) contains two existential sentences in apposition:

200)   ကၟၼ်   ၀ၼ်   ၟၼ်   ၾ   ၀ၼ် ။

kon²	yaŋ²	hən²	ma¹	yaŋ⁶
person	have	house	NEG	have.NEG

'There are people, but there are no houses.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (26), told by Ee Nyan Khet)

In example (201), the translation is 'He had one daughter', but syntactically there is only one argument of the verb, a noun phrase *luk⁴ sāu⁶ mün² kɔ⁴ nüŋ⁵* 'his one daughter', but the modifying classifier phrase has been postposed.

201)   လ   လၟၼ်   ၀ၟ   မိၼ်   ၀ၼ်   ကၟ   ၟ   ၀ိတ် ။

ne⁴	luk⁴	sāu⁶	mün²	yaŋ²	kɔ⁴	nüŋ⁵	wət⁴
DEF	child	female	3Sg	have	CLF	one	PRT

"He had one daughter."

(Phake Story, *Story of the foolish king*, No. (15), told by Ee Nyan Khet)

Sometimes an existential sentence is verbless, as in the first line of (202), which has the same syntax as a verbless copula sentence, namely a first noun phrase *pha¹ ta¹ wii¹ hau² nai³* 'this earth of ours' and a second noun phrase *sii¹ cun²* 'four parts'.

202) လꩠက္တ္တꩮꩼ�won့ꩻ   ꩻꩻ   ꩻ   ꩻ   ꩻ

    pha¹ ta¹ wii¹   hau²   nai³   sii¹   cun²
    earth   1Pl   this   four   part

    ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ‖

    sii¹   cun²   nai³   u¹ taŋ¹ ku¹ ru¹   nai³
    four   part   this   Utangkuru   this
    'This earth of ours has four parts, and one of those parts is Utangkuru.'

    (Aiton text, *Why Buddha was born in this world*, No. (2), told by Sa Cham Thoumoung.)

In both the first line of (202), which is a verbless existential sentence, and the second line of (202), which is a verbless copula sentence, the first noun phrase (copula subject) is marked by a demonstrative.

There are a few examples where *yaŋ²* appears to have two arguments, as in (203). This construction indicates possession:

203) ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ꩻ ꩻ

    puu¹   siŋ¹ phaa⁴   naŋ¹ kau¹   yaa⁵   nɛ⁴   yaŋ² sa¹ nɛ⁴
    grandfather Singpha   in addition   grandmother   DEF   have PRT

    ꩻ   ꩻ   ꩻ

    me²   mum²   sa¹ nɛ⁴
    wife   3Sg   PRT
    'In addition Grandfather Singpha had Grandmother (as) his wife.'

    (Tai Khamyang Manuscript, *The book of calling back the Khon*, No. (5.1), read by Sa Myat Chowlik)

The core existential sentence of (203) is as analysed in (204), and is then followed by a postposed NP, *me² mum²* 'his wife', which confirms this reading.

204)

1st participant (possessor)	LINK	2nd participant (possessee)	V
puu¹ siŋ¹ phaa⁴	naŋ¹ kau¹	yaa⁵ nɛ⁴	yaŋ²
grandfather Singpha	in addition	this grandmother	have

Sentences like (203) are rare. Generally possession is expressed by a construction such as (205), which is literally 'her one child there was'.

205) ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ꩻ   ꩻ ‖

    luk³   man²   kɔ³   luŋ²   yaŋ²   (0'5)   luk³   nan³   phu³   yiŋ²
    child   3Sg   CLF   one   have   child   that   person   girl
    'She had a single child, a daughter.'

    (Aiton Story, *Story of the old woman*, No. (4), told by Nabin Shyam).

Where the item being possessed is an alienable item, such as money, the construction in (206) is possible.

206) ꩻ   ꩻ   ꩻ   ꩻ   ꩻ ‖

    tī⁵   mün²   sām⁶ sip¹   trā²   yaŋ²
    at   3Sg   thirty   rupee   have
    'He/she has 30 rupees.'
    (Phake sentence elicited from Aithown Che Chakap).

Aithown Che Chakap indicated that any possessive sentences using *tī⁵* as in (206) would be ungrammatical or at least strongly dispreferred if the possessee was something inalienable like a relative, spouse or body part. The preferred syntax of possessive sentences is summarised in (207)

207) Alienable possession → tii²/tī⁵ POSSESSOR POSSESSEE yaŋ²
      Inalienable possession → POSSESSEE POSSESSOR yaŋ²

8.5.3 Definitional sentences

In example (193) above, we saw that the relation of naming persons was achieved in the Tai languages by use of a verbless copula, where the copula subject was headed by the noun *cur²/cü⁵* 'name'. An alternative way of naming people is to use the verb *waa²/wā⁵* 'say', as in (208):

208) ကော့ဃသိဘိကော့   ၆ိ၆   သ၁၆   ဟရ   ၆ိ၆   ကိမ့်   ၁   ၍ ၊

tɔ¹ sii¹ nai³ kɔ³   muu²   nan³   caa²   mum² aai³ maau² pai² waa²
in spite of this   time   that   to   3Sg   Ai Mau   go   say
'However, at that time he was called Ai Mau.'

(Aiton text, *History of the Aiton*, No. (15.5) told by Nang Wimala Thoumoung)

In (208) the entity to be named is *myn²*'3Sg' and is marked in a prepositional phrase. The name to be given is *aai³ maau²*. A similar structure is used when things are named, as in example (209), where *pɔk¹ khau³*'the skin of rice' is the entity to be named and *kyɛp⁴*'chaff' is the name given.

209.1) ၆ိ၆   ၀၍   ၆ိ၆   ကျ၆   မ၆   ၍

pɔk¹   khau³   mün²   (...)   kyɛp⁴   mun⁵   wā⁵
skin   rice   3Sg   chaff   dust   say
'It is the skin of rice, we call it chaff dust'

209.2) ၍၆   ၆ိ၆   ၀၍   က   ၍   ကျ၆ ၊

haŋ²   pɔk¹   khau³   nɛ⁴   wā⁵   kyɛp⁴
to   skin   rice   PRT   say   chaff
'The skin of rice is called 'Chaff''.
(Phake sentence, uttered by Ee Nyan Khet)

In (209.1) there was a noticeable pause[68] between the words *mün²*and *kyɛp⁴*. This is necessary to separate the two noun phrases *pɔk¹ khau³ mün²*and *kyɛp⁴ mun⁵*, as there is no prepositional marking. It was not necessary to have a pause in (209.2), however, because one of these noun phrases, in the slightly reduced form of *pɔk¹ khau³*'skin of rice', the entity to be named, is marked in a prepositional phrase. The name to be given to this entity then follows the verb.

When the entity to be named, the 'known' or reference object, is marked in a prepositional phrase, it is marked with *haŋ²*or *caa²/cā⁵*, in the same way as the marking of non-agent animate participants in core prepositional phrases (see above 8.3.3). This is the only circumstance in which inanimate noun phrases can be marked with these prepositions.

---

[68] Unfortunately the length of the pause is not know. This sentence was not recorded; it was spoken during the process of translating a story.

Sometimes, in definitional sentences, the verb is omitted altogether, as in (210):

210) ကိ၆   ၍ာ၆   က   ၍၆   ၆၆ၮၢ

ə²   ñān¹   nɛ⁴   hüŋ²   piŋ¹ ñā¹
YES   knowledge   DEF   to   knowledge
'Yes, *ñān²*means knowledge.'

(Tai Phake Song, *Poem in the khe² khyāŋ²style: How Stephen Morey came to the Tai Phakes*,. No (10.3), read by Ai Che Let Hailung)

That *piŋ¹ ñā¹*is the entity to be named or reference object is clear from its being in the prepositional phrase. The name it is to be given is *ñān¹*'knowledge'.

8.5.4 One participant events

Predications with one participant can have as the head of the verbal complex either an intransitive verb, as in example (211) or an adjective, as in example (212):

211) ၍၆   က   ကော့   ၀၍ ၊

hə²   nɛ⁴   kɔ⁴   khom³
boat   this   LINK   capsize
'The boat capsized'

(Phake Story, *The Dolphin, the Crow and the Mosquito*, No (29), told by Aithown Che Chakap)

212) ၀ျ၀၆   က   ၀ၻ၆   ၀၆၆ ၊

pa¹ sāu⁶   nɛ⁴   sɔp¹   phɛu⁶
girl   DEF   beautiful   excessive
'That girl was very beautiful.'

(Phake Story, *Story of the foolish king*, No. (19), told by Ee Nyan Khet)

Sometimes a one participant event can have more than one predication, as in (213):

213) ၀၆ကော့   ၁ဲ   ၀၆၆   ၁ဲ   ၍   ၍၆   ၀၆၆   ကော့   ၀၆   ၀ၽ၆ ၊

taŋ¹ kaa¹ nam³ lik³   nam³ nɔi¹   hɛŋ³ cuum²   kaa¹   cam³   ȝau³
all   river small   river small dry   evaporate GO   PRT   FINISHED
'All the small rivers have dried up and evaporated.'

(Aiton manuscript, *The creation of the world*, No. (8))

In example (213), the main predicate is the adjective *hɛŋ³* 'dry'. The verb *cuum²* is arguably a completive in this case, or alternatively this is a serial verb structure.

There are many sentences in the Tai languages in which only one participant is stated, but in which there are one or more understood. These are to be understood as two or three participant events, depending on the number of participants that are not expressed.

8.5.5 Two participant events

The prototypical two participant event has an agent and a patient, as in (214):

214) ၍ လ �034 လ 0J 00E 804 J0 ą လ ‖
    hā¹   nɛ⁴   sɔ⁶   nɛ⁴   kāp⁴   tüŋ¹   wɔt⁴   haŋ²   mü⁶   nɛ⁴
    time   DEF   tiger   DEF   bite   all   PRT   to   pig   DEF
    'The tiger ate the pig up.'
    (Phake Story, *Story of Deception*, No. (39), told by Ee Nyan Khet) .

As discussed above in detail in 8.3.3.1, the patient, if animate, is often marked in a prepositional phrase headed by *haŋ²* or *caa²/cā⁵*, although this marking is not obligatory.

8.5.6 Three participant events

The most common way of expressing three participant events is with the second animate participant, the prototypical beneficiary, in a prepositional phrase, as exemplified above in 8.3.3.2. A ditransitive strategy, although rare, is attested and verb strings or verb serialisation can also be used. In the Tai languages, ditransitive sentences have the following structure:

215)   NP   V   NP   NP
       S   V   O   I
       Agent   Predicate   Theme   Beneficiary

If neither animate participant is marked in a prepositional phrase, then only the order in (215) appears to be permitted, as in example (169) in section 8.4 above. This type of sentence is quite rare in the Tai languages, but within a few lines, Ee Nyan Khet repeated this sentence with the beneficiary marked in a prepositional phrase, as in (216):

216) ၐ္ဇ   ကျ   ဟၟ   ၐ္ဇ   ၸၟ   ၈   ၮ   ‖
    kau²   kā¹   soŋ¹   khau³   haŋ²   phī⁶   nam³
    1Sg   go   send   rice   to   spirit   ancestor
    'I am taking rice to the ancestor spirits,'
    (Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (22), told by Ee Nyan Khet)

The story *Chaw Mahosatha meets Nang Pingya*, although told as an oral narrative, is based on a large Tai manuscript, itself based on a Pali text (See Morey 2001b). It may be that the syntax found in (215) is an archaic form, because the only other examples of this are also from manuscripts. Example (217) is taken from *The Creation of the World*, an Aiton manuscript several hundred years old, which is not yet fully translated.

217) ၏ၟၐ္ဇၮ   ၮ   ၸ   ၮ   ဟၟ   ၐ္ဇ   ၮၸ္ဇ   ၸၟၮ   ‖
    mɔ tɔk   phuu³   luŋ²   maa²   suŋ¹   hau³   maak¹ khɯ¹   cau³ phaa³
    ??   person   one   come   send   give   brinjal   king
    'A person came to give a brinjal to the king.'
    (Aiton manuscript, *The creation of the world*,)

Unfortunately, some aspects of this sentence remain unclear, and need to be rechecked with native speakers.

Where the agent participant is not expressed, and yet the beneficiary is not marked in a prepositional phrase, it is arguable that the syntax is the same as (215) and (216) with the dropping of one noun phrase. (218) and (219) occur in the *Tai Aiton History*, a manuscript dating back at least 100 years (Morey 1999b). In the passages presented below, King Tai Pum is dividing his country between his heirs, Chaw Tai Lung and

Pa Miu Pum, who are therefore the beneficiaries. In these examples, neither beneficiary has been placed in a prepositional phrase, and both beneficiary NPs appear in front of the verb.

Only when the agent is not expressed is it possible for a beneficiary unmarked in a prepositional phrase to occur in front of the verb in a three participant event. The beneficiary role is then assigned to the first expressed animate participant. Thus the order IVO becomes possible, even although there is no preposition:

218) ဟွင်္တဝ္ဝဝ္ရုင်္   ဗ်ဝဲ   ၍   ကဳဝ်   ရိင်   ဟွှ်ယ်   ကိ် ။

cau³ tai² luŋ¹   cam³   hauɯ³   kin²   :muŋ²   yun²   kɔi³
Chau Tai Lung   PRT   give   eat   country   Yun   PRT
'He gave Chau Tai Lung the country of Muang Yun to rule[69].'

219) ၆   ဗ်ဝဲ   ဟဝ်စဝ်ရုံ   ၍   ကဳဝ်   မွဝ်လုင်   ၯုဝ်   ဝေ့ ။

ŋii²   cam³   pa miu pum   hauɯ³   kin²   maau²   luŋ¹   put³   pɔ²
2nd   PRT   Pa Miu Pum[70]   give   eat   Mau Lung   inherit   father
'The second son, Pa Miu Pum was given Mau Lung to rule in place of his father.'
(Aiton manuscript, *History from the time of the ancestor Chaw Tai Lung up until Sukapha*, Nos. (7) & (8))

Another strategy for expressing three places is to use a serial verb strategy. This is discussed in detail below in example (289) in section 8.5.9.

8.5.7 Tense, Aspect and Modality

8.5.7.1 Preliminary Theoretical Considerations

Tai languages are usually categorised as being 'tenseless'. Comrie (1985:9), defines tense as a "grammaticalised expression of location in time.", and adds that:

> "Indeed, given that no restrictions are placed by the definition on what kind of location in time is to be considered, it is probable that most of the world's languages will turn out to have tense,..." (Comrie 1985:9)

---

[69] The phrase 'eat country' is a common euphemism for 'to rule a country'.
[70] The correct pronunciation of this King's name is no longer known. Tone marks have therefore not been notated

In discussing the difference between tense marking, a form of grammaticalisation, and lexicalisation, Comrie observes that "grammaticalisation refers to integration into the grammatical system of a language, while lexicalisation refers merely to integration into the lexicon of the language." (1985:10). Full grammaticalisation requires two criteria, obligatory expression and morphological boundedness. As he concludes: "The clearest instances of grammaticalisation satisfy both these criteria (they are obligatory and morphologically bound), the clearest instances of lexicalisation satisfy neither, while there will be many borderline cases which the criteria do not assign unequivocally to grammaticalisation or lexicalisation."

By this definition, for example, the morpheme kaa^1/ka^1 (see below 8.4.7.5.1) is a borderline case, but it does often behave as a past tense marker.

Many of the languages of Southeast Asia are categorised as having aspect rather than tense. Comrie (1985:51) discusses Burmese in some detail, contrasting the realis particle *te/tha/ta/hta* with the irrealis particle *me/ma/hma*. In Comrie's discussion of Burmese particles, the realis can be seen as an indicator of non-future, as in (220.1), whereas the irrealis must be used in the future as in (220.2)

220.1) Saturday-every grass cut *te*   '(he) cuts the grass every Saturday'
220.2) tomorrow begin-*me*   '(we) will begin tomorrow'

The irrealis particle can be in present or past time "provided the reference is not restricted to our actual world", as in (221)

221)   tamarind:fruit eat-ever-*me* think-*te*   '(I) think (he) must have eaten tamarinds before'

Comrie concludes that"

> "What Burmese shows us, then, is a language where time reference per se is not grammaticalised, i.e. there is no tense. It is, of course, possible for time

reference to be expressed in other ways (for instance lexically, by the used of adverbials like ... 'tomorrow'" (1985:51)

The situation described for Burmese is not the same as the situation that will described below in Tai (see 8.5.7.3). The morpheme *kaa¹/kā¹* cannot be regarded as an aspect marker, because it does not appear to "view the internal temporal constituency of a situation", the definition given aspect by Comrie (1976:3).

Foley makes the distinction between aspect and tense clear:

"Aspect says nothing about the relationship between the time of the event described by the predicate and the time of the speech act: ... Aspect delimits the predicate: is it extended or not? Tense delimits the whole sentence by situating it in time with regard to the present moment of speaking: does it precede, overlap with or follow the moment of speaking?" (1986:142)

By this definition, *kaa¹/kā¹* is not aspect, since in most cases its presence tells us simply that the sentence being uttered preceded the moment of speaking.

Diller (1992:24-25) very concisely discusses the marking of Aiton verbs for time. He discusses the marker *kaa¹/kā¹*, pointing out that it:

"occasionally occurs as a main verb in Aiton, but more often it ... occurs postverbally to indicate several time-related factors, most commonly past time or possibly perfective aspect... In this function, articulation sometimes approaches postclitic status." (1992:24)

8.5.7.2 Previous studies of the Tai languages

8.5.7.2.1 Needham

A traditional view of time marking in the Tai languages comes from Needham (1894), a grammar with several texts and a sixty page English to Khamti vocabulary. The remainder of this section is taken from Needham's discussion of the verbs of Khamti, by reference to the 'conjugation' of the Tai verb *kin* 'to eat' (1894:40-41).

CONJUGATIONAL EXAMPLE:

Root verb ꩡꩦ = (kin) to take food, or liquid, Assamese *khau*.

INDICATIVE MOOD

Present Tense
The bare root of the verb is used to form this tense.

Present Indefinite Tense	
The present definite is formed by adding ꩦ (û) to the root as:	

Singular		Plural	
kau kin û 1 3 2	= I am eating 1 2 3	tû kin û 1 3 2	= We are eating 1 2 3
maü kin û 1 3 2	= You are eating 1 2 3	sû kin û 1 3 2	= You are eating 1 2 3
man kin û 1 3 2	= He is eating 1 2 3	khau kin û 1 3 2	= They are eating 1 2 3

This tense may also be used by adding ꩦ = sî participial suffix to the root followed by ꩦ (yang) ((sic.)), as :

khau kin si û ' 1 3 2	= He is eating (and has not yet risen) 1 2 3

Past Tense
'The past tense suffix is ꩦ (kâ), though ꩦ (mâ) is often used to mark it.'
Example from paradigm:

kau kin kâ 1 3 2	= I ate 1 2

Examples showing the use of ꩦ (mâ):

man mâ mâ tai khâng kau mai 1 2 3 5 4	= He came near (to) me. 1 2 3 4 5

Perfect Tense

'This tense is made by suffixing ဃ္ (yau), a particle denoting *completion*, to the past tense suffix, as –'

ကှ် ကိှ် ကၞ ဃ္	= I have eaten
	1 2 3
kau kin kâ yau	
1 3 2	

'ဃ္ (yau) may be and often is used *alone, i.e.,* without ကၞ (kâ), as:

သိပ် ဝန် ထိုန် ဃ္	Ten days ago
	1 2 3
sip wan thôn yau	
1 2 3	

Future Tense

'This tense is made by prefixing the particle တိ (tî) to the root as:

ကှ် တိ ကိှ်	= I will eat
	1 2 3
kau tî kin	
1 2 3	

'တိ (tî) may be termed the *indefinite* future prefix particle. The present tense suffix ကၞ (û) is often used to express *future* time as:

ဖော် ကှ် မၞ ကၞ စမ် ကှ် မိ ဟှစ် ၡၞ ကၞ	When my husband comes,
	1 2 3 4
phô kau mâ kâ cham kau mai het châ û	(he) (will) scold me
3 2 4 6 5	5 6

'ဃ္ (yau) the *completive* particle may be added and often is used in conjunction with the future tense prefix တိ (tî) as:

ကှန် မန် ကၞ တိ စမ် ချၞ ဃ္	His men also will (be) inconvenienced.
	1 2 3 4 5
kôn man kaw tî cham chaü yau	
2 1 3 4 5	

Other particles discussed and exemplified by Needham include:

- နၞ (nam) the particle denoting *certainty* or *affirmation (vide* page 37 *ante)* is also used *in conjunction* with the future tense prefix တိ (tî)
- A *definite* future is made by suffixing ကၟတ် (kat) to the root
- *Physical power* or *capacity* is expressed by adding the word ဝၞ = (pen) *able* to the root.
- Conditional sentences are constructed by means of participles. The *conditional* participial suffix is ကၞ ဃၟ (kâ yê), ကၞ ၡၟ (kâ cham) is also used as a *conditional* participial suffix
- The imperative suffix proper is ဏၞ (tâ). It is only used for the second person singular and plural

- ၡၟ (haü) = *permit, allow give,* is used as an auxiliary *causal* imperative, indicating authority
- The *prohibitive* particle is ၡ (pî) prefixed to the root verb.

8.5.7.2.2 Banchob

A different approach was followed by Banchob (1987). The following words are found in Banchob referring to the marking of time:

Phake word		Gloss	Reference in Banchob
ဃၟ	$sɛ^6$ / $sɛ^6$	a word denoting past participle	(1987:133)
ဏၟ, ဏၟက်	ta^1, tak^1	shall, will, must	(1987:163)
ၡၟ္မ်	$nüm^1$	a final particle denoting the near future statement	(1987:236)
ဃၟ	ya^2	a final particle denoting the present continuous tense	(1987:336)
ဃ္	yau^4	finished, a word denoting the past tense	(1987:347)
ကၞ	$ü^1$	an emphatic word	(1987:427)

Table 28: Words expressing tense and time from Banchob (1987)

Banchob exemplified these as follows:

$sɛ^6/sɛ^6$

222) မၞန် ကၞ ဃၟ ၡၟက် မၞ
$mün^2$ $kā^1$ $sɛ^6$ $pauk^4$ $mā^2$
3Sg go PRT return COME
'Having gone (somewhere), he came back.'

ta^1

223) ကှ် ဏၞ ကၞ မၞန်
kau^2 ta^1 $kā^1$ $mān^1$ $mān^1$
1Sg WILL go surely
'I must go definitely.'

nüm¹

224) ကဵ ကိၚ် နုမ်

kau²   kin²   nüm¹
1Sg   eat   PRT
'I am going to eat now.'

ya²

225) ကဵ ဟဵတ် ယ

kau²   het¹   ya²
1Sg   do   PRT
'I am doing'

yau⁴

226) ဟဵတ် ယဵ က

het¹   yau⁴   kā¹
do   PRT   PRT
'It has been done.'

ū¹

227) ကဵ ကိၚ် အူ

kau²   kin²   ū¹
1Sg   eat   PRT
'I do eat it.'

8.5.7.2.3 Aimya Khang

The influence of Needham is very strong in present day Assam, probably because of the reprinting of key parts of Needham's analysis in Grierson (1904), with Grierson being the only widely available source on the languages. For example, the late Aimya Khang Gohain sent me the following: (pers. comn. 23/6/1999)

ဝိအူ	sī⁶ ū¹	present continuous tense
ကိၚ်ဝိအူ	kin² sī⁶ ū¹	is / has been eating
ကဝိအူ	kā¹ sī⁶ ū¹	is / has been going

Table 29: 'Present Continuous Tense' according to Aimya Khang Gohain

Aimya Khang Gohain also presented two further examples with translations into both English and Assamese, which are detailed in Table (30). The Assamese for the first example in Table (30) is described in Assamese grammars, such as Dasgupta (1993), as 'present perfect tense', whereas the second is 'past indefinite tense'.

ၜမ္မ	ā² mā² kā¹	The aunt came (and still is).	পেহী আহিলে
ၜမ္ဝ	ā² mā² wai⁴	The aunt came (but had left).	পেহী আহিছিল

Table 30: Two examples of past tense after Aimya Khang Gohain

In Table (30), the Assamese translations are not the same as the English gloss.

8.5.7.3 The present investigation

In this study, the TAM morphemes will be divided into groups according to their syntax. The first are those which precede the verb, of which there is only one, *tak¹/ta¹/ti¹* 'WILL' (see 8.5.7.4.1). These will be referred to as TAM₁ markers.

Several TAM morphemes, *kaa¹/kā¹* 'GO', *maa²/kā²* 'COME' and *wai³/wai⁴* 'KEEP' can both immediately follow the verb, or be placed at the end of the core sentence. These are discussed in 8.5.7.5 and will be referred to as TAM₂ markers.

A number of morphemes occur only at the end of the core sentence. These are discussed in 8.5.7.6 and will be referred to as TAM₃ markers. A number of other morphemes with TAM functions are discussed in 8.5.7.7.

8.5.7.4 TAM morphemes which precede the verb (TAM₁)
8.5.7.4.1 tak¹ / ta¹ / ti¹:

Diller (1992:24) observed that *ta¹* (which also occurs in the corpus as *ti¹* and *tak¹*) "appears to cover many of the auxiliary functions of Central Thai *ca:* irrealis or future

marking, sometimes involving intention or volition ..., normally without accomplishment."

Comrie (1985:45), discussing the future tense in German pointed out that the present tense is usually used for future events, such as *Ich gehe morgen* 'I will go tomorrow', and the fut construction, *ich werde gehen* 'I will go', would only be used "where there would otherwise be a danger of misunderstanding in the direction of present time reference". This has led linguists to posit a PAST~NON PAST dichotomy in such languages

In Tai, on the other hand, this does not appear to be so. All future events, except for some in the negative (see (231) below) are obligatorily marked with the morpheme *tak¹/ta¹/ti¹*.

The marker *tak¹/ta¹/ti¹* is glossed in this thesis as 'WILL'. In manuscripts it is usually written ဆက် (*tak¹*) but it is more often pronounced as *ta¹* or *ti¹*. In oral texts, where it has been pronounced *ta¹* or *ti¹*, I have usually written it as it is pronounced. The late Aimya Khang (pers. comm.) criticised me for this, stating[71] that:

> "The pre-position word ဆက် (*tak¹*) in future tense should also be corrected and should be written in literary way as in old Puthis, as ဆက် (*tak¹*) and not တိ (*ti¹*) ... At most we may shorten to ဆကွာ (*ta¹ kā¹* 'WILL-go') not 'တိကွာ', which may mean 'place to go' (*tī⁵ kā¹*)"

The prototypical use of *tak¹/ta¹/ti¹* is to express future time, as in (228). It is very frequent in conversation and seems to be obligatory wherever future time is meant:

228) ပိ   �212   ဆိုင်   မ   ကိ   ဝည်   တ   ဦ

pī²   nā³   süŋ⁶   mā²   kī¹   wan²   ta¹   ü¹
year   next   if   come   how many day   will   stay
"If you come next year, how long will you stay."
(Phake sentence, spoken by Yehom Buragohain)
Audio link for this sentence (\wave files\aiton\extracts\2-9-2-1\87.mp3)

---

[71] I have added phonemicisations and some glosses in parentheses for the sake of clarity.

Following Diller's characterisation of *tak¹/ta¹/ti¹* as being sometimes intentional or volitional, example (228) could be translated as 'How long do you intend to stay'. When reaching a village, the hosts will often ask *kī¹ wan² ta¹ u¹*, and this arguably implies intention or volition. Some examples of *tak¹/ta¹/ti¹*, on the other hand, do not involve intention, as the in (229):

229) ပါ ရုက် ၃   တ   ဝိတ်   ဝစိတ်   ဝည်   ဝိ   ဝတိ   ဒီကွာ   ။

pa¹ ruk¹   nai³   ta¹   pin²   cit¹   wan²   ta¹   taai²   dii² ɔ¹
tomorrow   this   WILL   be   7   days   WILL die   PRT
'Tomorrow it will be the seventh day, and we will die.'
(Aiton Story, *The twelve questions*, No (62), told by Sa Cham Thoumoung)

Diller (1992:24) indicated that *ta¹* was "normally without accomplishment". In example (229), there are two tokens of *ta¹*, the first of which is in the predication 'tomorrow will be the seventh day'. This is something which was accomplished, and of which the accomplishment was certain, for the sentence was spoken on the sixth day and the seventh day will certainly follow.

The second use of *ta¹* in (229), on the other hand, is in the predication 'we will die'. This was not accomplished in the story of *The twelve questions*, because those who spoke this question did in fact survive.

The marker *tak¹/ta¹/ti¹* can also be used with the negative, as in (230):

230) ဟ   တ   မ   ယ   ဝ   နုင်   ၃

hau²   ta¹   ma¹   yam⁶   wā⁵   nüŋ¹   nai⁴
1Pl   WILL   NEG   wet.NEG say   like   this
"We will not get wet.' He spoke like this.'
(Phake Story, *The birth and early life of Chaw Mahosatha*, No. (29), told by Ee Nyan Khet)

However with negatives its use appears to be non-obligatory, as shown in (41), where the positive *nai³* 'get' is preceded by *ta¹*, but the negative *nai³* in the second clause of the sentence is not:

231) ꩦꩥ ꩯ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ॥
luk⁴ yɛ⁴ ŭ³ ta¹ nai³ mɛ⁵ ma¹ nai³ wã⁵ nũŋ¹ nai⁴
child PRT father WILL get mother NEG get say like this
'"As to this chick, the father will get it, the mother will not", he spoke like this.'

(Phake Story, *Story of the kum⁴ bird*, No. (62), told by Ee Nyan Khet)

The marker *tak¹/ta¹/ti¹* is not confined to future meaning, and also expresses the English 'should' as in (232):

232) ꩦ ꩦ ꩦ ꩦ ꩦ English ꩦ
soi³ ɛ¹ ka³ saŋ⁶ ta¹ wã⁵ English soi³
jealous HESIT what will say English jealous
'*soi³* (jealous), what will/should we call *soi³* in English?'

(Phake sentence, spoken by Yehom Buragohain)

Audio link for this sentence (\wave files\aiton\extracts\2-4-2-2\20-4.mp3)

8.5.7.5 TAM morphemes which may either immediately follow the verb or be placed at the end of the core sentence (TAM₂)

As indicated above in example (183), section 8.4.3, some TAM markers can immediately follow the verb, preceding the object noun phrase or any other non core phrase that follows the subject.

In other Tai languages, where the object is expressed, the TAM morphology will be separated from the verb and be placed utterance final, as in (233):

233) ฉัน ซื้อ ข้าว มาแล้ว
chán sứ: khâ:w ma: lé:w
I buy rice COME-ALREADY
S V₋ O (−V₋)

(Standard Thai sentence, after Vichin Panupong 1970:20)

This is called a discontinuous constituent by Vichin Panupong. She marked this constituent V₋ ... (−V₋), showing that there was a transitive verb (V₋) complex, interrupted by the object. Vichin regards the structure in (233) as a special case of the structure S V₋ O, because it would be possible to antepose the object, thus making the discontinuous constituent "continuous" (1970:20).

We can typify the pattern of (233) as in (234):

234) S   V   O   TAM marker

In (234), either TAM₂ or TAM₃ markers can occupy the TAM position. If the Object is moved to the preverbal position, as in (235), or is topicalised as in (236), or is elided for any reason, as in (237), then the verb will end up adjacent to the TAM marker.

235) S   (O)   V   TAM marker

236) (O)   S   V   TAM marker

237) S   V   TAM marker

It appears that among the Tai languages of Assam, the proximity of TAM markers like *kaa¹* to the verb or verbal complex is leading to a reanalysis of the position of this marker as being adjacent to the verb, as in (238):

238) ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ꩦ ॥
mən² aa³ paak¹ nɔi³ mən² pai² kap³ kaa¹ kɔn³ nɯ³
3Sg open mouth PRT 3Sg go bite GO piece meat
'He opened his mouth to grab that piece of meat.'

(Aiton Story, *Story of the Crow and the Fox*, No. (56), told by Ong Cham).

This sentence is of the pattern:

239) S   V   TAM₂ O

This process of reanalysing *kaa¹/kā¹* as a morpheme which is adjacent to the verb was first observed by Diller (1992:24) who wrote that "articulation sometimes approaches postclitic status." Since the reanalysis is not complete, and sentences with the pattern of (234) are still found, the morphemes discussed in this section are categorised as TAM morphemes which may either immediately follow the verb or be placed at the end of the core sentence, with the abbreviation TAM₂.

8.5.7.5.1 kaa¹/kā¹:

The TAM marker *kaa¹/kā¹*, a grammaticalisation of the verb 'to go', is very frequently used to mark past time, especially in narrative texts. Diller (1992:24) categorised it as indicating "several time-related factors, most commonly past time or possibly perfective aspect.". In example (240) it marks simple past time, indicating that the action of the verb is 'gone'.

240) ဝၢ မၢၵ် ၼႃ ၵၣ်ၵ် ဢုၵ်ၵ် ၵၣ်ၵ် မုၵ်ၵ် ၼ္ဂၢတ်လ္ၢ ၼ္ၢ ၼႃ ။
      caa² man² kɔ³ kun² suŋ¹ kun² muŋ² sɔt¹ phai² hau³ kaa¹
      to   3Sg   LINK person country person country burn fire GIVE GO
      'The countrymen burned his body.'
      (Aiton text, *History of the Aiton*, No. (71) told by Nang Wimala Thoumoung)

Sometimes, in a text with more than one verb in the sentence, all of those verbs will be marked with *kaa¹/kā¹* as in (241):

241) ၃ ၼႃ ၃ ၼႃ ၃ ၼ္ဂ ၼႃ ၼႃ မ္ၢၵ်ၵ္ၢ ။
      pai² kaa¹ pai² kaa¹ pai² uu¹ kaa¹ kɔ¹ mɔn¹ rai³
      go   GO   go   GO   go   stay GO   cluster name of a tree
      'He went, and went, and went and rested at a cluster of Mon Rai.'
      (Aiton Story, *Story of the monkey and the fox*, No. (15), told by Bidya Thoumoung).

In some texts it is marked both on the main verb and on the completive, as in (242), which is part of a story in which two brothers go out hunting. The first shoots and misses, but the second gets his prey:

242.1) ၎ၵ် ၼႃ ၽ္ ဢ္ၵ်ၵ် ၼႃ
        ʒɯ² kaa¹ ma¹ thuk¹ kaa¹
        shoot GO   NEG   touch GO
        'He shot but didn't hit it.'

242.2) ၎ၵ် ၼႃ ၃ ဢ္ၵ်ၵ် ၽ္ ၅ၵ် ၵ္ၵ် ၼ္ၢၵ်
        ʒɯ² kaa¹ pai² thuk¹ kaa¹ nuk³ kuk³ aṅ²
        shoot GO   go   touch GO   bird kuk³ CLF
        'He shot and hit the *kuk³*bird.'
        (Aiton sentences spoken by Sa Cham Thoumoung)
        Audio Links for these two sentences are found at \wave files\aiton\extracts
        \2-1-2-2\a.mp3 and b.mp3.

When used with adjectives, the TAM marker *kaa¹/kā¹* suggests ingressive meaning, as in (243) (see also above example (52) in section 8.2.4).

243) ၵ္ၢၵ် ၎္ၵ် ဝၥ်ဝၵ် ၼႃ ၎္ၵ်
      haŋ² hau² ʒaap¹ ʒaak³ kaa¹ ʒau³
      to   1Pl   difficult   GO   FINISHED
      'It became very difficult for us.'
      (Aiton text, *History of Barhula*, No. (35) told by Ong Thun

This ingressive meaning is similar to that observed with the perfective marker in combination with adjectives in Chinese (Comrie 1976:20). However the TAM₂ marker *kaa¹/kā¹* is not a perfective, which in these languages is *yau³/yau⁴* (see below 8.5.7.6.3).

Even though *kaa¹/kā¹* often appears as past tense marker, it is not bleached of its original meaning of 'go'. It is glossed as 'GO', to show the original meaning; capital letters are used to show that some grammaticalisation has occurred. The same morpheme *kaa¹/kā¹* can also be used as a full verb, or in a serial construction with other motion verbs (see example (288) in 8.5.8 below). An related imperative form *ka³* is also found (see below 8.6.4). Example (244) illustrates *kā¹* in a serial construction and *ka³* as an imperative.

244) ꩡ꩟ꩠ ꩡꩢ ꩣ ꩡꩢ ꩤ ꩥ ꩦ ꩧ
    phə¹ se⁶ nai³ yɛ⁴ kau² ta¹ nɔn² haɯ³
    despite PRT get PRT 1Pi WILL sleep GIVE

    ꩨ ꩩ ꩪ ꩫ ꩬ ॥
    maɯ² wɛn⁶ khãm² kã¹ ka³
    2Sg jump across go GO.IMP
    'In spite of this quarrel, I will lie down and you jump over me.'
    (Phake Story, *The story of the two goats*, No. (12), told by Yehom Buragohain)

8.5.7.5.2 maa²/mã²:

The TAM marker *maa²/mã²* is used to mark past time, just like *kaa¹/kã¹* above, but in the specific context of the action referring towards the speaker or place where the speaker is. For example, in (245), Lord Sikkya (Sakka) is in heaven looking down to the earth where the teller of the story is. In this circumstance, *maa²/mã²* is used.

245) ꩭꩮ ꩯ ꩰ ꩱ ꩲ ꩳ ꩴꩵ ॥
    lai³ pɯ² nai³ cau³ sik¹ kyaa² duu² lɛm² maa² muɯŋ² kaaŋ²
    therefore RESP Sakka look look at COME the earth
    'And so Lord Sakka looked (down) at the earth.'
    (Aiton text, *History of the Aiton*, No. (6.1) told by Nang Wimala Thoumoung)

Since the full verb *maa²/mã²* 'come' is always indicating motion towards the speaker, it will usually be marked by the TAM marker *maa²/mã²* in situations of past time, such as in example (246):

246) ꩶ ꩷꩸ ꩹ ꩺ ꩻ ꩼ ꩽ
    luk⁴ thai¹ lɛn¹ khī¹ laiŋ⁶ men² mã² mã²
    from Thailand ride vehicle fly come COME
    'He came from Thailand by aeroplane.'
    (Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to the Tai Phakes*, No (26.1), read by Ai Che Let Hailung)

This usage of *maa²/mã²* is also found in historical manuscripts, such as (247), from the Phake chronicle. The TAM marker *maa²/mã²* is used here because the movement is both towards the present time, and, gradually, towards the place where the

manuscript was written, Assam. This chronicle tells of the movement of the Phake people towards Assam.

247) ꪀ ꪁ ꪂ ꪃ ꪄ ꪅ ꪆ
    kin² khau³ khāu⁶ māu² (loŋ⁶) sauŋ² tīn² sāi¹ nam⁴
    eat rice white Mau (Lung) two feet wash

    ꪇ ꪈ ꪉ
    māu² loŋ⁶ mã² cam⁴
    Mau Lung come PRT
    'He ate the white rice of Mau Lung and washed his feet in the Mau Lung river.'
    (from the Phake Chronicle)

A similar phenomenon is observed in Aiton, as shown in (248), where King Sukapha is making his way, gradually, to Assam. In the narrative at this point, he had battled with the Nagas, won the battle and taken their women to be wives for him and his men. This done, (*yau³ maa²*), sentence (248) is uttered.

248) ꪊ ꪋ ꪌ ꪍ ꪎ ꪏ ꪐꪑ ॥
    ʒau³ maa² maa² thuɯŋ¹ tii² doi² pat¹ kai³
    FINISHED COME come reach place mountain Pat Kai
    'That being done, (they) reached the Pat Kai mountains.'
    (Aiton text, *History of the Aiton*, No. (42) told by Nang Wimala Thoumoung)

Sometimes, the TAM markers *kaa¹/kā¹* and *maa²/mã²* are interchangeable. Example (249) was the version in the manuscript of this text written out by Bidya Thoumoung. When Nang Wimala discussed this text with me, she indicated that in her version, *kaa¹* 'GO' was used in place of *maa²* 'COME' in the first line, but that this made no difference to the meaning.

249) ꪒ ꪓ ꪔ ꪕ ꪖ ꪗ ꪘ ꪙ ꪚ
    duu² han¹ maa² cau³ phaa³ ʒaau² ai³ taŋ¹ kun² suŋ¹
    look see COME king Yau PRT all person country

ကုၟ်   ၛိင်   ကုၟ်ောင်   ယူတၟင်   ၟ   ကၟင်   ၟ   လၞ   ။

kun² mum² kup¹ taŋ¹ suu² tɔŋ² haɯ² kaaŋ³ haɯ³ laai¹
person country with   pray give wide give many

'He saw that king Yau and all the people were praying for more people.'

(Aiton text, *History of the Aiton*, No. (7) told by Nang Wimala Thoumoung)

The context of sentence is that it immediately follows (245) above. The scope of the verb *han¹* is therefore towards the speaker, and as a consequence *maa²/mā²* is appropriate. The fact that *kaa¹/kā¹* can also be used in this circumstance strongly suggested that *kaa¹/kā¹* is the more generalised marker and that it is bleached of its directionality. On the other hand, *maa²/mā²* is used only when the action of the verb is towards the speaker.

A very good example of the difference between *kaa¹/kā¹* and *maa²/mā²* is (250), where there are two clauses in the past tense, the first marked by *maa²* and the second by *kaa¹*.

250) ၟ်ကုၟ်   ကလင်   ယၟ်ကရ   ၟၟ   ၛိင်   ဗၟ   .

khum² kɔn¹ ka¹ laŋ¹ suı¹ kaa² phaa³ khɔt¹ mum²   maa²
also   after   Sukapha   leave country COME

လၞ   ပက်   ၜ   ကုၟ်   ၛိင်   ဗၟ်   ၟ

sii¹ paak¹ pii² kun² mum² maan² khau¹
4   100   year people country Burma 3Sg

ၟ   ၟ   ကၟ   ၛိင်   ကၟင်   ။

him¹ au² kaa¹ mum² kɔŋ²
snatch take GO   country Kong

'Four hundred years after Sukapha left his country, the Burmese took Muang Kong.'

(Aiton text, *History of the Aiton*, No. (73) told by Nang Wimala Thoumoung)

In (250), *mā²* can be used to mark the first clause, because the place to which Sukapha came, Assam, is the place in which the speaker is situated. The second clause, on the other hand, cannot be marked by *mā²*, because it refers to events in another place, which in no sense come towards the speaker.

As with *kaa¹/kā¹* the TAM marker *maa²/mā²* can be used with adjectives, and as with *kaa¹/kā¹* its meaning is ingressive. However, in contrast to *kaa¹/kā¹*, *maa²/mā²* with adjectives implies future time, (i.e. time that is to come) as in (251).

251) ၟ   ၟ်   ၟ   ၟ   ယၟ   ၟ

au² mum² phraa² traa² saŋ¹ khaa¹ caŋ¹ nai³
take glorious Buddha law monks   now

ၟ   ကၟ   ၜ   ဗၟ   ကၟ   ။

cɔn¹ cɔn¹ uu¹   dii²/nii² maa² kɔ³
little live good   come also

'By the grace of the Buddha, his teaching and the monks, little by little we will became well.'

(Aiton letter, No (11), written by Nabin Shyam Phalung, 1999)

The different functions of *kaa¹/kā¹* and *maa²/mā²* are summarised in Table (31), below.

8.5.7.5.3 wai³/wai⁴:

A third morpheme which is used with a past tense sense is *wai³/wai⁴*, which also occurs as a main verb meaning 'keep' (see example (25.2) in 8.2.2.3 above, for an occurrence of *wai⁴* as a completive verb meaning 'keep'). The past tense use of *wai³/wai⁴* is exemplified in (252):

252) ၟ   ၟ   ၟ   ၟ   ကၟ   ၟ   ၜ

cüŋ² nai⁴ cāi² sū⁶ tüŋ¹ lɔŋ⁶ kā¹ yeu⁵ wai⁴
now   male 2Pl all   go pass urine KEEP

ၜ   ၟ   ၟ   ၟ   ။

tī⁵ hɔn² kau² yɛ⁴
at   house 1Sg   PRT

'Now you have all come and urinated at my house.'

(Phake Story, *A story of justice*, No. (20), told by Sam Thun Wingkyen)

In (252) more is implied than simply an event which is in the past. There are longer term results which remain, and it is for this reason that *wai³/wai⁴* is found. In (252),

for example, it is the fact that the results of the urination remain around that causes the owner of the house to request compensation. In (253) the subject of the sentence wishes not only to be able to write the language, but to keep that skill.

253) ရှင်း   သဝ်   ကုံ   ဖို့   ထုက်   တမ်   ခ
    mün²   cam²   com² phaü¹   thuk⁴   tɛm³ wai⁴
    he   self   glad   desire, to be interested to copy   write KEEP
'He was glad and interested to write (our language).'
(Tai Phake Song, *Words of gladness - in honour of Stephen Morey*, No. (17)), sung by Am Saeu Khyo)

However, *wai³/wai⁴* is also used to express a past state which no longer exists, as in (254):

254) ရှင်း   ဟိတ်   ဟယ်ငြ   ခ
    mün² hit¹   sa¹ lā¹   wai⁴
    3Sg   do   teacher   KEEP
'He was a teacher.'
(Phake sentence)

Aithown Che Chakap stated that (254) should be literally translated 'At some time he was a teacher, but he is not now' and that the morpheme *wai³/wai⁴* was best translated as 'at some time past'.

It might be expected that the other TAM₂ markers *kaa¹/kā¹* and *maa²/mā²*, being verbs of motion as main verbs, could not be used as past tense markers on sentences such as (254) which are relating to states rather than actions. No example of *maa²/mā²* being used to mark past time on states has been recorded, but (255) shows *kaa¹/kā¹* in this role. This suggests that *kaa¹/kā¹* is a more generalised past tense marker than the other TAM₂ markers.

    ... ...ဝ်   ပဖ့ုဝ်   ကူ
    ၉...? cau³ mun² kaa¹
    he   monk   GO
'I was a monk.'

(Aiton text, *Explanation of the spirit house*, No. (27), told by Sui Khong Thoumoung.)

Some of the postverbal occurrences of *wai³/wai⁴* are analysed as completive verbs rather than TAM₂ markers. In example (256), the action is in the future so a past time interpretation is not appropriate.

256) ကူဝ်   သို့   ကာဝ် ၃ဝ် ဗိုဝ်ကူဝ်
    kun²   thau³ an²   nan³ khum¹ kɔn¹
    person old   CLF   that   again

    ဟို   ကော   ဝင်   ၃   ကာဝ် ကဝ် ၃   ဝို ။
    hau² kɔ³   saŋ¹   daa¹ kan²   phan² nai³ sii¹
    1PI   LINK   if   quarrel RECIP like   this   PRT

    ဟို   ကော   တ   မဝ်   ကို ခ   ဝို ။
    hau² kɔ³   ta¹   maak¹ kai³ kai³ wai³ sii¹
    1PI   LINK   will   rich   quickly   KEEP PRT

'That neighbour said, "If we quarrel like this we too will become rich very soon.'
(Aiton Story, *Story of the forest ghosts and the opium pipes*, No. (25), told by Mohendra Shyam).

Table (31) compares the use of the three TAM morphemes which can mark past time, among other functions:

	as TAM marker with verbs	as TAM marker with adjectives	other uses
kaa¹/kā¹	General marker of past time	ingressive meaning in the past	as a main verb meaning 'to go'
maa²/mā²	Marker of past time - used when the scope of the process indicated by the verb is in some way towards the speaker	ingressive meaning in the future	as a main verb meaning 'to come'
wai³/wai⁴	Marker of past time, used when the process indicated by the verb is intended to have long term results, or past marker with states	–	as a main verb and completive verb meaning 'to keep'

Table 31: TAM₂ markers in Phake and Aiton

8.5.7.6 TAM morphemes which occur at the end of the core sentence (TAM₃)

As discussed above in 8.5.7.5, there are TAM markers which can only occur at the end of the core sentence. These are $uu^1/\bar{u}^1$ 'STAY' (see 8.5.7.6.1), $suu^1/s\bar{u}^1$ (see 8.5.7.6.2), and yau^3/yau^4 (see 8.5.7.6.3).

8.5.7.6.1 uu¹/ū¹ and ɔ¹:

Needham (1894) stated that the morpheme $uu^1/\bar{u}^1$ was used to mark a 'present indefinite tense', see 8.5.7.2.1, which modern linguistics would categorise as continuous or progressive aspect. Banchob (1987), on the other hand, categorised $\bar{u}^1$ as being an emphatic particle. Diller (1992:24) reported that $uu^1/\bar{u}^1$ (together with $suu^1/s\bar{u}^1$) "function postverbally to suggest progressive aspect.". In both of the examples given by Diller, the progressive aspect is marked by $suu^1/s\bar{u}^1$.

Yehom Buragohain explained that in Phake $s\bar{u}^1$ and $\bar{u}^1$ had the same meaning. She translated $kau^2\ kin^2\ s\bar{u}^1$ (1Sg-eat-sū¹) as 'I am eating'. As a TAM marker, $uu^1/\bar{u}^1$ occurs frequently in elicited sentences, such as (257), where it marks the progressive, and (258), where it was stated to be emphatic:

257) က... ... ...

kon² haü³ pāp⁴ lik⁴ nai⁴ haŋ² kau² nai⁴
person give book this to 1Sg this

kin² māk¹ moŋ⁵ sī⁶ ū¹
1Sg mango PRT STAY
"The man from whom I received this book is eating mangoes."
(Phake sentence elicited from Aimya Khang Gohain)

258) ... ...

ɔ̃² caü ū¹ kat¹ kɔi⁴
YES YES STAY cold PRT
'Yes, its definitely cold.
(Phake sentence elicited from Aimya Khang Gohain)

However, in the texts that I have collected, $uu^1/\bar{u}^1$ has not been found in the progressive/continuous meaning, although it is found with emphatic meaning, often in combination with adjectives, as in (259):

259) ... ...

tai² kɔ⁴ mün² wə̃⁵ nī² ū¹
friend 3Sg say good STAY
'His friend said 'Yes, Very good.'
(Phake Story, *The blind man and the man with scabies*, No. (9), told by Ee Nyan Khet)

The word $uu^1/\bar{u}^1$ is very frequently found as a main verb meaning 'to stay', 'to live at' .It is therefore glossed 'STAY', showing that its verbal meaning is not fully bleached.

There is another particle, especially in Aiton, pronounced $ɔ^1$, which is probably the same morpheme. It is exemplified in (260):

260) ... ...

uı² hauı³ lɛ¹ kau² cau³ uıi²
YES GIVE PRT 1Sg RESP* EXCL

kau² cuu² ɔ¹ cit¹ wan² ‖
I agree STAY seven day
'Yes, my lord, I agree (to give them) seven days.'
(Aiton Story, *The twelve questions*, No (32), told by Sa Cham Thoumoung)

In (260), it appears that the meaning of $ɔ^1$ is emphatic, as in (258) and (259). Nabin Shyam stated that in Aiton, $ɔ^1$ was used to mark the present tense, as in $mun^2\ m\check{a}^2\ ɔ^1$ (3Sg-come-ɔ¹) which he translated as 'He is (still) coming'. Nabin added that the meaning of $ɔ^1$ and uu^1 were the same.

In example (261), $ɔ^1$ is found in conjunction with an adjective, with ingressive meaning:

261)

   saŋ¹   khɔn¹   pɔk³   maa²   ɛ¹   khai³   kɔ³   nii²   ɔ¹
   if   *khon*   return   come   PRT   illness   LINK   good   STAY

"If the *khon* returns, the illness will get better."

(Aiton text, *Explanation of Hong Khon ceremony*, No. (6), told by Nabin Shyam.)

It is not clear what the difference is between *khai³ kɔ³ nii² ɔ¹* in (261) and *khai³ kɔ³ ta¹ nii²*(fever-LINK-WILL-good).

8.5.7.6.2 suu¹/sū¹:

The verb *suu¹/sū¹* is variously translated 'to arrive' or 'to progress towards'. As a TAM marker it is glossed as 'TOWARDS', and is used for progressive aspect, as in (262):

262)

   mɛ⁵   kai¹   nam²   to²   nüŋ⁵   ŋam²   sū¹
   mother   chicken   black   body   one   cover   TOWARDS

"'At the pagoda there is a single black hen sitting on her eggs.'"

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (55), told by Ee Nyan Khet)

That this is a marker of aspect not of tense is shown in example (263), where it marks the progressive in past time. It is possible to read *sū¹* in example (263) as the main verb 'to arrive', rather than a TAM morpheme.

263)

   mün²   lɛ¹   sū¹   pā³ lan⁴   naü²   mɔ⁵   nan³
   3Sg   roam   TOWARDS   there   in   time   that

'She was away travelling there, at that time.'

(Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to the Tai Phakes*,. No (39.1), read by Ai Che Let Hailung)

Unlike the morphemes in 8.5.7.5 above, *suu¹/sū¹* always appears at the end of the core sentence, as in (264):

264)

   a¹ māt¹   nɛ⁴   au²   kā¹   ye⁴   kā¹   haü³   kin²   kā¹   nai⁴
   minister   PRT   take   GO   PRT   go   give   eat   GO   this

   haŋ²   khun⁶ hɔ⁶ kham²   sū¹   wā⁵   nüŋ¹   nai³   wüt⁴
   to   king   TOWARDS   say   like   that   PRT

"The minister took it and went to give it to the king to eat, it is said like that."

(Phake Story, *Story of the foolish king*, No. (42), told by Ee Nyan Khet)

In this example we clearly see *suu¹/sū¹* is in final position within the core sentence, whereas *kā¹* is attached to the verb complex *kā¹ haü³ kin².* The constituents of the core sentence are analysed in (265):

265)

Verb complex	TAM		Object	TAM
kā¹ haü³ kin²	kā¹	nai⁴	haŋ² khun⁶ hɔ⁶ kham²	sū¹
go and give ... to eat	GO	this	to the king	TOWARDS

Another example of the usage of *suu¹/sū¹* is in (266), in which Sa Cham Thoumoung is telling about the various migrations of the Aiton, and then moves into present time to refer to something that is the case now.

266)

   puun²   cɛ²   lau²   suu¹   kham² tii²   khau¹   lau²   suu¹   ɛ¹
   other   now   say   TOWARDS   Khamti   3Pl   say   PRT   PRT

   naa²   aai³ tɔn²   naa²   aai³ tɔn²
   field   Aiton   field   Aiton

'In the place that the Khamtis now call the Aiton fields,'

(Aiton text, *History of the Tai*, No. (26) told by Sa Cham Thoumoung)

8.5.7.6.3 ʒau³/yau⁴:

Diller (1992:24) categorised *ʒau³/yau⁴* as a temporal adverbial, translated as 'already'. An excellent example of *ʒau³/yau⁴* is (267), which is taken from a text telling about the death (Parinibbana) and cremation of the Buddha. Example (267) relates the

complete burning up of the Buddha's body. It is the only time in this text that *ʒau³/yau⁴* occurs.

267) ၛၖၘ ၚ ၖ ၛ ၖၖ ၖၖ ၖၖ ၜၖ ၖ ။
    muɪ¹ haai¹ waaŋ² naɯ² cuŋ² caŋ¹ haaŋ² tɔ¹ ʒau³
    gone gone inside in coffin then body until FINISHED
'Then the body inside that coffin was gone.'

(Tai Khamyang text, *The Parinibbana*, No. (32), read by Chaw Cha Seng)

From example (267) we can derive the analysis that *ʒau³/yau⁴* is TAM marker of perfective aspect. It is glossed as 'FINISHED' in this thesis.

Example (268) shows that an action marked by *ʒau³/yau⁴* is completely over before the next action commences:

268) ၛ ၖၖ ၖ ၛၖ ၖ ၖၖ ၖ ၖၖ
    waa² naŋ¹ nai³ ʒau³ mɛ² thau³ ɔk¹ pai² kaa¹
    say like this FINISHED woman old go out go GO
'When she had said this, the old woman went out.'

(Aiton Story, *Story of the old woman*, No. (14), told by Nabin Shyam).

Example (269) is from Ongthun Shyam's *History of Barhula*. He has been talking about the fact that in days long past, there was teaching of the Tai scripts in the village. The introduction of Assamese schools with Indian independence meant that only Assamese script was being learned, and that finished off the Tai script.

269.1) ၖၖ [ၖၖၖ] ၖၖ ၖၖ ၖ
    lik³ ma¹ num² ku¹ laa¹ sɔn¹ ʒau³
    book Assam foreigner teach FINISHED
'(Only) Assamese writing was taught.'

269.2) ၖၖ ၖၖ ၖ ၖၖ ၖ ၖၖ ၖ ၖၖ ၖၖ
    aa¹ luk³ mɯ² ʒaam² nan³ mɔk¹ lik³ tai² lik³ maan²
    from time time that since book Tai book Burma

ၖၖ ၛ ၖၖ ၖၖ ၖ ၖ
an² hau² sɔn¹ ɛ¹ ʒɔm² maa² sii¹
CLF IPL teach PRT get less COME PRT
'After that, the Tai and Burmese writing which we had learned, got less.'

(Aiton text, *History of Barhula*, No. (34) told by Ong Thun

The TAM morpheme *ʒau³/yau⁴* can also be used in sentence initial position to indicate that the item previously being discussed is completed and a new topic has begun, as in (270):

270) ၖ ၖ ၖ ၖၖ ၖ ၛ ၖၖၖ ။
    ʒau³ maa² maa² thɯɯŋ¹ tii² dɔi² pat¹ kai³
    FINISHED COME come reach place mountain Pat Kai
'That being done, (they) reached the Pat Kai mountains.'

(Aiton text, *History of the Aiton*, No. (42) told by Nang Wimala Thoumoung)

The TAM marker *ʒau³/yau⁴* is also found in future contexts, showing clearly that it marks aspect rather than tense, as can be seen in (271):

271) ၖ ၖၖ ၖ ၖ ၖ ၖၖ ၖ ၖၖၖ ။
    kau² ta¹ taai² ʒau³ kau² pɔ² ma¹ dai³ ŋin²
    I will die FINISHED I if NEG hear
'(By then) I will have died if I do not get to hear the answer.'

(Aiton Story, *The twelve questions*, No (21), told by Sa Cham Thoumoung)

This use of *ʒau³/yau⁴* in the future is very common in certain manuscripts, as in (272):

272) ၖၖ ၖ ၖၖ ၖၖၖ ၖ ၖ ၖ ၖၖ ၖ
    luk³ sii¹ wan² aaŋ¹ kaa¹ cam³ pii² maɯ¹ nɯɯ² maɯ¹
    from PRT* day Tuesday PRT year new month new

ၖ ၖၖ ၖ ၖ
nai³ wan² luŋ² ʒau³
get day one FINISHED
'From Tuesday, it will be the first day of the new year and the new month.'

(Aiton manuscript, *Calendar for the year 2001/2002*, No. (8), told by Sui Khong Thoumoung.)

In (272), *ʒau³* co-occurs with *nai³* 'GET' (see below 8.5.7.7.1), but the meaning is clearly in the future, because at the time at which the manuscript was written, the beginning of the new year was still in the future. The literal meaning is 'got the first day of the new year'.

8.5.7.7 Other TAM morphemes

8.5.7.7.1 dai³/nai³

The word *dai³/nai³* has a basic meaning of 'get' (see example (231) above). In all of the languages of the Tai family, this word has a number of meanings. In a study of Lao texts, Enfield (2000:123) found that it occurred as a main verb in 15% of occasions, as a preverb 45% of the time, and the remainder of the time postverbally.

No statistical analysis of the texts in this thesis has been done, but it does appear that *dai³/nai³* occurs much less frequently as a preverb than in the Lao texts studied by Enfield. It does occur postverbally as a completive verb with the meaning 'can', as in (273) and preverbally with the meaning of achievement, as exemplified in (274):

273) saŋ¹ ma¹ ca(a)p¹ ɛ¹ khɔn¹ ɛ¹ hɔŋ³ ma¹ dai³
if NEG join PRT *khon* PRT call NEG GET
'If it doesn't adhere, the *khon* can't be called.'
(Aiton text, *Explanation of Hong Khon ceremony*, No. (159), told by Nabin Shyam.)

274) pān² khiŋ² thüŋ⁶ mã² a¹ ni¹ cã⁴ tak¹ kaü³
duration of life body reach come changing will near

khãm² saŋ⁶ mau¹ nai³ saŋ¹ sɔn⁶ wai⁴ ॥
word any NEG GET advise teach to keep
'The duration of my life has nearly reached it's end (i.e. will change), and yet I have not taught any words of advice.'

(Tai Phake Text, *Grandfather teaches Grandchildren*, Intro No. (7), read by Yehom Buragohain)

Example (274) is from a manuscript and the use of *dai³/nai³* here probably does not reflect common spoken usage.

For Lao, Enfield (2000:181) categorised the preverbal *dai³/nai³* as (275).

275) *preverbal modal* 'result of prior event'
daj-V
V
this is the case because something else happened before this.

Enfield added that *dai³/nai³* should not be thought of as a 'past tense' marker, adding that

> "As many of the examples ... show *daj* is neither required by an expression's being set in the past, present or future, nor does its presence unequivocally denote any particular tense locus... It does, however, have *aspectual* properties related to a successful completion of actions/events ... and a common 'past tense' interpretation can arise from this."

The most frequent occurrence of *dai³/nai³* in preverbal position is in combination with *hau³/haü³* 'to give', as in (276):

276) an² khai³ an² cep¹ an² nɔt¹ an² nãu⁶ pi¹ haü³
CLF fever CLF pain CLF hot CLF cold don't GIVE

nai³ yaŋ² ॥
get have
'May you not have fever, pain, heat or cold.'
(Phake Blessing *Blessing*, No (15), intoned by Nang Pe)

This is a different preverbal usage from that found by Enfield for Lao and presented above in (275). Unfortunately there are not enough examples of preverbal *dai³/nai³* in

412

the texts that form the basis of this thesis to be able to discuss its usage in any more depth.

8.5.7.8 Combinations of TAM morphemes

Sometimes more than one of the TAM morphemes discussed in some detail above co-occur, as in example (277):

277) ယှိင် ᠊ᠣᠣᡳ ᠷᡙ ᠷᠫᠠᡅ᠊ ᠷᠣ ᠷᠶᠷᠣᠠ ᠣᠣᡳ ᠷᡙ ᠷᠣ
nüŋ¹ an² kau² khaü³ caü³ nɛ⁴ hɔ¹ phã¹ lãŋ⁶ nai⁴ nɛ⁴
as CLF 1Pl think PRT meeting CLF this PRT

ᠷᡙ᠊ ᠷᠷᠣᠠᡳ ᠷᡙ᠊ ᠷᠷᠣᠣᡳ ᠷᡙ ᠷᠷᡙᠠᡳ ᠷᡙ ᠷᠷᠣᠣ
mə⁵ nüŋ⁵ wan² nüŋ⁵ süŋ⁶ het¹ kā¹ ye⁴
time one day one if do GO PRT

ᠷᠣᠠᠣᡳ ᠷᡙ᠊ ᠷᠣ ᠷᠣᠠ ᠷᠣᠠᡅᡳ ᠷᡙ ᠷᠣᠠ ᠷᡙᠠᡳ ᠷᡙ
ā¹ mü² hau² nɛ⁴ ta¹ phen³ kā¹ ta¹ cuŋ¹ kā¹
work 1Pl PRT will tidy GO will complete GO
'I also think that our work will be tidy and complete if our meeting is done today.'

(Tai Phake *Speech*, No (40), given by Ai Chanta)

In this example, the adjectives *phen³* 'tidy' and *cuŋ¹* 'complete' are marked by both the future/irrealis marker *ta¹* and *kā¹*. As indicated above in Table 31, *kā¹* marks ingressive with adjectives, so the co-occurrence of these two markers in not inconsistent. In example (278), on the other hand, *ta¹* and *mā²* co-occur with a verb.

278) ᠷᠣᠠᡳ ᠷᡙ ᠷᠣᠠ ᠷᡙ ᠷᠣ ᠷᡙ
kɔ⁴ kau² ta¹ kā¹ khī³ mā²
friend 1Sg will go defecate COME
'I am going to relieve myself.'

(Phake Story, *The blind man and the man with scabies*, No. (15), told by Ee Nyan Khet)

It may be that (278) should be translated 'I am going to relieve myself (and then) come back.'

413

A more common co-occurrence is *kā¹* with *ȝau³*, as in (279), in which it appears that *kā¹* marks the fact that this sentence is in past time with relation to the time of speaking, and *ȝau³* marks the completion of the entire event.

279) ᠣᠠᡳ ᠷᡙ ᠷᠣ ᠷᠷᡙᠠ ᠷᠣ ᠷᠷᠣᡳ ᠷᡙ ᠷᠣᠠᡳ
pɔk³ maa² tii² hum² taŋ¹ luk³ saau¹ man²
return come to house with child female 3Sg

ᠷᠣ ᠷᡙ ᠷᠣ ᠷᡙ ᠷᠣᠠ ᠷᡙ ‖
hum² ȝuu¹ hum² sau² kaa¹ ȝau³
together stay together stay GO FINISHED
'And returned to the house and lived with her daughter (happily ever after).'

(Aiton Story, *Story of the old woman*, No. (23), told by Nabin Shyam).

8.5.7.9 Utterances unmarked by TAM morphemes:

The marking of Tai sentences with TAM morphemes is not obligatory, except in the case of the marking of future events with *ta¹/ti·/tak¹*. Habitual action is often unmarked, as in Example (280):

280) ᠷᠣ ᠷᠷᡙᠷᠣ ᠷᠣᠠᠣᡳ ᠷᠣ ᠷᠣᠠᡳ ᠷᡙ ᠷᠣᠠᡅᠣᡳ ᠷᠣᠠᡳ
nai⁴ a¹ khiŋ¹ mə⁵ hü⁴ ye⁴ pa¹ khā⁶ khau⁶ cüŋ¹ nai⁴ auk¹
this time summer PRT dolphin PLU now come out

ma cuM fUD ma xa hJwq mnq wJtq
mā² com² phü² com² phü² mā² khā⁶ hɔ² mün² wüt⁴
come sink float sink float come search boat 3Sg PRT
'So, in summer, the dolphins jump up and down in and out of the water, searching for its boat.'

(Phake Story, *The Dolphin, the Crow and the Mosquito*, No (47), told by Aithown Che Chakap)

Not all unmarked sentences are habitual. Sometimes TAM morphemes are not required because of the presence of a time expression, as in (281) and sometimes because of the context, as in (282):

281) မြစ်ဟဲ၁၁ဟ်   ဟ္ခောက်ဟ္ဂု   ရှ၁ဲ်   ၈၁၁ဲ်   ၈ဲ်   ပွံ ။
mɯ² ʒaam² nat³   cɔk¹ khaa¹ mɯm²   an²   tam² pauŋ¹
then   bag   3Sg   CLF   pound through
'At that time her bag was pounded through with holes.'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (69), told by Sa Cham Thoumoung)

282) ၀ဲ   ၈၁၁ဲ်   ဟ္ဂက်ဟ္ဂ၈်   ၈ုဲ်   ၁၁   ဟ၁ဲ်   ၈ဂု   ဟ္လှီ၈်
pai² han¹   muk³ sɯ²   khau¹ nai³/dai³ phaan² tuu²   lɯŋ²
go   see   hunter   3Pl   get   deer   CLF   one

တ္ဂ၈်   ၈၁၁ဲ်   ဟ္ဂှ်   ၈၁၁ဲ်
tɔp   kan²   sɔi²   kan²
cut   RECIP cut to pieces, to slice RECIP
'He saw that the hunters had got a deer, and were slicing it up.'

(Aiton Story, *Story of the Crow and the Fox*, No. (9), told by Ong Cham).

In example (282), it is clear that the getting of the deer preceded the seeing, because the hunters were already slicing it up. The context of the story tells us that the whole thing is in the historical past; pragmatic common sense tells us the hunters must have caught the deer before slicing it up, and the previous sentence in the story, in which the crow spots the hunter's fire, tells us that the deer was already caught when the crow saw it, and that the crow did not see the deer being caught.

8.5.7.10 Time expressions:

Table (32) lists the time expressions found in Aiton and Phake.

	Phake	Aiton	Gloss
မြစ်ဟ	mə⁵ nai⁴	mɯ² nai³	now, today
မြစ်ဟ	mə⁵ nan⁴	mɯ² nan³	then
မြစ်ဟဲ၁ဟဲ်	mə⁵ yām² nan⁴	mɯ² ʒaam² nan³	then
၁ဝ၈်ဟ	cüŋ¹ nai⁴	caŋ¹ nai³	now
၈ဟ	hā¹ nai⁴		then

တ၁ဲ်	wan² nai⁴	wan² nai³	today
၈ဂ	ma¹ ŋā²	ma¹ ŋaa²	yesterday
၈၈ဲ်	ma¹ sūn²	ma¹ sɯn²	the day before yesterday
၈ဟ္ဂက်	ma¹ phuk⁴	ma¹ phuk³, ma¹ phruk³	tomorrow
၈ရှ၈်	ma¹ hū²		the day after tomorrow
ဟ၈ဟ္ခ	kāŋ² naü⁶	kaan² naɯ¹	morning
ဟ၈၈၈်	kāŋ² wan²	kaan² wan²	midday
ဟ၈ဟ	kāŋ² kham²	kaan² kham²	evening

Table 32: Time expressions

These time expressions most frquently occur clause initially, as in example (283), although they can occur clause finally.

283) ၈ဟ၈်   ၈ဂ၈ဲ်   ၈၁၁ဲ် ၈ဂု ဟ္ဂ၈် ၈ု တ၈်   ၈၈   ။
hā¹ nai⁴   tauŋ⁴ mai³   an² ta¹ kin² ma¹ yaŋ⁶   yɛ⁴
at that time hungry   CLF will eat   NEG have.NEG PRT
'At that time they were hungry, but there was nothing to eat.'

(Phake Story, *The two brothers*, No. (4), told by Ee Nyan Khet)

8.5.8 Multi-Verb Sequences

In Tai languages verbs can be arranged serially, as in (284):

284) ၈ဂ၈်   ၈ဟ၈   ၈   ၈ဂ   ၈ဲ်   ၈ ။
pa¹ rɯ²   kɔ³   pai² khaa¹ um¹ dai³
day after tomorrow LINK go   search NEG   get
'On the second day they went and searched but did not get (the answer).

(Aiton Story, *The twelve questions*, No (38), told by Sa Cham Thoumoung)

The two verbs in (284) could both be followed by noun phrases, as in (285), and it is argued here that noun phrases something like *mɯŋ² kai³*'far country' and *kɛ³ pa¹ sa¹ naa¹*'answer to the question' are at least understood in (284).

285) ဒ မိၚ် ကဲ ၵ႑ ၵေ ပ႑ယၵႃ႑ ꩦၚ် ဒ ။

pai² muŋ² kai³ khaa¹ kɛ³ pa¹ sa¹ naa¹ uŋ¹ dai³
go country far search answer question NEG get
'They went to a far country to search for answers to the questions, but did not get them.'

There are many actual examples where serialised verbs are followed by objects, as in (286) where *khau³* 'rice' is the object of *khaa¹* 'search: and *luk³ saau¹ man²* 'her daughter' is the object of *liŋ³* 'feed'.

286) ၵ႑ ၵ်ႄ ဝ႒ သ္ၚ် ꩫ္ႄꩫ် ၵ္ꩪ် မꩫ် ꩫ် ။

khaa¹ khau³ sii¹ liŋ³ luk³ saau¹ man² nam¹
search rice PRT to feed child female 3Sg PRT
'... she searched for rice to feed her daughter
(Aiton Story, *Story of the old woman*, No. (5), told by Nabin Shyam).

Sometimes the multi verb strategy is used to express events which are sequential as in (287). The meaning here appears to be 'He saw a tiger and was shocked', but could equally be 'He saw a tiger causing him to be shocked'.

287) မꩫ် ꩁꩫ် သ္ꩥ် ꩮꩫꩥꩫ်

mün² han⁶ sə⁶ pɔ² lãp¹
3Sg see tiger to be startled, shocked
'seeing a tiger, he felt shocked'
(Banchob 1987:368)

It is not necessary for each of the verbs to be able to have an unstated object. This especially the case with motion verbs, which are often arranged as in (288), where a verb of manner of motion (*phut¹* 'rise up') is followed by one of a restricted set of verbs of path (*ɔk¹* 'come out') and then one of direction (*maa²* 'come'):

288) မိꩫ် ꩫꩫ် ꩫ႒ ꩫ် ꩮꩪꩫ် ꩫꩫ် မꩫ ။

muɪ² nan³ phii¹ nam³ phut¹ ɔk¹ maa²
when that ghost water rise up come out come
'Then the ghost of that river rose up out of the water.'
(Aiton Story, *Story of the old woman*, No. (16), told by Nabin Shyam).

This possibility of the verbs arranged serially having different objects allows three participant events to be expressed using more than one verb, as in example (289):

289) ꩫꩫ် ꩫ႒ ꩤꩫ် ꩪꩪ် ꩤ ꩫꩫ် မ႒ ꩫ႒ ꩫꩫ ꩫꩫ် ။

nok⁴ au² khun⁶ mɛn³ sɪ⁶ pauk⁴ mã² cɪ⁴ kɔ³ cãŋ⁴
bird take hair porcupine PRT return come show friend elephant
'The bird took the porcupine's hair and returned and showed it to his friend the elephant.'

(Phake Story, *The bird with the red bottom*, No. (19), told by Ngi Kheng Chakap)

In example (289) *nok⁴* 'bird' is the agent of the verb *au²* 'take' and indirectly of *cɪ⁴* 'show', and *khun⁶ mɛn³* 'porcupine's hair' is the patient or second participant of both verbs, although only indirectly of *cɪ⁴*. The other participant, *kɔ³ cãŋ⁴* 'his friend the elephant', is the goal of *cɪ⁴* 'show'. It is not necessary to mark it in a prepositional phrase because it is the only direct participant of that verb, the others being understood. It is clear that *kɔ³ cãŋ⁴* can only be the goal of *cɪ⁴*.

8.5.9 Completive Verbs:

A special type of multi verb construction is the completive construction where one verb completes the action of another. In contrast to the multi verb strategy in 8.5.8 above, a completive verb cannot have an object. Completive verbs have been discussed above in 8.2.3.4.

Example (290) exemplifies this. The main verb is *ʒuɪ²* 'shoot' and its action is completed (or not as in this case) by *thuk¹* 'touch'.

290) ꩮꩪꩫ် ꩫ႑ မ႑ ꩮꩫꩫ် ꩫ႑

ʒuɪ² kaa¹ ma¹ thuk¹ kaa¹
shoot GO NEG touch GO
'He shot but didn't hit it.'

(Aiton sentence spoken by Sa Cham Thoumoung)
Audio Link for this sentence (\wave files\aiton\extracts\2-1-2-2\a.mp3)

A rather poetic example of a completive verb is (291). Aimya Khang Gohain explained that the verb *ɲin²* means 'to listen with understanding', and is the completion of the action of *thɔm¹* 'to listen, obey.'

291) လူ   ဝင့်   ရံဝ့်   ၛ   ကိုဉ်   ၛိၚ် ॥
     sɯ⁶   thɔm¹   nam⁴ khām² wā⁵ kau²   ɲin²
     you all to listen, obey speaking   say   1Sg   hear
     'You should all listen to the words that I speak.'

(Tai Phake Text, *Grandfather teaches Grandchildren*, Intro No. (16), read by Yehom Buragohaiñ)

The noun *nam⁴ khām²* is literally 'water-word', which is translated by Tai informants as a single lexeme with the meaning 'speaking'. The structure of (291) is given in (292):

292) Subject Verb₁   Object   Verb₂
     sɯ⁶   thɔm¹   nam⁴ khām² wā⁵ kau²   ɲin²
     you   listen   speaking-say-I   hear

Example (291) could be translated literally as 'You (should) listen to my speaking and hear it (with understanding).'

Frequently there is more than one possible completive. Example (293) shows the word *hɔt³* as the completive of the main verb *khaɯ³ caɯ²* 'think':

293) ၝိုဝ်ဃၚာၚ်   ၯ္ဝင့်   ဃၚင် ကိုၚ် ၛၚၐ် ၛိဉ်   မၚ   �္ဒိလၚၚ ॥
     mɯ² ʒaam² nan³ khaɯ³ caɯ² saŋ¹ ɯŋ¹ hɔt³ ŋɯ³   maa² dii² ɔ¹
     then   think   what NEG clearly not clever come PRT
     'Then they couldn't think clearer, they were confused.'

(Aiton Story, *The twelve questions*, No (63), told by Sa Cham Thoumoung)

When this example was being translated, Nabin Shyam gave the meaning of *hɔt³* as 'clearly', but it seems to be the verb 'to arrive'. Nabin Shyam gave five ways of saying 'could not think', presented as (294):

294.1) ၯ္ဝင့်   မၚ   ၛၚၐ်
       khaɯ³ caɯ²   ma¹   hɔt³
       think   NEG   reach

294.2) ၯ္ဝင့်   မၚ   ဝ္ဒိၚ
       khaɯ³ caɯ²   ma¹   thuɰ¹
       think   NEG   reach

294.3) ၯ္ဝင့်   မၚ   ဝ္ဒိၐ်
       khaɯ³ caɯ²   ma¹   thuɰ¹
       think   NEG   meet/get

294.4) ၯ္ဝင့်   မၚ   ၚဝ
       khaɯ³ caɯ²   ma¹   pɛ³
       think   NEG   can

294.5) ၯ္ဝင့်   မၚ   ၝ
       khaɯ³ caɯ²   ma¹   dai³
       think   NEG   GET

Nabin Shyam did not indicate whether the five examples in (294) had different meanings, or would be likely to be used in different circumstances.

Sometimes, as in (295), an adjective can be followed by a completive verb.

295) တၚၚ်ကၚ ၚၚ ၚၐ်ၛ်ၐ် ၚၚ   ၙ   ၞၚ ၯၐ်   ကၚ ဝၐ် ဃ္ဇၐ် ॥
     taŋ¹ kaa¹ nam³ lik³ nam³ nɔi¹ hɛŋ³ cuɰm² kaa¹ cam³ ʒau³
     all   river small river small dry evaporate GO PRT FINISHED
     'All the small rivers have dried up and evaporated.'

(Aiton manuscript, *The creation of the world*, No. (8))

In example (295), the main predicate is the adjective *hɛŋ³* 'dry'. The verb *cuɰm²* is arguably a completive, and the whole predication is marked by the TAM₂ marker *kaa¹*, which shows that this event was in past time. An alternative analysis of (295) would suggest that this is a serial verb structure with the first verb in the series being an adjective.

8.6 Non-declarative sentences

8.6.1 Theoretical considerations

The three processes of interrogative, imperative and negative have been grouped together because of certain common features. The first two processes, interrogative and imperative, are often called 'mood', and are generally more marked than declarative sentences, often termed the indicative mood.

These three are treated together because of a common property: the fact that all three processes can be marked by changed tone in Phake (see for example below 8.6.2.3). It is also probably the case that the TAM markers are less frequently employed in interrogative, imperative and negative sentences.

8.6.2 Questions

Questions in the Tai languages of Assam are expressed by the use of a question particle (see 8.6.2.1) or a WH- type question word (see 8.6.2.2).

8.6.2.1 Polar questions

Polar questions use the particle $n\sigma^6$ in Phake or naa^1 in Aiton, as in example (296):

296) ဝဒ ယှ် ဧင်္ဒ ကော ဘ ဝဒ ယှ် ॥
taai² ʒau³ siŋ³ kɔ³ naa¹ taai² ʒau³
die FINISHED end LINK PRT.QN die FINISHED
'When we die, is that the finish?'
(Aiton text, *On Buddhism*, No. (12), told by Sa Cham Thoumoung.)

Polar questions occur less frequently in the texts than WH- questions. Therefore some of the data about them will be drawn from elicited sentences. In (297), a simple polar question and its affirmative and negative answers are exemplified:

297.1) ယှ် ဗင်္ဗ ဘ ကော ကော
maü² pen² hən² kā¹ nɔ⁶
2Sg be, become house GO PRT.QN
'Are you married?' (asked to a woman)

(Phake sentence elicited from Aimya Khang Gohain)

Affirmative answer:

297.2) ဘ ကော ဗင်္ဗ ဘ ယှ်)
ʒ² kau² pen² hən² yau⁴
Yes 1Sg be, become house FINISHED
'Yes, I am married.'
(Phake sentence elicited from Aimya Khang Gohain)

Negative answer:

297.3) ယှ် ဒ (ဝင်္ဒ)
mau¹ pai¹ loŋ²
NEG not yet go down
'I am not yet married' (i.e. I have not come down)
(Phake sentence elicited from Aimya Khang Gohain)

It is also possible to ask a question expecting the answer yes, as in (298):

298) ဘ ကော် ယှ် ဘ ကော
nai⁴ kat¹ mau² caü⁵ nɔ⁶
this cold NEG yes PRT.QN
'It is cold, isn't it?'
(Phake sentence elicited from Aimya Khang Gohain)

Such sentences also occur in the texts, as in example (299):

299) ယှ် ဗင်္ဗ ဘ ဝင်္ဗ ဘ ဘ ဘ ॥
ʒau³ muŋ² hau² suŋ¹ hau² uŋ¹ u²
FINISHED country 1Pl country 1PL NEG YES
'And then (he said) 'This is our country, isn't it?'
(Aiton text, *History of the Tai*, No. (90) told by Sa Cham Thoumoung)

Where a question is asking for a one of a set of alternatives, it is not necessary to use a question word, rather it is only necessary to put two propositions (here two noun phrases) in appositions. This is exemplified in (300):

300) ဋ္ဌိ ဩ ကျ ကၐ ၐ္ဍ ကၐ ဟ

maü[2] ta[1] kā[1] kāŋ[2] naü[6] kāŋ[2] kham[5]

2Sg will go middle morning middle evening

Will you go in the morning or in the evening?

(Phake sentence elicited from Aimya Khang Gohain)

A question with a final 'or not' would be expressed with two clauses, one positive and the other negative, as in (301):

301) ဋ္ဌိ ဩ ၑ ကၐၐ ၐ္ဍ ၐ္ဌိ ၑ ကၐၐ

maü[2] ta[1] mā[2] kin[2] khau[3] mau[1] mā[2] kin[2]

2Sg will come eat rice NEG come eat

'Are you coming to eat rice or not?'

(Phake sentence elicited from Aimya Khang Gohain)

8.6.2.2 WH–questions:

As discussed in 8.2.2.3 above, WH-question words occupy the position of noun phrases or prepositional phrases in the syntax, as in (302).

302) ဋ္ဌိ ဩ ကျ ၐ္ဋ္ဌ ॥

maü[2] ta[1] kā[1] thaü[6]

you will go where?

'Where are you going?'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (16), told by Ee Nyan Khet)

The WH–question words, together with the types of phrases that answer them, were listed in Table 5 in section 8.2.2.3.

8.6.2.3 Questions expressed by word order or tone alteration

In Phake, in some circumstances an alteration of word order or a reassignment of tone can change a statement into a question.

8.6.2.3.1 Word order

The late Aimya Khang Gohain explained that the order of constituents in a question depended on which item was being questioned. In (303.1) the question is whether the subject ('you') has eaten or not, whereas in (303.2) the question is whether it is rice that has been eaten, or some other food:

303.1) ဿ ကၐၐ ကျ ၐ္ဍ

sü[6] kin[2] kā khau[3]

2Pl eat GO rice

'Have you eaten rice?'

303.2) ၐ္ဍ ဿ ကၐၐ ကျ

khau[3] sü[6] kin[2] kā

rice 2Pl eat GO

'Have you eaten rice?'

(Phake sentences, elicited from Aimya Khang Gohain)

Audio Link for these two sentences, with an explanation by Aimya Khang Gohain (\wave files\phake\sounds\aimya.mp3)

Aimya Khang added that (303.1) would become a statement if ကျ *kā[1]* 'GO' were utterance final, as in (304):

304) ဿ ကၐၐ ၐ္ဍ ကျ

sü[6] kin[2] khau[3] kā[1]

2Pl eat rice GO

'You have eaten rice.'

In (303) tone marking has been omitted from the words *kā* 'GO'. In citation form the word is certainly *kā[1]*, but, as discussed above in 6.2.4.5.1, when *kā* is pronounced with the sixth or rising tone in a sentence like (303.2), the sentence becomes of question. Unfortunately Aimya Khang Gohain died before I was able to question him again about these examples; on listening to them again it does appear that both tokens of *kā* are pronounce with a rising tone, and should therefore probably be written as *kā[6]* with the gloss 'GO.QN'.

Possibly, therefore, constituent order variation alone will not lead to either (303.1) or (303.2) being interpreted as questions. Of course sentences such as these with 2nd person subjects are much more likely to be questions. It would be a quite marked situation if either (303.1) or (303.2) were declarative sentences.

8.6.2.3.2 Tonal alternation

As discussed above in 6.2.4.5.1, tonal alteration may be used to express questions. Banchob (1987) was the first to identify this phenomenon in Phake. Table (33) lists some of these tone alternations found in her dictionary:

ฟู๋ย	caü[6]	an interrogative form of caü[5]	(1987: 101)
ฟู๋ย	caü[5]	Yes	(1987: 101)
ฟู๋	mɔ[2]	to go back home	(1987: 307)
ฟู๋	mɔ[6]	to go back home (in questions, negative, and future tense)	(1987: 307)
ฟู๋ย	het[1] yau[4] kā[1]	it has been done	(1987: 347)
ฟู๋ย	het[1] yau[4] kā[6]	has it been done?	(1987: 347)
ฟู๋ย	ik[1] mā[2]	please give more.	(1987: 433)
ฟู๋ย	ta[1] au[2] ik[6]	will you want some more?	(1987: 433)
ฟู๋ย	mün[2] mā[5] mā[6] kə[4] nī[6]	If he does not come, it is also good	(1987: 5)

Table 33: Tone alternations for negatives and questions, after Banchob (1987)

It appears that this process is quite widespread among Phake speakers, and that certain very common verbs can change from the first or second tone to the sixth tone if the utterance is a question or a negative.

Diller (1992:19) also discussed this phenomenon[72], describing it as "a morphophonemic tone sandhi rule", indicating that it operates on words with the 2nd tone after "the preclitic negative marker [m-]" but not after *mau[1]*.

There is also a special tone for questioning in Phake, which was described above in example 6.2.4.5.1. The questioning tone is not found only on verbs, but can be found on the final lexical item of an utterance, regardless of its word class. In example (305), it is found on a noun. This example is a question that was asked when making a tour of a village and observing that an old house had been demolished and a new one was about to be erected.

305) ฟู๋ย ฟู๋ย ‖
      hən[2] nin[6(2)]
      house earth.QN
      '(Will it be) a house built on the ground?[73]'
      Audio Link for this sentence example (\wave files\phake\extracts\Ai Chanta \a.mp3)

The tape recorder was not running when this sentence was uttered. The same evening Ai Chanta repeated it for the tape, but actually said *ta[1] het[1] hən[2] nin[63]* (WILL-do-house-earth.QN).

There are restrictions on the use of the questioning tone. It is not used if the question particle *nɔ[6]* is used, as in the following examples:

306) ฟู๋ย   ฟู๋   ฟู๋ย ฟู๋ ‖
      phak[1]   nai[3] kin[2] nī[6(2)]
      vegetables/dishes this eat   good.QN
      'Is this dish[74] good to eat?'
      Audio Link for this sentence example (\wave files\phake\extracts\1-10-2-2 \a.mp3)

---

[72] In discussing Diller's findings, I have used the tone numbering in Banchob (1987), to avoid confusion.
[73] Traditional Tai houses are built on stilts and called *hən[2] hāŋ[3]* in Phake. Many Tai people are now demolishing their traditional style houses and building brick houses at ground level. These are called *hən[2] nin[2]*.
[74] The Tai word *phak[1]* literally means "vegetable", but has come to mean the dishes that are eaten with rice. It can include meat dishes.

307) ဧꩡꩯ ꩲ꩝ ꩯꩪ ॥

    kin² nī² nɔ⁶

    eat good PRT.QN

    'Is it good to eat?'

    Audio Link for this sentence example ₍ vave files\phake\extracts\1-10-2-2 \b.mp3)

8.6.3 Negation

Table (34) lists the negative morphemes from Banchob (1987), together with some examples of their use:

Word			Example		
ꩡꩯ	tap⁴	need not	꩝꩜ꩡꩯꩪ	maü² mã⁵ tap⁴ kã¹	you need not go
ꩩ	pai¹	not yet	ꩦꩩꩫꩯ	mã⁵ pai¹ het¹	not yet done
ꩩ	mã⁵	not	ꩡꩦꩫꩯꩬꩯ	kau² mã⁵ hɔn¹ han⁶	I have never seen
ꩬꩯ	hɔn¹	not	ꩬꩯꩪ	hɔn¹ kã¹	won't go
ꩮꩦꩨ	üŋ²	no	ꩶꩯꩮꩦꩨꩼꩨ	ŋün² üŋ² yaŋ²	no money

Table 34: Negative morphemes after Banchob (1987)

Negative particles in the Tai languages appear to fall into two categories. The simple negators recorded in this thesis are listed in Table (35):

Tai	Phake pronunciation	Aiton pronunciation	gloss	notes
ꩥꩨ	mau¹	mau¹	'NEG'	this is the form used in writing
ꩩ, ꩩꩬ	ma¹	ma¹	'NEG'	infrequent in Aiton
ꩮꩦꩨ, ꩮꩦꩨ	üŋ¹, üm¹, m¹	uɯŋ¹, uɯm¹	'NEG'	

Table 35: Simple negators in Tai

The negators listed in Table (35) occur preverbally, following TAM₁ markers such as the future/irrealis marker *ta¹/ti¹/tak¹* (see example (230) in section 8.5.8.4.1). These are exemplified in (308):

308) ꩡꩬ ꩪ ꩯꩪ ꩮꩦꩨ ꩨ ॥

    lɛ¹ khaa¹ kɔ³ uɯŋ¹ dai³

    roam search LINK NEG get

    'She went to look for it but could not get it.'

    (Aiton Story, *The story of the forest spirit and the first daughter*, No (7), told by Sa Cham Thoumoung)

In Table (35), the morpheme ꩥꩨ is characterised as 'the form used in writing'. This is certainly the case for old manuscripts, but most present day Aiton speakers would write the form ꩮꩦꩨ *uɯŋ¹* as in (302). Diller (1992:19) characterised this as a preclitic negative marker and wrote it as [*m-*].

In Phake, there are two forms which commonly occur, *ma¹*, as in (309) and *üŋ¹/üm¹*, which is exemplified in (312.3) below.

309) ꩯꩪꩨ ꩩꩨ ꩮꩯ ꩪ ꩡꩯ ꩩ ꩡꩬ ꩡꩬ ॥

    hã¹ nai⁴ tauŋ⁴ mai³ an² ta¹ kin² ma¹ yaŋ⁶ yɛ⁴

    at that time hungry CLF will eat NEG have.NEG PRT

    'At that time they were hungry, but there was nothing to eat.'

    (Phake Story, *The two brothers*, No. (4), told by Ee Nyan Khet)

The late Aimya Khang Gohain strongly urged that the negator always be written ꩥꩨ *mau¹*, to avoid confusion with the word ꩩ *mã²*'come'. This is discussed below in 9.1.

In example (309), the main verb is 'have', which in citation form is *yaŋ²*. As already discussed above in 6.2.4.5.2, some Phake verbs undergo tone change after the negative particle *ma¹*. This is most common with verbs that have the 2nd tone in their citation forms, but can be observed with verbs that have other tones, such as in (310).

310) ꩮ ꩫꩥ ꧟ ꩬꩳ ॥
pū¹   naï⁴   ma¹   haü⁶
grandfather that NEG give.NEG
'(But), the old man didn't give (them) the food.'
(Phake Story, *The two brothers*, No. (8), told by Ee Nyan Khet)

The main verb of (310) is 'give' which *haü³* in citation. The third tone is creaky, but in (310), the creakiness in absent and the 6th or rising tone has been employed in its place. Interestingly, a few moments after uttering (310), Ee Nyan Khet repeated the sentence, and the negative tone was not apparent, as in (311):

311) ꩮ ꩫꩥ ꧟ ꩬꩳ ꧏꩳ ॥
pū¹   naï⁴   ma¹   haü³   wɔt⁴
grandfather this NEG give   PRT
'(and so) the old man did not give (them the food).
(Phake Story, *The two brothers*, No. (10), told by Ee Nyan Khet)

In the examples of negative tone discussed so far, it appears to be redundant, in that it does not occur without the negative particle *ma¹*, nor does it appear to carry any meaning that would be lost if it were not present. This is in contrast to the questioning tone, the absence of which in a sentence like example (305) above would change the sentence from an interrogative to a declarative sentence.

However Ngo Ong, a Phake from Ningkam Phake village, said that the negative tone is used only for the habitual negative as in (312.2) and not when the action is non-habitual, as in (312.3).

312.1) ꧟ ꧘ꩳ ꨯꨮ ॥
ma¹   kin⁶   yɛ⁴
NEG eat.NEG PRT
'As to *ma¹ kin⁶*.'

312.2) ꩩ ꩫꩳ ꧟ ꧟ ꧘ꩳ ꩫ ꧟ ꧘ꩳ ॥
kū⁵   wan² kau² ma¹   kin⁶   nɛ⁴   ma¹   kin⁶
every day   1Sg   NEG eat.NEG DEF NEG eat.NEG
'(If I) don't eat fish every day, I would say *ma¹ kin⁶*.'

312.3) ꩬꩳ ꧟ ꧟ꩳ ꧘ꩳ ॥
mɔ⁵ naï⁴ kau²   üm¹   kin²
today   1Sg   NEG   eat
'(If I didn't eat) today, I would say *üm¹ kin²*.'
(Phake sentences spoken by Ngo Ong).
Audio Link for these sentence examples (\wave files\phake\extracts\1-10-2-9 \ngoaung.mp3)

The main verb of both (312.2) and (312.3) is *kin²* 'eat'. In (312.3) it is realised with the tone from its citation form, but in (312.2), it is realised with the negative tone, as *kin⁶*.

Ee Nyan Khet offered a different explanation for why the tone sometimes does not change. In the case of (313), the main verb *yaŋ²* retains its citation form, and does not appear as negative *yaŋ⁶*. These words were spoken in anger, and therefore the pronunciation was, as Ee Nyan Khet put it *cɔn² nüŋ⁵ haiŋ²* 'a little strong.'

313) ꧟ꩳ ꩻꩳ ꩹ꩳ ꧟ ॥
üŋ¹   yaŋ²   luk⁴   maü²
NEG have   child   2Sg
'You shouldn't have our child!'
(Phake Story, *Story of the kum⁴ bird*, No. (55), told by Ee Nyan Khet)

In both (312.3) and (313), the verb is preceded by *üŋ¹*, rather than *ma¹*. It may be that the negative tone is usually associated only with the negator *ma¹*.

The simple negators listed in Table (35) above occur before auxiliary verbs as in (314):

314) ꧟ ꩬꩳ ꧘ꩳ ॥
ma¹   khaü³   kin²
NEG want   eat
'(I) don't want to eat.'
(Phake sentence)

In (314), the negative tone is blocked by the presence of a auxiliary verb. The pronunciation *ma¹ khaü³ kin⁶* was declared ungrammatical by Phake informants.

Example (315) gives the syntactic frame for the negators with other preverbal elements:

315)   (TAM₁)   NEG   (AUXILIARY)   VERB

Sometimes even in manuscripts negation is understood on the word ໝ໌ *yaŋ* 'have', being interpreted by the Phakes as having the negative tone *yaŋ⁶*, rather than *yaŋ²*. One example where the negative reading is preferred was (316):

316)
ອງ໌	ໝ	ໝ໌	ໝ	ໝ໌	ໝ	ໝ໌ ‖
co⁵	nai⁴	cāt⁴	nā³	yaŋ⁶	nī²	pen²
time	this	life	future	have.NEG	good	be

'The present and the next life will (not) be good.'

(Tai Phake Text, *Grandfather teaches Grandchildren*, Intro No. (28), read by Yehom Buragohain)

Aimya Khang indicated that as this clause is linked to the previous one, in which there is a negative, the verb in (316) should be read as negative. When reading this text, however, Yehom Buragohain read *yaŋ²*.

A second type of negative morpheme, such as *yaa¹/yā¹* 'don't', *pai¹/pi¹/pa¹* 'don't', 'not yet' or *tap⁴* 'need not', are prohibitive in meaning. These can appear by themselves, or in combination with the simple negators discussed above. Some of these have already been exemplified after Banchob (1987) in Table 34 above. In Aiton, *pa¹* is frequently found, as in (317):

317)
ໝ	ໝ	ໝ	ໝ໌	ໝ ‖
pa¹	naa¹	hauɯ³	caak³	khɔ¹
NEG	PRT	give	dig	hoe

'Don't make them dig.'

(Aiton text, *History of the Tai*, No. (147) told by Sa Cham Thoumoung)

Finally, there is one case of possible double negation in (318):

318)
ໝ໌	ໝ	ໝ	ໝ໌	ໝ	ໝ	ໝ	ໝ ‖
khan⁶	hɔt⁴	cū²	cūŋ² nai⁴	mau¹	yā¹	com²	khā⁶
up to this	arrive	time	now	NEG	don't	follow	search

'Up to this time we have not failed to follow our grandparents.'

(Tai Khamyang Manuscript, *The book of calling back the Khon*, No. (7), read by Sa Myat Chowlik)

It might be better to gloss *yā¹* as 'cease' in (318). Diller (pers. comm.) pointed out that in some early Thai inscriptions, the phrase *bɔB3 yaaB3* is found. The example given was *bɔB3 yaaB3 khaatD1* (NEG-don't-separate), used when King Lue Thai[75] was chasing an elephant and meaning 'without giving up'. This may suggest that *yaaB3* was originally a main verb meaning 'to leave, stop doing something'

8.6.4 Imperative

Commands in Tai Aiton and Tai Phake can be expressed in several ways. One is by the use of hortative or imperative particles, often found in manuscripts, as demonstrated in example (319):

319)
ໝ	ໝ	ໝ	ໝ	ໝ໌	ໝ	ໝ	ໝ	ໝ	ໝ ‖
suu¹	hɛm²	khau³	hɛm²	lau³	kin²	kaa¹	ta³	waa²	nai³
2Pl	prepare	rice	prepare	wine	eat	GO	PRT.HORT	say	this

'You prepare rice and wine for us to eat, so they spoke.'

(Aiton manuscript, *The treaty between the Aitons and the Turungs*, No. (88).)

The hortative particle *ta³* occurs in utterance final position. It is less frequently found in spoken texts. Example (320) is from the *Story of the old woman*, which Nabin Shyam wrote down and read, although it is composed in spoken style.

320)
ໝ໌	ໝ	ໝ	ໝ	ໝ໌	ໝ	ໝ	ໝ
mauɯ²	pai²	uu¹	kaa¹	tii²	hum²	pum²	ta³
you	go	stay	PAST	at	other	other people	PRT.HORT

"'Go and stay at the house(s) of other people.'"

(Aiton Story, *Story of the old woman*, No. (13), told by Nabin Shyam).

---

[75] A Thai king from Sukothai, reigned 1317-1354.

Another method of expressing a command is to use the TAM₂ markers. While investigating the morpheme *kaa¹/kā¹*, I recorded several examples where the marking of past time was clearly not its function. One of these, which was written down but not recorded, is presented in (321), where the verb is followed by *kaa*, here notated without any tone marking.

321) ꩠꩠꩠ ꩠ ꩠꩠ ꩠꩠ
  ɯ² khai² kaa wɯm³
  yes tell GO PRT
  'Yes, tell it!'
  (Aiton sentence, spoken by Sui Khong Thoumoung.)

This use of *kaa/kā* in an imperative sentence remained a mystery until Aithown Che Chakap was assisting with the translation of his story, *The widow*. When discussing example (322), he pointed out that there was an imperative variant of *kā¹*, which was pronounced *ka³*.

322) ꩠꩠ ꩠꩠ ꩠꩠꩠꩠ ꩠꩠꩠ ꩠꩠꩠ ꩠꩠ ꩠꩠꩠꩠ ꩠꩠꩠ ꩠꩠꩠ ॥
  au² ka³ nɯŋ¹ kau¹ kɔ⁴ nɯŋ⁵ me² nɯŋ¹ kau¹ kɔ⁴ nɯŋ⁵
  take GO.IMP again   CLF one wife again   CLF one
  'Take another one, take another wife!'
  (Phake Story, *The widow*, No (12), told by Aithown Che Chakap)

Aithown Che explained that each of the three TAM₂ markers (see 8.5.7.5), had an imperative form, spoken with a short vowel and on the third tone, as in (322). Each of these three imperative markers still maintains some of its original meaning. In Phake, *wai³* would be used in a sentence like *tɛm³ wai³* (write-KEEP.IMP) 'Write it!', which suggests not only that the thing be done but that its results would remain. If calling a child to eat, one might say *kin² ma³* (eat-COME.IMP), because the child would likely be in another location and would need to come towards the speaker to eat.

This usage of imperative particles is summarised in Table (36)

	TAM₂ as markers of imperative
ka³	General imperative marker
ma³	Marker of imperative - used when the scope of the process indicated by the verb is in some way towards the speaker
wai³	Marker of imperative, used when the process indicated by the verb is intended to have long term results, also used as imperative past marker of states

Table 36: Imperative markers in Phake and Aiton

It appears that these markers are also found in Aiton, as in (323), where *ai³* is a spoken form of *wai³*. When the recording of the *History of the Tai* was copied by Bidya Thoumoung, he wrote ꩠ *wai³*.

323) ꩠ ꩠꩠꩠꩠ ꩠꩠ ꩠꩠ ꩠꩠ ꩠꩠ ꩠꩠ ॥
  mau² sɔ¹ lai¹ ai³ kau² ti¹ pɔk³ mɯŋ²
  2Sg   rule   PRT 1Sg will return country
  'You keep ruling, I will return home.'
  (Aiton text, *History of the Tai*, No. (137) told by Sa Cham Thoumoung)

It is not clear which tone these markers have in Aiton. It appears in (324) that it might be the second tone, but the sense is still that of a command.

324) ꩠꩠ ꩠꩠꩠ ꩠꩠꩠꩠ ꩠꩠ ꩠꩠ ꩠꩠꩠꩠꩠ
  phii¹ thum¹ ɯi² cii³ ma² sim² phraaŋ¹
  spirit forest VOC show COME.IMP spade

  ꩠꩠ ꩠꩠ ꩠꩠꩠꩠꩠ ॥
  cii³ ma² sim² phraaŋ¹
  show COME.IMP spade
  'Hey, Forest Ghost, show me my spade.'
  (Aiton Story, *The story of the forest spirit and the first daughter*, No (19), told by Sa Cham Thoumoung)

8.7 Complex sentences

In this thesis the term *complex sentence* will be taken to refer to sentences which are
made up of more than one clause. Vichin Panupong (1970:4-8) identifies three types
of complex sentences for Standard Thai, which she calls *complex*, *compound* and
linked sentences. A complex sentence for Vichin Panupong is "a sentence one or more
of whose constituents is in itself a downgraded sentence", where a downgraded
sentence is one "whose status can be reduced to becoming simply a unit of structure
within a larger construction". She exemplifies this in (325), where the downgraded
sentence is in italics:

325)   *กุหลาบ*   *ที่*   *ยัง*   *ตูม*   อยู่   ไม่   ควร   เด็ด

      kulà:p   *thî*   *yaŋ*   *tu:m*   *yù:*   mây   khuən   dèt

      rose   *which*   *still*   *bud*   PROG   NEG   should   pick

      'The roses which are still in bud should not be picked.'

Such complex sentences in Vichin Panupong's terms are those involving a noun
phrase with a relative clause or some other sentential modifier. The relative clauses
have already been discussed at 8.3.2.5.

A compound sentence, on the other hand, is one which contains two or more simple
sentences, or a simple and complex sentence, as in (326), linked by a class of words
called linkers by Vichin Panupong. The linker is italicised:

326)   ท่าน   ทราบ   ไหม   ครับ   *ว่า*   เขา   จะ   กลับ   เมื่อไร

      tân   sâ:p   máy   khráp   *wâ:*   kháw   cà   klàp   mûərày

      RESP   know   QN   PRT   *LINK*   3Sg   WILL   return   when

      'Do you know when he'll be back?'

A linked sentence is one which is preceded by a sentence linker. Usually these are part
of a compound sentence as in (327), where the linked sentence is shown in italics.

327)   ดี   ที่   ฝน   ไม่   ตก

      di: *thî*   *fǒn*   *mây*   *tòk*

      good   *which*   *rain*   *NEG*   *fall*

      'It's a good job *it didn't rain.*'

In the Tai languages of Assam, several types of sentences are made up of more than
one clause. In some cases, two sentences are linked without any implication of one
being in any way dependent on the other. Such sentences are linked by words like
sam⁴(nüŋ³), as in example (328):

328)   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   ।

      yɛ³ hɔm⁶   yaŋ²   ti¹   ti⁵   kɛn²   məŋ²

      Yehom*   have   stay   at   heart   country

      [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   ।

      sam⁴   kau⁴   au²   khon²   cü⁵   nɔ¹ pin²

      repeat   again   take   mix   name   Nabin*

      '"There is Yehom who lives in the heart of that country, in addition there is a
man called Nabin."'

      (Tai Phake Song, *Poem in the khe² khyāŋ² style: How Stephen Morey came to
the Tai Phakes*,. Nos. (24) and (25), read by Ai Che Let Hailung)

More frequently when two sentences are combined, one is in some sense dependent
on the other. Sometimes this is achieved by simple apposition of two sentences, as in
(329), the literal meaning of which is: '(there was) an earthquake - (so) the earth
shook'.

329)   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]

      iŋ¹ san¹   laŋ⁶ nin²   yon⁵   pai²   yon⁵   mā²

      earthquake   the earth   rock   go   rock   come

      "during an earthquake shock, the earth shook."

      (Banchob 1987:354)

When speech is being reported, it is introduced by the word *wā⁵* 'say', as in (330):

330)   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]   [Tai script]

      mün²   wā⁵   süŋ⁶   phā⁴   phon⁶   yɛ⁴   kau²   mau¹   pauk⁴

      3Sg   say   if   weather   rain   PRT   1Sg   NEG   return

      [Tai script]   [Tai script]   [Tai script]   ॥

      ta¹   sɔn⁴   nā²

      will   take shelter   field

      'She said, "If it rains, I won't return, I'll take shelter in the fields."'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (47), told by Ee Nyan Khet)

Sometimes speech or thought is introduced by other words, such as *hɯ⁴* 'know' in (331). This example is complicated by the fact that Ee Nyan Khet repeats the link (*hɯ⁴ thɔm¹ kā¹*) before introducing the second clause of this complex structure.

331)

ne⁴	cau³	ma¹ hɔ¹	ne⁴	hɯ⁴	kā¹	hɯ⁴	thɔm¹	kā¹
DEF	RESP	Mahosatha	PRT	know	GO	know	listen	GO

ɔ¹	ū³	mün²	ne⁴	kā¹	thai⁶	nā²
EXCL	father	3Sg	DEF	go	plough	field

'Chaw Mahosatha knew that her father had gone to plough the fields.'
(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (33), told by Ee Nyan Khet)

The second clause in (331) is in the form of direct speech. This sentence was difficult to render into English adequately, because in English direct speech is not used after 'knows/knew that'.

In addition to structures like (331), there are complex sentences in which one event has occurred before another. These sentences can be linked with clause initial conjunctions such as *ka¹ laŋ¹*, as in (332):

332)

khuɯ² kɔn¹	ka¹ laŋ¹	sɯ¹ kaa² phaa³	khɔt¹	muɯŋ²	maa²
also	after	Sukapha	abandon	country	COME

sii¹	paak¹	pii²	kun²	muɯŋ²	maan²	khau¹
4	100	year	people	country	Burma	3Sg

him¹	au²	kaa¹	muɯŋ²	kɔŋ²
snatch	take	GO	country	Kong

'Four hundred years after Sukapha left his country, the Burmese took Muang Kong.'
(Aiton text, *History of the Aiton*, No. (73) told by Nang Wimala Thoumoung)

In addition, there are the complex sentences in which one part of the complex is dependent on the other. In (333), the dependent or subordinate sentence is introduced by *kɔp⁴ pɔ⁵* 'because':

333)

kɔp⁴ pɔ⁵	kin²	kā¹ nī¹	phān⁶
because	eat	opium	poor

"He becomes poor, because of smoking opium."
(Banchob 1987: 37)

A particular type of subordinate relationship is that of condition, in which the subordinate clause is introduced by *if*, which is most often expressed in the Tai languages of Assam by *saŋ¹/sün⁶*, or less frequently by *pɛk³/paik³*, which we see in (334)

334.1)

man²	aa¹ saŋ¹ nai³	pɛk³	puɯ²	sip¹	kɔ³	maa²
he	HESIT	if	other	10	person	come

au²	mit³	maa²	tii²	maa²	bup¹	kɔ³
bring	knife	come	to	come	beat	PRT

muɯ²	nan³	man²	maa¹ naa³	tik¹ thii¹	ai³
time	that	3Sg	pride	wrong thinking	PRT

334.2)

man²	uɯ¹	pɛ³	kɔm³	man²	pam²
3Sg	NEG	can	bow	3Sg	ready to fight

ꩡ ꧤ ꩮ ꩦ ။

haw³ naa² caa² kau²
GIVE come to 1Sg

'If ten people came, bringing knives, came to beat him, and if in that time he was proud and haughty, he could not bow down, he is ready to fight, saying 'Let them come to fight me!'"

(Aiton Story, *The twelve questions*, Nos (167)-(168), told by Sa Cham Thoumoung)

Example (334) is rather too long to be regarded as a single sentence and when preparing the analysis of *The Twelve Questions*, Nabin Shyam suggested these be written as two sentences. However it is regarded as a single sentence because (334.1) is laying down the two conditions, namely *If ten men came* and *If he was proud and haughty*, and the main sentence is found in (334.2).

Example (335) also shows subordinate structures, here marked by *lai³* 'so':

335)

hat¹ hat¹ kai² kai² lai³ phraa² pin² cau³ pɔ¹ nɔi khyɔt³
very angrys   so   Buddha be RESP born PRT preach

lai³ hat¹ lai³ kai²
so   short temper so   angry

'They were very angry, so Buddha was born (here), and came to preach because of this short-temperedness.'

(Aiton text, *Why Buddha was born in this world*, No. (9), told by Sa Cham Thoumoung.)

8.7.1 Causatives & Purposive constructions

The verbs *au²* 'take' and *hauɯ³/haɯ³* 'give' both have secondary functions as the formative in causative and purposive constructions. Diller (1992:24) reported that *au²* is "the main causation formative ... The direct object ("causee") may follow this group, in which construction it is often "case marked" by the preposition *haŋ²*."

However, as discussed above in 8.3.3, it appears that the use of *haŋ²* is not related to the presence of one of these causation forming verbs.

In example (336), the causative is necessary because a Buddha image cannot of itself walk; hence it has to be caused to be gone. The agent is unstated.

336)

haam¹ phraa² haam¹ au² pai²
carry Buddha carry TAKE go

'Carry the Buddha image (and cause it to go).'

(Aiton text, *Explanation of Balipathar Temple*, No. (13), told by Chaw En Lai).

In example (337), the "causee" is in an prepositional phrase, headed by *caa²*. In this construction the intransitive verb *lap¹* 'to be afraid' in combination with *au²* becomes almost lexicalised as the transitive verb *au² lap¹* 'to frighten'.

337)

pai² au² lap¹ caa² mɛ² pɔ¹ haa¹ rii¹
go   TAKE afraid to woman businessman

mɛ² pɔ¹ haa¹ rii¹ lap¹ sii¹ ။
woman businessman frightened PRT

'They frightened the woman businessman. She was afraid'

(Aiton Story, *Story of the monkey and the fox*, No. (4), told by Bidya Thoumoung).

Probably the most lexicalised combination with *au²* is *au² taai²* (TAKE-die) 'to kill'. In example (338), on the other hand, there is no lexicalisation.

338)

phii¹ thum¹ an² cii³ hauɯ³ maa² cɔk¹ khaa¹ an²
spirit forest CLF point GIVE come basket   CLF

�%နeတင်   မိုင်   ကုၚ်   ဇိ   ဒ   ဒ   ကၚ   ॥

sim² phraaŋ¹   mɯɯn²   au²   haai¹   daai²   pai²   kaa¹
spade   3Sg   TAKE   disappear lost   go   GO
'He pointed out the basket and caused her spade to disappear.'

(Aiton Story, *The story of the forest spirit and the first daughter*, No (11), told by Sa Cham Thoumoung)

In this thesis *au²* has been glossed as 'TAKE', using upper case letters in any case where it is felt that there is some grammaticalisation of the original meaning. The word *haɯ³/haü³* is also used for causatives, as in (339)

339)   ဇိ   ၚင်   ၚၚ်   ဝှိ

haɯ³   naaŋ²   nan³   thɔm¹
GIVE   lady   that   listen
'To get her to listen.'

(Aiton text, *Advice to women*, No. (2.5), told by Nabin Shyam.)

The late Aimya Khang (pers. comm. 23/6/1999) stated that *haü³ pɔ²* meant "cause to, in order to" and gave several examples of its use similar to that in (340). In (340), there is some sense of purpose, although *haɯ³/haü³* could be interpreted variously. In this thesis, mindful of the fact that this word retains much of its original full verbal sense, it is glossed as 'GIVE'.

340)   ကၠုဟ္ၚင် ဟၚလ္ဍ   တၚက် လၚ   ဇိ   လၚ   ထိၚ် မၚ် ဟၟဒ
        a¹ khiŋ¹ kha¹ laü¹   tak¹   lā²   haü³   pɔ²   thüŋ⁶ mǎn³ pā¹ wai⁶
        time   how much will require GIVE enough reach village Pawai
        "How long does it take to reach Pawoi?"

(Phake sentence elicited from Aithown Che Chakap)

Sometimes there is no sense of either purpose or causation in its use, but it does involve transfer, as in (341):

341)   ၚၚ ၚိ   ၚၚင် ၛီဟၟၚ ၚ   ၚၟ   ၚၚင် ၚၟ   ၚၚၟ   ॥
        hā¹   nai⁴ nāŋ²   piŋ¹ ñā¹ ne⁴   lau⁵   haü³ haŋ²   cau³   ma¹ hɔ¹
        time this lady   Pingya DEF   say   give to   RESP   Mahosatha
        'Then Miss Pingya said to Chaw Mahosatha.'

(Phake Story, *Chaw Mahosatha meets Nang Pingya*, No. (56), told by Ee Nyan Khet)

In example (341) it is unclear whether it is more appropriate to gloss *haü³* as 'give', implying that the meaning of transfer is still salient, or as 'GIVE', suggesting that grammaticalisation has considerably advanced.

The word *haɯ³/haü³* also has a meaning of 'let' or 'should' as in (342):

342.1) ၚၟၠ   ၚၟၚ် ၚိ   ၚ   ၚၚင် ၚိ   ၚၟၚ်
        aa¹ muu²   an⁵   nii⁵ ne⁴   caŋ¹ haɯ³ het¹
        work   CLF   good this then GIVE do
        'This good work (only) they should do.

342.2) ၚၟၠ   ၚၟၚ် ၚး   ၚိ   ၚ   ၚး   ၚိ   ၚၟၚ်
        aa¹ muu²   an⁵   ma¹   nii⁵ ne⁴   ma¹   haɯ³ het¹
        work   CLF   NEG   good this not   GIVE do

        ၚၚင် ၚၟၚ်ၚၚ် ၚၚင် ၚၟၚ် ၚ   ၚိ   ၚၟၚ်
        haŋ²   luk⁴ ɔn¹   an⁵   khɯɯn³ maa² nai⁴   khau⁶
        to   children   CLF   increase COME these* 3Pl
        'The growing children should not be allowed to do these wrong deeds.'

(Khamyang Story, *Advice to children*, Nos. (30) and (31), told by Sa Myat Chowlik.)

Example (343) forms a neat quartet. In (343.1), *haü³* appears in its full verbal meaning of 'give', but in (343.2) it can be read as either a full verb in a serial verb construction or as a causative, or even benefactive construction. Example (343.3) is more clearly an applicative construction of cause or purpose. These three are then followed immediately by a causative construction with *au²* in (343.4):

343.1) ၚ   ၚ   ၚၚင် ၛီၚၟၚ ၚိ   ၚ
        ne⁴   me⁵   nāŋ²   piŋ¹ ñā¹ haü³   ne⁴
        DEF   mother lady   Pingya   give   DEF

        ၚ   ၚၟၚ် ၚၟၚ် ၚၟ   ၚၟၚ် ၚၟ   ၚၟ   ॥
        ne⁴   sɔ³   caik¹ phā³   caik¹ au²   mā²
        DEF   cloth tear   cloth tear   take   come
        'Miss Pingya's mother gave some torn cloth (to him).'

343.2) ఞ ఞ ఞ ఞ ఞ ఞ ॥
  mã²   haü²   cau³   ma¹ hɔ¹   nɛ⁴   yap⁴
  come   give   RESP   Mahosatha   DEF   sew
  'She gave them to the Mahosatha to sew.'

343.3) ఞ ఞ ఞ ఞ ఞ ॥
  phã³   hãi⁴   haü³   yap⁴   yap⁴ yap⁴
  cloth   bad   give   sew   sew
  'The bad clothes she gave him to sew.'

343.4) ఞ ఞ ఞ ఞ ఞ ॥
  au²   sɔp¹   au²   nĭ²   se⁶
  take   beautiful   take   good   PRT
  'He made them beautiful .'
  (Phake Story, *Chaw Mahosatha meets Nang Pingya*, Nos. (76)-(79), told by Ee Nyan Khet)
  Audio Link for this sentence example.
  Audio Link for the story of *Chaw Mahosatha meets Nang Pingya*

No better examples could finish the syntax chapter with the clear message that the syntactic construction in which a word occurs plays a key part in its interpretation. Words such as *haü³* continue to maintain at least part of their original meaning, no matter how grammaticalised they appear to be. I venture to suggest that there will be no occurrences of *haü³* in which the meaning of 'give', or the concept of transfer, is completely bleached.

9. The Literature of the Tai

ఞ ఞ ఞ ఞ ఞ ఞ

The stories of the Tai are so many
– Am Saeu Kyo

The literature of the Tai of Assam includes both written and oral texts. Every village temple possesses a large collection of manuscripts, mostly Buddhist texts, although these can only be interpreted by a relatively small number of people. Furthermore, many individuals possess both religious and non religious manuscripts, the latter including histories, astrologies and manuscripts dealing with traditional medicine.

Manuscripts are regarded with great respect in the Tai community. Those who can read and interpret them also make new copies to replace those older copies which will, in time, deteriorate and be discarded. Great merit is accrued in the copying of an important Buddhist manuscript. The temple libraries often contain multiple copies of a single text, because several individuals have copied the manuscript. Manuscripts are presented to monks at festivals. Sometimes, manuscripts are presented to honoured guests, and the writer has been presented with several manuscripts in this way.

Despite the importance of the manuscript tradition, much of what we will term the literature of the Tai is oral, some being taught by oral tradition, some being innovated by a story teller or singer, and some being based on manuscripts.

In order to discuss the place of literature, especially written literature, in the Tai community, it will first be necessary to briefly examine literacy in the community.

9.1 Literacy among the Tai

Although these communities have possessed writing for many generations, and although there is almost universal literacy in Assamese, the language of wider

communication, many Tai people cannot read the traditional Tai scripts, called *to²*
lik⁴ tai² in Phake.

One of the reasons for this is that the script has traditionally been taught by Buddhist
monks. Tais in Northeast India are Theravada Buddhists and every village has a
temple. At the present time, however, there are very few monks, and many villages
have no resident monk at all. Teaching the script is therefore only sporadic, and in any
case usually only available to men, because in the Theravada tradition followed in
Northeast India, only men can fully ordain.

There are, however, a number of women who have learned the script thoroughly,
especially in Namphakey village.

A second reason for the low level of literacy in the native script is the script itself. As
pointed out in 7.5 above, the script does not mark tone, nor does it mark all of the
vowel contrasts, nor even all of the initial consonants in all varieties. The script has
only 18 consonant symbols, even though Aiton, for example, as noted above in 6.3.1,
clearly has at least two consonant phonemes which are not represented in the script,
namely /b/ (usually notated as ᴜ <m>), and /d/, (usually notated as ᴜ <n>).

The problem is greater with the representation of vowels and tones. There are three
contrastive tones in Aiton and six in Phake, but there is no tone marking at all in the
writing system. Furthermore, there is, for example, no contrastive marking for front
vowels in closed syllables, and in Phake there are three contrastive front vowels. This
means that in Phake a syllable written ᴄ𝟠𝞮 <kiŋ> may be pronounced as kiŋ, keŋ or
kɛŋ on any of six different tones, a possible 18 different pronunciations of the same
graph.

This has lead to considerable difficulty in reading the texts, and also to some
controversy. Consider example (1), from the Phake *Grandfather teaches*
Grandchildren.

1)

khãm³	nam⁴	haü³	nɯ²	khɔɴ⁶
cross	water	GIVE	look	snag, log in water

pãn⁵	pai²	ɔn²	cüŋ¹	noi³
others	go	follow	maywith, together	

'When crossing the river, take care of the logs in the water. Let others go in
front.'

(Tai Phake Text, *Grandfather teaches Grandchildren*, Proverb No. 12, read by
Yehom Buragohain)

In the Phake community there is dispute as to the correct interpretation of the word at
the end of the first line. There are two possible readings, given in (2), of which the
first seems to be preferred by most members of the community:

2.1)   khɔn⁶ 'a snag, a log of wood, a stump or branch of a tree embedded in the
       bottom of a river'
2.2)   khɔn³ 'a wooden block, cut from the branch of a tree'. (In Shan given as 'a
       stick').

The late Aimya Khang Gohain, who preferred the second reading, explained it as
follows: "ᴜᴏᴠ𝟝 means 'a stick' (*khɔn³*), not *khɔn⁶* 'a log in water' One does not know
from the log in water, how deep the water is. But one can very well ascertain from a
stick in the hand about the depth of the water."

It has even been said that for a person to be able to read a Tai text, they need to know
its meaning beforehand. As Diller (1979:19fn) observed "Underdifferentiation of
tonal and segmental contrasts lead Lik-Tai readers regularly to spend time "puzzling
out texts". He further points out that Weidert (1979:322) reports a case of a Khamti
man requiring two days to decode a letter from his father!

The late Aimya Khang Gohain (pers. comm 23/6/1999)[76], commented on the need to
be as clear as possible in writing:

---

[76] The glosses and transcriptions in parentheses are added by me.

"The word ma^5 ('NEG') is used while speaking colloquially, but not used in literary writing, where ⟨ (mau¹ 'NEG') is used instead of ⟨ (ma¹ 'NEG'). Moreover the fifth tone is a long tone, while ma^5 'not' when spoken before a verb is a short tone. It may be written as ⟨. The use of the long tone sign here would confuse the meaning, e.g. ⟨ (ma¹ kin⁶')'not eating', ⟨ (ma² kin²') 'coming to eat', ⟨ (ma¹ nɔn⁶') 'not sleeping', ⟨ (ma² nɔn²) 'coming to sleep'. However, if one uses ⟨ (mau¹ kin²') 'not eating', ⟨ (mau¹ nɔn²') 'not sleeping', then there is no confusion at all."

The actual numbers of Tai people who can read the script is not known, but it is probably quite a small proportion of the total number of Tai speakers. Boonyong Kettatte (1998:168), writing about Khamti, noted that .

"... young Khamtis can speak their language but can neither read nor write it. Only those over 50 years of age can speak, read and write the language. However, Tai Khamti words and vocabularies are limited, so they borrow quite a few words from Hindi, Bengali, Assamese and English."

From my research, it is clear that there are people below the age of 50 who can read Tai script quite fluently. In Namphakey village, for example, most of the older men whom I have met are fully literate in Tai, and there are several women who have also learned the Tai script. There are also at least a small number of people under 35 who can read Tai script.

In the Aiton villages, the proportion of people who can read Tai script seems to be lower than among the Phake. I have only met one female who is fully conversant with the script, Nang Am in Bargaon, who stated that there were only three people in Bargaon who could read Tai.

In Duburoni the number of people who can read the Tai script is somewhat greater, and includes several younger men.

In the Khamyang village of Pawaimukh, there are said to be only two men who are literate in Tai, the younger of whom, Chaw Cha Seng, was born in 1928 and the older, Sa Myat Chowlik, in 1920. In addition to the two elderly Khamyangs, the only person in the village who is literate in Tai is the resident monk, the venerable Etika Bhikkhu, who is a Tai Phake and fluent in both Tai script and Burmese.

In Pawaimukh, however, the situation might be about to change. Due to the decline of the language (discussed by Deben Chowlik in *The future of the Khamyang language*), the old men of the village have met and decided that the children of the village should be taught Tai (see above 2.4.2). Sa Myat Chowlik's younger brother Mihingta has thus commenced teaching the language and script. In order to do this he has himself begun learning to read Tai.

9.2 Types of Literature

At least in part because of the difficulty of reading the script, oral literature holds a very important place in the community. The literature of the Tai will be briefly discussed in terms of three parameters: mode, age and genre.

There are both written and oral texts, but many of the oral texts are themselves partly or fully informed by written texts, as shown in Table (1):

Mode	Example
Manuscripts	*Chaw Mahosatha* Manuscript
Oral retellings of manuscripts	*Why the Buddha was born in this World*
Literary Oral narrative	*A story of Justice (Pu Noi Ce)*
Spontaneous Oral narrative	*A story of Justice (Sam Thun)*

Table 1: Modes of Tai Literature

In Table (1), a distinction between types of oral text has been drawn. A number of the literate members of the Tai community are able to tell stories that are clearly drawn from manuscripts, although it is possible that these stories were also passed down to them by word of mouth. The Aiton text *Why the Buddha was born in this World* is one such text, which, although presented in spoken style, tells a story which ultimately goes back to the Pali Canon, translated through several languages

eventually into Tai. Whether a manuscript containing this information still exists is not known. Presumably it existed in the past.

Two types of oral narratives are identified in Table (1), arising from the two versions of the *Story of Justice* recorded for this thesis. One of these, by Pu Noi Ce, is told in poetic language, with more elaborate vocabulary. The other, told by Sam Thun is much more spontaneous and vernacular in style.

I regret that during the process of collecting these stories, it was not always my habit to ask the informants about the provenance of the text that they were presenting. Further research on the literary forms of the Tai is certainly necessary.

The second parameter by which literature may be categorised is the age of the text. Most of the surviving manuscripts are up to 70 years old. The tradition of hand copying of old manuscripts remains, leading to the fact that most of the important manuscripts are present in several copies in temple libraries. For example in the Namphakey temple there were five copies of the *Chaw Mahosatha* manuscript when the village decided that I could take one copy to Australia for analysis. The age of the original texts from which these manuscripts were copied is indeterminate, but in some cases it may be up to several hundred years. This is suggested by the fact that many of the words which they contain are no longer in common use. Example (3) contains the word ဒီဘုဒီဘ *theu leu*, which was not known to all my Tai informants.

3)   ၆တ   ၈၃   က၃   ယက   ၮ
       paik⁴   tī⁵   kai²   yăn⁶   cū¹
       if   place far   far   long far place

       ဒီဘုဒီဘ   ၈၃   ယက   ကက ။
       theu² leu²   lāi⁶   săp¹   kɔ⁴
       very far   many   sides LINK
       "If it is in a very far away place, in any direction,"

(Tai Khamyang Manuscript, *The book of calling back the Khon*, No. 30, read by Sa Myat Chowlik)
Link to a photograph of the manuscript (\photos\khamyang\books\Lik Hong Khon\Hong Khon p2.jpg). This line is found on the 4th line, to the right.

In the process of translating the *The book of calling back the Khon*, leading Tai scholars from the Khamyang, Aiton and Phake communities were consulted. The first of these was Sa Myat Chowlik of the Khamyang, who gave the meaning of ဒီဘုဒီဘ *theu leu* as 'a place where men cannot go, where spirits can go but men cannot,' or as 'a place so far that you cannot see it'.

However when the translation was checked with Nabin Shyam Phalung, he did not know the meaning of ဒီဘုဒီဘ *theu leu*. It was only when this was rechecked with several Phake scholars, particularly Ai Che Let Hailung, that the meaning and tones for this word could be assigned.

The composition of new literature continues, and four poems were written in January 2000 to be read at the meeting in honour of the writer to express thanks for the printing of the *Phake Primer* (Morey 1999c). One example of this is the poem by Am Saeu Kyo, performed by her and several of her female friends on 29th January 2000. Example (4) is an extract from that poem.

4.1)   mɔŋ²   loŋ⁶   hau²   sau²   khɔŋ⁵   an²   kum¹
        country big   1Pl   to stay things   CLF   to gather
        'Those of us who live in this big country, gather your things.'

4.2)   ho⁶   caü²   com²   hom⁵ hom⁵   mā²   tum²
        head   heart   glad   gather   come   to gather
        'Our hearts are glad as we gather together.'

4.3)   tüŋ¹ loŋ⁶   caü²   küm²   hāŋ³   tā¹ nam¹
        all   heart   strong prepare PRT

        phuŋ⁶   kon²   lāi⁶   sam⁴   hom⁵   pauŋ²   kün²
        group   person many   repeat   gather   join   RECIP
        'All with strong hearts prepare for this festival, all the people gathered together.'

4.4) ပ္ဎ ၁ သ္ဎ ၣၜ ၮၢၣ် ၒၢၣ
phaü⁶ pai¹ caü⁶ lüm² auk¹ ta⁴ na²
who don't no forget come out PRT
'Don't forget, come out!'

4.5) ၣၜ ၮၢၣ် ၶၩၢၣ ၔၣ် ၛၖ ၣ္ဎ
khəŋ⁵ an² khe² khya² man³ hau² sau²
things CLF beautiful village 1Pl stay
'Prepare this beautiful village, where we live.'

(Tai Phake Song, *Song in honour of Stephen Morey* , Nos. 1-5, written and sung by Am Saeu Khyo.)

The third parameter is that of genre. Some of the genres are listed below in Table (2):

Genre	Example
Buddhist Manuscripts	*Mangala Sutta*
Stories from the life of the Buddha	*Chaw Mahosatha*
Moral tales	*Story of the Two Brothers*
Prayers	*Prayer of Blessings*
Songs	*Rice stamping song*
Histories	*The treaty between the Aiton and the Turung*
Ritual	*The book of calling back the Khon*
Astrologies	*Book of avoiding dangers*
Medical	*Tai Medical Manual*
Proverbs	*Grandfather teaches Grandchildren*

Table 2: Genres of Tai Literature

9.2.1 Cataloguing the Tai manuscripts

The Buddhist temple at Namphakey village keeps a very comprehensive catalogue of each of the books in the temple library, and the number of copies of each book that are held there. Most of the temple's collection are Tai translations or interpretations of Buddhist holy books, from the Pali canon. It also includes some Tai translations of Indian classics, such as the *Lama Mang*, which is a translation of the Indian Epic *Ramayana*.

Another very important library of Tai manuscripts is in the Department of Historical and Antiquarian Studies, Guwahati. Their collection includes manuscripts from all the

Tai communities of Assam, including the Ahom. A comprehensive and annotated catalogue is being prepared by Yehom Buragohain and Prof. J.N. Phukan.

Several Assamese publications have listed manuscripts. The most important of these is Gogoi (1993). Puspa Gogoi is a keen Ahom revivalist, and has attempted to not only list but also catalogue the manuscripts. Table (3) is an extract from his catalogue:

No. 6
Deo Buranji
Owner – Maniram Mahan Baruah, Maniki, Sibsagar
Subject – Book of creation
Date – uncertain, felt to be very old
Description – Sachipat size - 10" x 3"; 33 leaves. Each leaf contains five or six lines, language and script in Tai Ahom. The letters on cover page are damaged, covered by fumes of ti phai. e.g. fire place
A line – Kan phuk rang mung jen pin
mɛ̀ wɳ̃ ʌɛ̀ ʋ̃Lɛ̀ wʔɛ̀ ʋʔɛ̀
Content – It describes about the creation of the universe
Notes – The old manuscripts are kept in a box or bamboo tray hung over the fire place. It is one of the traditional method of preservation. The fumes of fire place keep the manuscripts well, drive away the mites and protects from fungus attack.

Table 3: Extract from Gogoi (1993)

Gogoi (1993) includes not only Ahom manuscripts, as in Table (3), but also Khamti manuscripts from Narayanpur on the north bank of the Brahmaputra.

Nomal Gogoi (1994:102) also lists the manuscripts held in the library of the temple at Namphakey.

9.2.2 The relationship between Ahom Texts and the texts of the other Tai groups

Terwiel (n.d.) made a careful study of an Ahom astrological text, *The Rotating Naga* and its Phake counterpart, and, while noting that the two texts are very similar and clearly have come from the same source, concluded that:
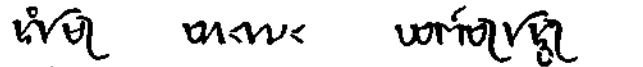
"It has been demonstrated that, although Ahom and Phakey share certain themes in their written traditions, there is no sign of these having been copied from each other during the last two hundred years. The differences in language,

script, content and style have been found consistent with the idea that the Ahom and Phakey texts have not been in contact for a very long period."

Another manuscript found in several Tai communities is the *Creation of the World*, for which Aiton, Phake and Ahom versions are known to exist. It must be assumed that they are all originally from the same source. Th Aiton *Creation of the world* has been partly translated.

On the other hand, Ahom manuscripts are often quite different from the Aiton/Khamyang/Phake texts with similar names. It seems, for example, that the Ahom *Rik Khwan* (calling the *Khon*) is not the same as the Khamyang *Book of calling the Khon*. Further investigation of these two texts is needed to establish what they have in common.

There is in addition a very interesting series of Ahom manuscripts, called Puthi, which are actually Buddhist texts. Their provenance and place in the Ahom literature is unclear, because it is generally believed that the Ahom were never Buddhists, unlike all the other Southwestern Tai peoples, and that as a consequence the literature they left behind is entirely pre-Buddhist. Nevertheless there are a number of such manuscripts, one of which has been analysed. That this is a Buddhist manuscript is clear from the opening phrase, as can be seen in (5)

5)

na[1] mɔ[1]	ta[1] sa[1]	phak[1] ka[1] wa[1] tɔ[1]
honour	to him	to one who is illustrious

a[1] ra[1] haa[1] tɔ[1]	sa[1] (maa[1]) sam[1] puk[1] (ta[1]) saa[1] thii[1]
to one who is worthy	to a true Buddha

'Honour to the one who is illustrious, worthy, a true Buddha.'
(Ahom Manuscript, *Advice to Women* >, No. (1)).

This phrase, the *namo tassa*, is a fixed form at the beginning of any Buddhist text in Pali as well as being written at the beginning of all manuscripts of the Tai Aitons and

Tai Phakes. Unfortunately the *Advice to Women* is an incomplete manuscript, but from what remains, it is clearly a Buddhist text.

9.2.3 The difficulties of reading Ahom Texts

When I first went to Assam in 1996, it was my wish to work on the Tai Ahom language, because of an interest in linguistic and cultural revival. It soon became apparent that working on Ahom manuscripts was not an easy task, and a great deal of study would be necessary before it would be possible to do it.

The problems with the Tai script, referred to above in 9.1 together with the fact that Ahom is no longer spoken as a mother tongue, has rendered the interpretation of Ahom manuscripts particularly challenging.

Terwiel (1988:294) advanced three reasons why the readings of Ahom manuscripts was so difficult. Firstly, most of the manuscripts were "copied by people who had, at best, only an inadequate knowledge of the script and language." If we assume that the scribes were already Assamese speakers, we might expect confusion between the letters w , historically /s/ and ᴎ , historically /c/ but now pronounced as /s/ by Assamese speakers. Such confusions would certainly lead to mistakes.

The second reason advanced by Terwiel is that the script itself was inadequate for rendering the living language, and therefore, "a reader who is unfamiliar with the text must read and re-read until he gets an inkling, from particular syllable combinations that cannot be misunderstood, what the general topic is about."

Thirdly, Terwiel strongly believes that only those who are well acquainted with "the principles of Tai grammar and who has access to a wide vocabulary can hope to create a meaningful translation of an Ahom text." It is hoped that this thesis might assist in this process.

9.2.4 The interpretation of Tai Aiton, Khamyang and Phake texts

In discussion with the Tai Phake scholars, it emerged that, just as Terwiel observed above, it is necessary to repeatedly read a manuscript in order to be able to find out its meaning. It is not necessary to be taught the meaning of the manuscript by another, although it seems that it is necessary at times to seek the advice of knowledgeable older persons as to the meaning of particular words and phrases.

The late Aimya Khang Gohain often invoked the authority of his late father when explaining to me why a particular phrase, such as that in (1) above, had the meaning that he ascribed to it.

For a non-Tai speaker to be able to read these manuscripts is clearly the study of many years. After nearly six years of study of the Tai languages, I certainly make no claim to be able to read their manuscripts.

9.3 The Importance of Literature

The various communities of Tai Aiton, Tai Khamyang and Tai Phake are very keen to have their languages studied and their traditions recorded. They are aware that as small communities of approximately two to three thousand, their languages are under threat from the Assamese language which is spoken by perhaps 20 million. They have already seen the Turung and Nora lose their Tai language completely.

As discussed above in 4.6, the recording and presentation of traditional literature must be accompanied by a respect for that literature as literature; and a respect for the fact that it is important not just to the community but to the whole of human culture that these literatures be recorded for their own sake, not just for the linguistic information we can mine from them.

The reader is enjoined to read some of these texts in English translation; and therefore in many cases full translations of the whole text have been provided, in addition to the line by line glosses necessary for linguistic analysis.

9.4 Bringing the Tai languages into the Computer Age - Presenting the literature for the benefit of the community:

The production of the computer fonts (see 7.8), has meant that, for the first time, printed books in the Tai language have appeared. Five such books were published and printed at Triograph Press in Dibrugarh as part of this PhD research project. These are listed in Table (4)

Book Title		Reference
လိက်ယွၢ�480 (ပြိုၐ်ဝ�\3ကၢ3တွၢ5)	Book for teaching the Tai Language - Aiton Primer	Morey 1999a
လိက်ဟိၐ်ၡိၐ်ဝ\3ကၢ3တွၢ5	Tai Aiton History	Morey 1999b
လိက်ယွက်ဟ္ၐ်ဝ\3 (ပြိုၐ်ဝ\3ဟဏ္ဏ)	Book for teaching the Tai Language - Phake Primer	Morey 1999c
လိက်ၡ္ဂၢ်ဟွၢ5	Book of calling the *Khon*	Morey 2001a
လိက်ၡ္ၐ်ယွၢ်ဟ3၃	Grandfather teaches Grandchildren	Morey 2001b

Table 4: Tai books published as part of this PhD Project

In addition, in 2001, several Tai Phake elders produced *The good way of Teaching* (Ngi Kheng Chakap and Ai Che Let Hailung 2001).

The onset of printing will lead to some changes in the way that literature is viewed in the community; it is to be hoped that the manuscript copying tradition will not be devalued, but as the Tai move into the modern age, it is likely that fewer people will take the time to hand copy manuscripts in the traditional way.

Since, as noted in 9.1 above, many Tai people who are fluent in the language are not necessarily literate in the Tai scripts, those books already published have presented the texts both in Tai script, and in Tai language using Assamese script. The ideal would be a four way presentation, as in Table (5):

- The text in Tai language and script.
- The text in Tai language, using the Assamese script (see 7.9 above).
- A translation of the text into English.
- A translation of the text into Assamese.

Table 5: Ideal for the presentation of Tai texts and translations.

This methodology was followed in the production of the *Tai Aiton History* (Morey 1999b).

9.5 The texts presented in this thesis:

Over seven hours of text have been analysed and translated, all of which are presented as an electronic appendix to this thesis. In the following tables, the text number, length of the text, name of the informant and name of the text are given. All of these files will be found in the \texts folder of the electronic version of this thesis.

Phake texts are listed in 9.5.1, Aiton texts in 9.5.2, Khamyang texts in 9.5.3 and Turung texts in 9.5.4.

9.5.1 Phake:

Stories

1-1-1-1	2'00"	Ee Nyan Khet	The blind man and the man with scabies
1-1-1-2	1'35"	Ee Nyan Khet	The story of the two brothers
1-1-1-3	4'50"	Ee Nyan Khet	Chaw Mahosatha meets Nang Pingya
1-1-1-4	2'52"	Aithown Che	The Dolphin, the Crow and the Mosquito
1-1-1-5	3'14	Aithown Che	The widow
1-1-1-8	1'15"	Pu Noi Ce	A Story of Justice
1-1-2-1	1'09"	Yehom Buragohain	Crossing a bridge
1-1-2-2	2'40"	Ee Nyan Khet	Chaw Mahosatha and the raising of cattle
1-1-2-3	2'50"	Ee Nyan Khet	The birth and early life of Chaw Mahosatha
1-1-2-4	6'00"	Ee Nyan Khet	Story of the *kum⁴* bird
1-1-3-1	2'00"	Ngi Kheng	The bird with the red bottom
1-1-3-2	1'24"	Sam Thun	A Story of Justice
1-1-3-3	1'22"	Sam Thun	Story of the dogs diving up their children
1-1-3-4	1'22"	Sam Thun	Story of the new king
1-1-3-6	2'19"	Ee Nyan Khet	Story of Deception
1-1-3-7	4'42"	Ee Nyan Khet	Story of the foolish king

Histories

1-2-2-1	13'	Ee Nyan Khet	Story of her youth

Prayers

1-3-1-3	3'06"	Sam Thun	Prayer of Blessings
1-3-2-1	3'30"	Nang Pe	Blessing
1-3-2-2	2'23"	Ai Thaan	Prayer of Blessings

Buddhist Texts

1-4-3-1	17'12"	Ee Nyan Khet	Chaw Mahosatha Manuscript - Chaw Mahosatha meets Nang Pingya

Other Cultural Texts

1-6-1-1	3'40"	Ee Nyan Khet	Riddles
1-6-1-2	1'10"	Sam Thun	Riddle - what has three heads and ten legs
1-6-1-3	1'10"	Ee Nyan Khet	Lullaby - the little rooster
1-6-1-4	2'	Sam Thun	Lakni - the Tai Calendar
1-6-2-1	10'47"	Ee Nyan Khet	Book for explaining the meaning of words to children
1-6-2-4	11'04"	Pingyajyoti Bhikkhu	Tai Calendar for the Year 2002-2003
1-6-3-2	15'01"	Yehom Buragohain & others	Grandfather teaches Grandchildren

Songs

1-7-1-3	2'58"	Ai Che Let	Rice Pounding Song
1-7-1-9	4'49"	Ee Nyan Khet	Rice Pounding Song
1-7-1-10	5'00"	Ai Che Let, Ngi Kheng & Sam Thun	Song in honour of Stephen Morey
1-7-1-11	4'03"	Am Saeu Khyo	Words of gladness - in honour of Stephen Morey
1-7-2-1	16'47"	Ai Che Let	Poem in the *khe² khyāŋ²* style: How Stephen Morey came to the Tai Phakes.
1-7-2-2	2'26"	Ee Kya	Song of hoisting the Buddhist banner
1-7-2-3	2'52"	Ee Kya	Song of teaching the Tai Language
1-7-3-1	0'26"	Ee Nyan Khet	Our fragrant land
1-7-3-2	0'45"	Ee Nyan Khet	Importance of Tai traditions
1-7-3-3	0'17"	Ee Nyan Khet	Song about fruits
1-7-3-4	0'39"	Ee Nyan Khet	Fish, why are you blind?
1-7-3-5	0'42"	Ee Nyan Khet	Numerical rhyme
1-7-3-6	0'38"	Ee Nyan Khet	Australian Visitor
1-7-3-9	7'09"	Ee Nyan Khet	The Australian researcher - in honour of Stephen Morey

Informal texts

1-9-4-1		Peim Thi Gohain	Letter

Grammatical texts

1-10-2-3	3'18"	Ai Che Let	The Phake Tones
:-10-2-7	~1	Ai Che Let	Names of the vowels

Speeches

1-11-1-6	4'04"	Ai Chanta	Speech at Meeting, 29/1/2000

9.5.2 Aiton

Stories

2-1-1-3	14'50"	Sa Cham	The Twelve Questions
2-1-1-4	4'35"	Sa Cham	Story of the forest spirit and the first daughter
2-1-3-1	0'50"	Chaw En Lai	How the Aitons and the Khamtis quarreled
2-1-3-2	3'09"	Ong Cham	Story of the crow and the fox
2-1-3-3	2'59"	Nabin Shyam	Story of the old woman
2-1-3-4	3'24"	Bidya Thoumoung	Story of the monkey and the fox
2-1-3-5	0'30"	Bidya Thoumoung	Story of the boy who wouldn't go to school
2-1-5-5	2'21"	Mohendra Shyam	Story of the forest spirits and the opium pipes

Histories

2-2-1-1	9'40"	Nang Wimala	History of the Aiton
2-2-2-1	7'28"	Nabin Shyam	Words in Praise of Khau Khau Mau Lung
2-2-3-1	5'56"	Ong Thun	Story of Barhula
2-2-3-2	10'37"	Sa Cham	History of the Tai
2-2-4-1	12'26"	Nabin Shyam	Treaty between the Aitons and the Turungs
2-2-4-2	2'15"	Nabin Shyam	Treaty between the Aitons and the Turungs - coda
2-2-4-3	8'21"	Nabin Shyam	Book of history (from) the time of the ancestor Chaw Tai Lung up until Sukapha
2-2-4-4	8'52"	Nabin Shyam	Creation of the world

Prayers

2-3-1-1	1'39"	Nang Am	Prayer
2-3-1-3	0'25"	Ven. Si Vicikta	Prayer (partial translation)

Buddhist Texts

2-4-1-4	5'53"	Ai Mya	Satipatthana
2-4-2-2	18'25"	Nabin Shyam	Advice to Women (Buddhist MS in Ahom script)
2-4-3-1	1'13"	Sa Cham	Why Buddha was born in this world

2-4-3-2	3'35"	Sa Cham	On Buddhism

Ritual Texts

2-5-1-9	3'	Silava	Butterfly Dance
2-5-2-2	7'36"	Nabin Shyam	Explanation of Hong Khon ceremony.
2-5-2-4	20'53"	Sui Khong	Manuscript of the Tai Calendar for the Year 2001/2002 (Sakkarit)

Other Cultural Texts

2-6-1-1	0'49"	Nang Am	Lullaby
2-6-3-3	0'33"	Nabin Shyam	Numerical Rhyme
2-6-4-1		Pradip Thoumoung	Aiton Unlucky Days
2-6-4-2		Pradip Thoumoung	Omens for building a new house
2-6-4-3		Pradip Thoumoung	The Ten Perfections
2-6-5-1	1'35"	Nang Am	Various Cultural Items

Songs

2-7-1-4	2'00"	Cham Mya	Love Song
2-7-1-6	3'18"	Ruhila & others	Song of the village ...
2-7-1-9	2'00"	Pradip	Sangkyen Song
2-7-1-15	4'30"	Ong Cham	Love Song
2-7-3-1	~0'30"	Ai Hom	Dhonsiri River
2-7-3-2	1'08"	Nabin	Blowing Wind
2-7-3-4	0'43"	Bidya	Love Song - Sangkyen
2-7-3-5	1'12"	Bidya	Love Song - Our eyes will meet on the road
2-7-3-6	0'49"	Bidya	Love Song - Oh Golden Girl

Explanations

2-8-1-6	6'50"	Sui Khong	Explanation of Spirit House
2-8-1-8	2'00"	Chaw En Lai	Explanation of Balipathar Temple

Informal

2-9-2-1	7'11"	Nabin Shyam	Conversation with Nabin Shyam
2-9-2-5	~2'50"	Nabin Shyam	Conversation with Nabin Shyam
2-9-4-1		Nabin Shyam	Letter from Nabin Shyam
2-9-4-2		Nabin Shyam	Sample Letter
2-9-4-3		Nabin Shyam	Letter from Nabin Shyam
2-9-4-4		Ngi Khang In	Letter

9.5.3 Khamyang:

Stories

3-1-1-3	8'38"	Sa Myat Chowlik	Of children and kings
3-1-2-3	1'37"	Deben Chowlik	The Hunter's Parrot
3-1-3-2	3'40"	Sa Myat Chowlik	Advice to Children
3-1-3-3	5'41"	Sa Myat Chowlik	The Bird and the King

Histories

3-2-1-1	7'32"	Sa Myat Chowlik	The second world war and its aftermath

Prayers

3-3-1-2	3'29"	Chaw Cha Seng	Parinibbana

Buddhist Texts

3-4-1-3	8'00"	Sa Myat Chowlik	Mangala Sutta
3-4-2-1	30"	Sa Myat Chowlik	Some words about religion (partially translated)

Ritual

3-5-1-1	14'29"	Sa Myat Chowlik	Book of calling back the *khon*

Other cultural texts

3-6-2-1	3'37"	Deben Chowlik	The Future of the Khamyang language
3-6-2-5	2'00+	Sa Myat Chowlik	The book of avoiding dangers

Grammatical texts

3-10-2-6	7'12"	Deben Chowlik	Khamyang Sentences
3-10-2-7	1'00"	Deben Chowlik	Khamyang Sentences

9.5.4 Turung

Stories

7-1-2-1	1'58"	Pong Chap	The old couple without any food (in Turung)
7-1-2-2	1'37"	Pong Chap	The old couple without any food (in Tai)

10. Lexicography

Two lexicons are presented as part of this thesis, the *Tai Aiton Dictionary* (see 10.1) and *Tai Phake Dictionary* (see 10.2).

10.1 Tai Aiton Dictionary

10.1.1 The word list of Bidya Thoumoung

In November 1999, Bidya Thoumoung, Nang Wi Thoumoung and Pradip Thoumoung were interviewed with regard to the production of a Tai Aiton word list. The first task was to identify the number of tones in use in the Aiton villages. After some discussion, it was agreed that there were probably five tones, although it appears that only three of these are contrastive (see above 6.3.4.1.3).

Bidya Thoumoung then commenced the task of listing as many Aiton words as possible. He wrote down each orthographic syllable, and then each of the meanings in Assamese for any possible pronunciation of that syllable. After discussion with him, I then added the phonemic transcription and English glosses. In almost all cases he listed monosyllabic words first and only added polysyllabic words with some difficulty.

Table (1) gives the full list of meanings and pronunciations which Bidya Thoumoung identified for the graph ကင် (kaŋ):

Aiton	Assamese	Phonemic	Gloss
ကင်	থুঁতৰি.	[kaaŋ²]	chin
ကင်	বহল	[kaaŋ³ / kaaŋ⁵]	wide
ကင်	কাড় ধনু	[kaaŋ¹ / kaaŋ⁴]	a bow
ကင်	হৰ পছ	[kaaŋ²]	deer
ကင်	দিনিল লাগি ধৰা	[kaŋ³ / kaŋ⁵]	to be stuck in the throat
ကင်	গলি যোৱা	[kaŋ²]	to spread (over) water

ကၢင်	ফিচা	[kaaŋ³]	fishbone
ကၢင်းၠ	পথাৰ	[kaaŋ² naa²]	paddy field
ကၢင်ၵ	ভেকুলী	[kaaŋ³ khaa³]	a type of frog that lives inside the house
ကၢင်�721	বাতি	[kaaŋ² kham²]	night
ကၢင်ၠ	বাতি পুৰা	[kaaŋ² nauɩ¹]	morning
ကၢင်ဝၢ	দিন	[kaaŋ² wan²]	daytime
ကၢင်ၢ	পানীৰ মাজত	[kaaŋ² nam³]	in the middle of the water

Table 1: Words with the written form ကၢင် according to Bidya Thoumoung

Altogether 1924 words were collected, and recorded on tape. There are without doubt thousands of other words which have not been included. Examples of word types which are certainly under-represented include polysyllabic words, proper names, names of animals and plants, taboo words and specialised vocabulary such as religious or astrological terms.

10.1.2 Other sources of lexicographical information on Aiton

There are two main sources for additional lexicographical information on Aiton. The first is the *Aiton–English–Thai Dictionary* (Banchob ms). In the future, the entire contents of this should be entered into the computer data base (see below 10.3).

The second source is the literature of the Aiton. The study of Aiton literature, whether Oral or written, will gradually reveal a large proportion of the words which were not listed by Bidya Thoumoung. As each text is analysed, any word which does not appear in Bidya's list is noted with its Phake cognate (Banchob 1987) or its Shan cognate (Sao Tern Moeng 1995) if there is one.

Table (2) lists the words in the first ᶠ··· ⫶·· ᶠ *Treaty between the Aitons and the Turungs* which were not in Bidya Thoumoung's list.

Aiton script	possible phonemic transcription	Phake	English	Shan	Reference
ꩰ	khuɩ²	khə²	lineage	'creeper, race, lineage'	1
	tu¹ ruŋ² / ta¹ ruŋ²		Turung		2
	seŋ¹	saŋ⁶	make an appointment, a promise		2
	huŋ³ mut³		Hung Mut, name of a year		3
	lak¹ nii³		year		3
	saaŋ¹ puŋ¹ an² lak¹ nii³		Lakni Board	'a spelling book'	Note to No 3
	iŋ¹ san¹	iŋ¹ san¹	earthquake	'quake, as the earth, earthquake'	5

Table 2: Sample of list of words not found in Bidya Thoumoung's list

The full list of Aiton words not found in Bidya Thoumoung's list but occurring in the texts, is presented in Additional Aiton words.doc (in \lexicon\aiton).

In future, it is hoped to use the information in these texts to make a dictionary of Tai Aiton, which not only defines the meanings of words, but also exemplifies these words with reference to Aiton literature. Example (1) shows how a dictionary entry for the word iŋ¹ san¹ "earthquake", might appear:

1)

[iŋ¹ san¹] an earthquake

muɩ² nan³ iŋ¹ san¹ luŋ¹ kaa¹ pɔ² thun³ pet¹ wan² nam¹ naa¹ (There was an earthquake lasting up to 8 days.) [*Treaty 5.*]

In (1), the abbreviation [*Treaty 5*] refers to the source, *Treaty between the Aiton and the Turung*, and the sentence number.

10.1.3 Current state of the Tai Aiton Dictionary

The entire list elicited from Bidya Thoumoung has been entered into the data base (see 10.3 below). The flexibility of the data base allows for any number of fields to be printed in any order. The *Tai Aiton Dictionary* included in this thesis presents five of these fields (see \lexicon\aiton\aitdic.doc):

> Head word in Aiton script
>
> Head word in Phonemic script
>
> Transcription into Assamese Script
>
> Gloss in English
>
> Gloss in Assamese.

This is exemplified in Table (3):

ကောက်ခေါ် kaak[1] khau[3] কাক বাও type of rice হেতা
ကောက် kak[3] কাক to keep one portion ভাগ কৰা
ကင် kaŋ[2] কাঙ to spread (over) water গলি যোৱা
ကင် kaŋ[3] / kaŋ[5] কাঙ to be stuck in the throat দিনিল লাগি ধৰা

Table 3: Sample from *Tai Aiton Dictionary*

In addition, an English–Aiton word finder has been derived from the data base. This is exemplified in Table (4):

keep ကောက် ။ kak[3] কাক to keep one portion
ဒ ။ wai[3] ৱাই to put, to keep
ဂုၚ် ။ o[1] ত্মা to keep carefully
ဂုၚ် ။ um[2] উম to hold in the mouth

Table 4: Sample from English–Aiton word finder, in *Tai Aiton Dictionary*

One of the fields in the data base is headed <english_heads>. Every word which has 'to keep' entered in this field has been listed under the heading *keep*, together with its phonetic transcription and gloss.

In time it is hoped to add more information about word class, etymology and other lexicographical information.

A small section of the Aiton dictionary has been linked to sound files, with Bidya Thoumoung pronouncing each word in Assamese and then in Tai Aiton.

10.2 Tai Phake Dictionary

Since Banchob (1987) has already produced the very substantial *Phake–English–Thai Dictionary*, it has not been necessary to collect a word list to form the basis of a lexicon as was the case with the Aiton.

Banchob's excellent work suffers from two major disadvantages: its very large bulk makes copying and use difficult, and the lack of Tai script makes it very difficult for most Phakes to understand it.

Therefore the entire *Phake–English–Thai Dictionary* has been entered into a data base (see 10.3), with the orthographic form of each word as the head entry. Using the various fonts devised as part of this project, it has been possible to produce a version of Banchob's dictionary with the head word, phonemic realisation, and English translation as given by Banchob.

Banchob (1987:8) limited her lexicographical work to the collection of "words for daily use". She hoped to collect literary words later, but was unable to do so.

As with Aiton, words identified in Phake texts which are not found in Banchob's dictionary have been listed, and will be added to the data base over time. As with the Aiton dictionary, it is hoped that eventually a comprehensive dictionary with examples from the literature can be published. A full list of Phake words not found in Banchob (1987), but which have been found in the texts, have been gathered as Additional Phake words.doc. (in \lexicon\phake). This list will form the basis of additions to the Phake Dictionary.

10.2.1 Current State of the Tai Phake Dictionary

The present *Tai Phake dictionary* (see \lexicon\phake\phakdic.doc) contains all the head words listed in Banchob (1987) and a considerable number of compound words from her list. Four fields have been presented:

Head word in Phake script

A head and sense number

Head word in Phonemic script

Gloss in English

As with the Aiton Dictionary, an English–Phake word finder has also been derived from the data base.

10.3 The computer data base:

10.3.1 Design of the data base:

With the assistance and advice of Doug Cooper, Southeast Asian Software Research Centre, Bangkok, I have set up a data base in a modified XML format. This format is maximally flexible and can easily be converted to a format which is able to be entered into the SIL shoebox programme or similar lexicographical software. It is somewhat more flexible than any lexicographical software currently available, in particular allowing for the free use of as many fonts as required within any particular field.

A single data base is being used for both the Aiton and the Phake dictionary.

The data base has been designed to include:

i)   a head word in the Tai script

ii)   the native orthography for each of the varieties, with an allowance for orthographic alternatives

iii)   the word in a phonemic representation in each of the target varieties

iv)   definitions in English, Thai and Assamese

v)   example phrases or sentences to illustrate the word.

vi)   a link with the computer data-base of the Thai English Dictionary of Haas (1964), which has been prepared by Doug Cooper, Southeast Asian Software Research Centre, Bangkok.

The three target languages for translation are Assamese, English and Thai. Assamese is included because it is the principal language of wider communication in Assam state, and is also the most direct threat to the long-term survival of these languages. English was chosen because it is a language of wider communication in India, as well as an international language of scholarship, and the native language of the writer. Thai is included both because Banchob (1987) gave translations into Thai, and also because Thai is a widely spoken national language closely related to the Tai languages of Assam.

Further fields can easily be added to this data base, and in the future it could incorporate fields such as:

<proto_southwesternTai_cognate>

<proto_Tai_cognate>

<proto_Tai-Kadai_cognate>

Table (5) is a sample entry from the data base:

```
<number>3
<phake_number>1,1
<lex term="kkq" head="1" sense="1" level="head"/>
<phake>ကောက်
<ahom>ꩫ
<aiton>ကောက်
<aiton_number>1,1
<aiton_ipa>kak³
<aiton_assamese_script>কাক
```

<aiton_english>to keep one portion

<aiton_assamese>ভাগ কৰা

<phake_ipa>kāk⁴

<khamyang_ipa>k

<english_gloss>keep

<phake_english_banchob>to keep one portion

<english_heads>to keep

<phake_assamese_aimya>ভাগ কৰি ৰখা

<assamese_heads>ৰখা/ভাগ <roman {CHK}>

<thai_cognate>

<thai_heads>เก็บ;ไว้

<thai_banchob>แบ่ง_เก็บ_ กัน_ไว้_ส่วน_หนึ่ง

<haas def>

<ex_banchob_tai>ꩫꩬ်_ဠ_ꧠ_ꩰ်_ꧠꩫ်

<ex_banchob_phake_ipa>kāk⁴ wai⁴ cā⁵ cau³ mon²

<ex_banchob_english_banchob>to keep one portion (of food) for a monk

<ex_banchob_assamese>ৰ

<ex_banchob_thai>แบ่ง_เก็บ_กัน_ไว้_สำหรับ_ถวาย_พระ

Table 5: Sample from the Tai dictionary data base

In this sample entry, in the field of <assamese_heads>, the symbol <roman > indicates that the words within the angle brackets should be printed in Roman script. The symbol {CHK} in curly brackets indicates that this entry should be checked. None of these brackets nor any of the information contained in them would appear in any final version of the dictionary.

11. Postscript - presenting this thesis in electronic format

When I commenced this thesis, I had no idea that it would end up being presented in two formats, as both a printed book and in electronic format. I had no conception that it would be possible to do what has been done here. This short chapter presents and discusses some of the issues involved in presenting the thesis in this way.

This chapter will date much more quickly than the rest of the thesis. Within a very short time some of the innovations discussed here may be normal practice in linguistics theses and barely worthy of comment. Future readers will no doubt be amused by some of the blind allies that I have followed. Nevertheless in 2002 it is important to discuss the challenges, problems and successes of presenting this thesis in an electronic format.

11.1 Creating and working with non-standard fonts

Even before I had decided that the study of the Tai languages of Assam would be my thesis topic, their study presented some electronic challenges. As mentioned above in 1.1, the first step that I took was the production of the fonts for the Tai scripts (see 7.8 above). In retrospect this was a very important step which has to some extent influenced everything that has followed. Since the fonts were created before I recorded any examples of spoken text, the Tai script became an integral tool in the analysis of these languages. Even before the phonological analysis was complete, data collection could be undertaken and translation could begin.

The process of creating a font involves not only making the right shapes, but deciding on a keyboard layout. At the time, I was using the Thai version of Windows. In that operating system, there is a switch key between the lower ASCII characters (the keys shown on the keyboard), and the upper ASCII, or hidden keyboard, where various diacritic forms like <â> are to be found. At that time, in Thai Windows, the Thai characters were all placed on the upper ASCII keyboard.

When operated, this switch key would give the character <Ê> when the letter <s> was typed. This is the place on a Thai keyboard where the letter ห /h/ is found. The first draft of the fonts used the Thai keyboard for the cognate characters in the Tai scripts of India, so that ꞕ h was found on the character <Ê>.

The first story collected in January 1998 was *How the Aitons and the Khamtis quarrelled*, which Chaw En Lai Phalung told me during a long walk on the day of a bandh. In the evening I entered the story into the computer, using the Thai based keyboard described above.

I have since decided that it is better to place the Tai characters in place of the Roman characters in the lower ASCII keyboard, as can be seen in the various documents explaining the fonts, (see Table 27 in 7.8 above). However I did not always remember to go back and retype the stories which had been entered using the first drafts of the fonts. Some of this retyping was still being done in the week before the thesis was complete, and it is possible that some examples of the old keyboard are still present in the thesis.

Together with the fonts themselves, explanatory documents have been produced. These briefly how I went about making each font, and list all the characters included in the font, together with the keystrokes required to access those characters. For example, phake.doc (in the \fonts folder) lists all of the details of the Phake font.

Non-standard fonts such as those used in this thesis do cause problems, at least for older versions of the computer operating systems currently available. For example, several of the fonts used in this thesis will not work properly in Windows 95, but will work in Windows 98. Several of the fonts will not work properly in Word 98 but will work in Word 2000.

The fonts used in the earlier versions of Windows with which I began this thesis were 8 bit fonts, meaning that they had 2^8-1 or 255 characters. More recently, 16 bit fonts have been devised, having $2^{16}-1$ or 16,383 characters. With that number of spaces, it

is believed that all the writing systems of the world can be accommodated, and the Unicode consortium was set up to arrange this. Already today Unicode fonts are becoming standard and have many alphabets. A commonly available Unicode font in 2002 is Arial Unicode MS[77] which has all the characters for national writing systems (and those of the states of India), except for Burmese and Cambodian.

It is envisaged by the proponents of Unicode, that it will have a place for all of the writing systems in use in the world. We might hope that a place will be found there for the Tai scripts of Assam and when it is, the non-standard fonts produced for this thesis may become obsolete. However the non-standard fonts used here are necessary to present this thesis in 2002, and it would be surprising if the Tai scripts of Assam were present in Unicode fonts before the year 2010. Even then, Unicode is never likely to cover all of the possible fonts that we might want to use in a thesis like this. Even if it does eventually include characters that will allow us to reproduce Dr. Banchob's orthography for Phake, which is the basis of all the orthographies used in this thesis, it is unlikely that Unicode will ever have a place for the orthography devised by Weidert for Khamti (see 3.2.14.1). Despite the best intentions of the designers of Unicode, non-standard fonts may continue to be necessary.

11.2 Presenting the sound files

It seems self-evident that a thesis on language will be greatly enriched by having access not only to the transcription and translation of language examples, but also by the opportunity to listen to those language examples.

Several issues will be canvassed here. The recording of the texts is discussed in 11.2.1 and the transfer of those recordings into digital format is discussed in 11.2.2. The linking of these sound files to the format in which this thesis is written is discussed in 11.2.3.

---

[77] Should the reader not know about Unicode fonts, but wish to explore them, I suggest that you open a document and select Arial Unicode MS. Then select Insert Symbol (if you are a PC user) and then scroll down the fonts and explore the thousands of characters and dozens of writing systems present there.

11.2.1 Recording the texts

The first stories were collected for this thesis in 1998. They were recorded on a poor quality tape recorder. They are listed, with their dates of collection, in Table (1):

Story name	informant	community	text number	date of collection
How the Aitons and the Khamtis quarrelled	Chaw En Lai	Aiton	2-1-3-1	9/1/1998
Story of the crow and the fox	Ong Cham	Aiton	2-1-3-2	10/1/1998
Story of the old woman	Nabin Shyam	Aiton	2-1-3-3	20/1/1998
The bird with the red bottom	Ngi Kheng Chakap	Phake	1-1-3-1	15/1/1998
Advice to children	Sa Myat Chowlik	Khamyang	3-1-3-2	13/1/1998

Table 1: Stories recorded in 1998.

Each one of these texts has been used in the thesis and several sentences from each are used to exemplify different grammatical features. The poor quality of the recordings will be immediately obvious to the listener, and some may question their use. In part they are used here because they contain examples of grammatical forms not otherwise recorded, but in part they are here to make a point about recording. Any recording is better than none, and these recordings were the best that I could make at the time. Only since 1998 has it been possible to purchase the high quality equipment that I can now use to produce better recordings.

Four different methods of recording have been used. First the cheap tape recorder mentioned above. All other cassette recordings were done on a much better quality cassette recorder. More recently I was able to purchase first a video camera and then a minidisc recorder.

No video recor᷈   ; are presented in this thesis, although sound files of several of these video recordings are included. Presenting video texts would indeed have enriched this thesis, and I venture to suggest that if future linguistics theses include sound as this one has done, they will include video also.

11.2.2 Digitising the texts

Altogether around 60 hours of text has been recorded; only a portion of this could be analysed for this thesis, but all of the texts which are analysed and presented here have been digitised. The texts were recorded onto computer as .wav files, using the SIL Speech Analyzer 1.5.

In total, the various .wav files took up nearly 2 gigabytes of space, which took up three CDs. Therefore, for presentation in this thesis, all of the sound files have been converted to .mp3 format, taking up only 300 megabytes. It is therefore possible to include all of the sound files for this thesis on a single CD.

The Speech Analyzer programme shows wave forms and can produce spectrograms and pitch analyses, some of which were captured as bitmaps linked to sections of Chapter 6 (see for example Table (7) in 6.1.2).

11.2.3 Linking the texts to the thesis

As it became gradually clear to me that it would be possible to present this thesis in electronic format, I was faced with the question of what format that should be. Several people advised me to consider producing a web-page.

Early discussions on this issue with my supervisor, Dr Heather Bowe, revolved around what kind of web-page program might be purchased and how to go about learning to do it. A very small experimental website was created which demonstrated some of the difficulties that would be faced.

The normal format for websites is .html, but .html files cannot support tabulations and non-standard fonts. To change all the tabulations to table format would be possible, although it would have been very time consuming.

To overcome the problem of the use of non-standard fonts, it was suggested that the entire thesis could be converted to .pdf format. This format is now very commonly used for the distribution of documents around the internet, but it is quite inflexible compared with the Word format used in this thesis. The problem with .pdf files is that they cannot be easily searched, and they cannot be easily altered. Furthermore they do not appear as clearly on the screen as this Word document does.

During one stopover in Bangkok, I sought the advice of Doug Cooper, of the Southeast Asian Software Research Centre. He suggested I use the Word format in which the entire thesis has been written.

The most recent versions of the Word program allow for the insertion of hyperlinks. Those hyperlinks can link to any document in any format. They can link to any part of a Word document if that part is bookmarked. Furthermore, the creation of a Table of Contents such as occurs at the head of each chapter builds in its own links.

Therefore the links created here could all be set up without any special skills having to be acquired. Furthermore, all the software necessary to read any part of the thesis is now widely available. A draft of the electronic version of this thesis has even been read in Assam.

The only outstanding problem was how to link to a single sentence within an .mp3 file. Nearly every one of the over 500 language examples in the thesis is linked to a sound file, and most of those are single sentences from a text. The most flexible way would be to effectively bookmark the .mp3 file for a particular text, asking the hyperlink to go to a particular point in the .mp3 file and play a particular length of text.

Programs do exist which will allow this, at least for .wav files, but a direct link out of the word file was not possible. Consequently, it was necessary to go back to the Sound Analyzer program each time I wanted to include a sentence in the text. The .wav file for that text would be opened, the particular sentence found and then cut and paste into a new file which was named according to the number of that sentence

in the analysis of the particular text. Although this was time consuming, it turned out to be a good way of re-checking those sentences.

11.3 The role of the researcher in the electronic age.

In the past, when writing such a thesis, the researcher would assemble all the relevant examples and build up a case for the analysis, just as I have done. Usually, several short texts were analysed and presented in an appendix, but in general the reader had to accept the data as the researcher presented it.

In this thesis, the role of the researcher is different. I have presented my analysis of these languages, just as researchers in the past have done. In addition, I have given the reader the opportunity to go back to the data and check any claim that I have made.

Since becoming aware that a thesis could be presented in electronic format, it has been my primary aim to make it as easy as possible for the reader to check the claims I make. That is why sentence examples are linked in the way they are. The reader may wish to know what kind of text the sentence comes from. They may wish to see the context from which an utterance comes, and I would expect this to be a very valuable tool for languages in which context is so important. And of course the reader will want to hear at least some of the sentences.

By providing a rich data set, linked to a comprehensive analysis of the Tai languages, it is hoped that the readers will be able to get all the information about the languages that they want, and further to be able to check the data and see things that I have not been able to see.

11.4 Archiving the data

Although it is a primary aim of this thesis to provide documentation of the Tai languages for posterity, we cannot be certain that the data contained in this thesis will survive in its present form. Past generations of humans have recorded valuable

information on stone tablets, on papyrus, in tape recordings or in books. No doubt they also hoped that the media they used would survive, but stones may break and papyrus may crumble; tape recordings may degenerate and books may get eaten by moths, or burned in a fire.

With electronic presentation, destruction is always possible, but another danger may present itself. What if the electronic format, the computer program, used today is simply not available in fifty or a hundred years time?

With this in mind, the decision was taken to use the most widely available electronic formats. The thesis is written in Microsoft Word Documents. Documents produced by earlier versions of Microsoft Word can be read on the latest versions, because each update has taken into account the past versions. It is hard to imagine that for the foreseeable future this will not continue to be so.

The other types of files presented in this thesis, .jpg and .bmp for images and .mp3 for sound, are also very widely available, and it is expected that future updates will make allowance for these formats when .jpg and .mp3 files are regarded as somewhat passé.

When Dr. Banchob recorded her Aiton tapes in the 1960s, she used the best available equipment. By the year 2001, her heirs could no longer access the tapes that she had made; those tapes needed to be brought to Australia and copied onto compact discs, the latest technology available. How do we know that in another forty years the compact disc on which this thesis is written will still be readable?

It is hoped that somewhere in the world, some archivists will take on the responsibility of ensuring that this thesis is updated into the latest formats, once the writer is no longer available to do so.

11.5 Possibilities that were not followed up

A PhD thesis is a finite activity, even if at times the contrary appears to be the case. In such a finite activity all the possible ways of undertaking analysis cannot be followed up. At least one area which has not been followed up is the possibility of using a concordance program to go through every example of every word in the texts. This thesis might have been better if this had been done, but it was not.

11.6 Conclusion.

I hope that this thesis has given a good account of the analysis of the Tai languages of Assam. I hope my readers, both in the year 2002 in which this work was written and into the future, will find this analysis useful. I also hope that readers will agree that the inclusion of the texts and the links to sound files are a useful innovation.

In a hundred years time, though aspects of my analysis may not have stood the test of time, the richness of the corpus of texts will ensure that this work is still useful.

ကျွန်ုပ်ဆုတောင်း

Bibliography

In this thesis, the decision has been taken to use the original dates of publication when referring to it in a text, even when it is a later edition that has been consulted. This is felt necessary to give the readers a clear indication of the date at which the reference was written. The casual reader, unaware perhaps of *Linguistic Survey of India*, might see a reference to Grierson (1966) and assume that it was a work of the mid 20th century. Therefore this work will always be referred to as Grierson (1904), with details of the reprints in this bibliography as in the following example:

Grierson, Sir George. 1904. *Linguistic Survey of India*, Vol. 2. *Mon-Khmer and Siamese-Chinese Families*. Calcutta: Office of the Superintendent of Government Printing. Republished in 1966, Delhi: Motilal Banarsidass.

Also, in this thesis, it has been decided to refer to people of Tai origin by their formal names in their own language, rather than by the surnames which Indian citizenship has conferred on them. Thus *The Good Way of Teaching* by Ngi Kheng Chakap and Ai Che Let Hailung will be found in the bibliography under the heading 'Ngi'.

A similar principle will be followed with names of scholars from Thailand. In Thai language bibliographies are listed in the alphabetical order of first names, and this has been followed in this thesis. It was decided that it would inappropriate to refer to Dr Banchob Bandhumedha under her surname, which is little known, particularly in India. Therefore in this bibliography, Thai scholars are listed under their first name, with a reference under their surname to this first name.

Abramson, Arthur S. (ed.) 1997. *Southeast Asian Linguistic Studies in honour of Vichin Panupong*. Bangkok: Chulalongkorn University Press.

Aikhenvald, Alexandra Y. 2000. *Classifiers: a typology of noun categorisation devices*. Oxford: Oxford University Press.

Aimya Khang Gohain. 1991. 'The Tai language as spoken by the Tai-Phake' in *Pan Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics*. Bangkok: Chulalongkorn University. 44-59

Aimya Khang Gohain. (শ্রী আইমা খে গোহাঁই). 1997 (১৯৯৬). *Elementary Tai Primer* (টাই ভাষাৰ প্ৰাথমিক পাঠ ৰুৰবণেৰে হৈতে). Dibrugarh University: Department of Assamese (in Assamese and English).

Aimya Khang Gohain. (trans.). 1999. *Sucha Naraha - Teachings of Gautama*. St. Petersburg, Florida: published by James L. Edwards.

Anderson, Stephen and Bernard Comrie (eds.). 1991. *Tense and Aspect in Eight Langauges of Cameroon*. Arlington: Summer Institute of Linguistics and University of Texas at Arlington.

Banchob Bandhumedha (พรรจบ พันธุเมธา). 1977. (๒๕๒๐). 'Language and Literature - Tai Words, Aiton Language.' (ภาษาและหนังสือ คำไท - ไทย ภาษาอายตอน) in *Satri San* (in Thai).

Banchob Bandhumedha (พรรจบ พันธุเมธา). 1987. (๒๕๓๐). *Phake - Thai - English Dictionary* (พจนานุกรมพาเก - ไทย - อังกฤษ). Published by the Author. (in English and Thai).

Banchob Bandhumedha (พรรจบ พันธุเมธา). ms. *Aiton - Thai - English Dictionary* (พจนานุกรมอายทอน - ไทย - อังกฤษ). (in English and Thai).

Bandhumedha, Banchob. See Banchob Bandhumedha

Bandhumedha, Navavan. See Navavan Bandhumedha

Barua, B. 1966. 'Influence of the Tai-Ahom on Assamese Language'. in *Lik Phān Tai (Journal of the Tai Historical and Cultural Association of Assam)*. 1:60-67.

Barua, B. 1975. 'A note on the Tai Āhom Language'. *Journal of the Department of Assamese* Dibrugarh: Dibrugarh University. 1:60-61.

Barua, Bimala Kanta and N.N. Deodhai Phukan. 1964. *Ahom Lexicons, Based on Original Tai Manuscripts*. Guwahati: Department of Historical and Antiquarian Studies.

Barua, Ghan Kanta. (শ্রী ঘনকান্ত বৰুৱা) 1936. *Ahom Primer* (আহোম প্ৰাইমাৰ). reprinted in 1987 in Guwahati: Department of Historical and Antiquarian Studies in Assam. (In Assamese and English).

Barua, Golap Chandra. 1920. *Ahom-Assamese-English Dictionary*. Calcutta: Baptist Mission Press (printed under the authority of the Assam Administration).

Barua, Golap Chandra. 1930. *Ahom Buranji - from the earliest time to the end of Ahom rule.* reprinted in 1985 in Guwahati: Spectrum Publications.

Baruah, Girin Mohun. (গিৰীন মহৰ বৰুৱা)1999. Loti Amra. (লটি অমৰ) published by the author (In Assamese).

Barz, R.K. and A.V.N. Diller. 1985. 'Classifiers and standardisation: some South and South-East Asian comparisons', in David Bradley (ed) *Papers in South-East Asian Linguistics No. 9 – Language Policy, Language Planning and Sociolinguistics in South-East Asia.* Canberra: Pacific Linguistics.

Bauer, Robert S. and Paul K. Benedict. 1997. *Modern Cantonese Phonology.* Berlin: Mouton de Gruyter.

Becker, A.L. 1993. 'The elusive figures of Burmese grammar: An Essay.' In Foley, William A., (ed.) *The Role of Theory in Language Description.* 61-85. Berlin: Mouton de Gruyter.

Bisang, Walter. 1998. 'Grammaticalization and Language Contact, Constructions and Positions.' in Ramat, Anna Giacalone and Paul Hopper (eds.). *The Limits of Grammaticalization.* Amsterdam: John Benjamins.

Biswas, S. 1966. 'A note on the Khamtis and their Language' in Neog, Maheswar and Mukunda Madhava Sharma (eds.) *Professor Birinchi Kumar Barua Commemoration Volume,* Gauhati: All India Oriental Conference, XXII Session. 169-178.

Boas, Franz and Ella Deloria. 1941. *Dakota Grammar.* Washington: Memoirs of the National Academy of Sciences, Volume XXIII.

Boonyang Kettate, 1998. 'Some aspects of the Tai Khamtis Socio Cultural Traits', in Shalardchai Ramitanondh, Virada Somswasdi and Ranoo Wichasin (ฉลาดชาย รมิตานนท์, วระดา สมสวัสดี และ เรณู วิชาศิลป์) *In ออ: Tai.* Chiang Mai: Chiang Mai University. 144-170.

Boruah, Bhim Kanta, 1992. 'Thai Languages in India: A Linguistic Analysis' in *Pan Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics.* Bangkok: Chulalongkorn University. 2: 854-860.

Boruah, Bhimkanta. 2001. 'Tai Language in India - an Introduction' in *Indian Journal of Tai Studies.* Moranhat, Assam: Institute of Tai Studies and Research. 1:30-33.

Boruah, P. N. Dutta. 1980. *An intensive course in Assamese.* Mysore: Central Institute of Indian Languages.

Bowe, Heather. 1990. *Categories, Constituents and Constituent Order in Pitjantjatjara, an Aboriginal language of Australia.* London: Routledge.

Bowe, Heather & Stephen Morey. 2000. *Yorta Yorta (Bangerang) language of the Murray-Goulburn including Yabula Yabula.* Canberra: Pacific Linguistics.

Bradley, David and Maya Bradley. (eds) 2002. *Language Maintenance for Endangered Languages: An Active Approach.* London: Curzon Press.

Brown, J. Marvin. 1985. *From Ancient Thai to Modern Dialects.* Bangkok: White Lotus.

Brown, Rev. N., 1837a. 'Alphabets of the Tai Language.' in *Journal of the Asiatic Society of Bengal* 6:17-21.

Brown, Rev. N., 1837b. 'Interpretation of the Ahom extract, published as Plate IV of the January number of the present volume. By Major F. Jenkins, Commissioner in Assam.' in *Journal of the Asiatic Society of Bengal* 6:980-983.

Brugman, Claudia. 1996. 'Inalienability and the Interpretation of Modified Noun Phrases', in Shibatani, Masayoshi and Sandra A. Thompson. (eds.) *Grammatical Constructions - Their Form and Meaning.* Oxford: Clarendon Press.

Buchanan, Francis. 1799. 'A comparative Vocabulary of some of the Languages spoken in the Burma Empire' in *Asiatick Researches* 5:2:9ff

Buragohain, Jaya. 1998. 'The Tai Aitons of Assam' in Shalardchai Ramitanondh, Virada Somswasdi and Ranoo Wichasin (ฉลาดชาย รมิตานนท์, วระดา สมสวัสดี และ เรณู วิชาศิลป์) *In ออ: Tai.* Chiang Mai: Chiang Mai University. 57-113.

Buragohain, Ye Hom. 1981. 'Old units of time of the Tai people of North-Eastern India'. in *Lik Phān Tai (Journal of the Tai Historical and Cultural Association of Assam)* Vol. II. 6-9.

Buragohain, Ye Hom. 1998. 'Some notes on the Tai Phakes of Assam', in Shalardchai Ramitanondh, Virada Somswasdi and Ranoo Wichasin (ฉลาดชาย รมิตานนท์, วระดา สมสวัสดี และ เรณู วิชาศิลป์) *In ออ: Tai.* Chiang Mai: Chiang Mai University. 126-143.

Bybee, Joan; Revere Perkins and William Pagliuca. 1994. *The Evolution of Grammar - Tense, Aspect and Modality in the Languages of the World.* Chicago: University of Chicago Press.

Cameron, Deborah, Elizabeth Frazer, Penelope Harvey, M.B.H. Rampton and Kay Richardson. 1992. *Researching Language - Issues of power and method.* London: Routledge.

Campbell, Sir George. 1874. *Specimens of the Languages of India, including those of the Aboriginal Tribes of Bengal, the Central Provinces and the Eastern Frontier.* Calcutta: Bengal Secretariat Press.

Chamberlain, James R. 1975. 'A new look at the history and classification of the Tai languages' in J. Harris and J. Chamberlain (eds) *Studies in Tai linguistics in honor of William J. Gedney.* Bangkok: Central Institute of English Language. 49-66.

Chau Khouk Manpoong. ('ÇÈ à¢Ò¡ ÁÒ'È»Ø§). 1993. *New Tai Reader* (Êâ' Áã µäö) Chongkham, Arunachal Pradesh: Tai Literature Committee. 2 volumes. (in Khamti).

Chelliah, Shobhana L. 1992. 'Tone in Manipuri' in Ratliff, Martha and Eric Schiller, (eds). *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society 1991.* Arizona State University. 65-85.

Codrington, R.H. & Palmer B.D. 1896. *A Dictionary of the Language of Mota* London: Society for Promoting Christian Knowledge .

Comrie, Bernard. 1976. *Aspect*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press.

Comrie, Bernard. 1989. *Language Universals and Linguistic Typology*. Chicago: University of Chicago Press. (2nd edition)

Comrie, Bernard and Norval Smith. 1977. 'Lingua Descriptive Studies: questionnaire.' *Lingua* 42:1-72.

Cooper, Doug. 1993. *Oh! Pascal!*. New York: W.W. Norton & Company.

Cowell, E.B., 1907. *Jātaka or Stories of the Buddha's former Births*. Cambridge University P- s. Volume 6. (republished in 1990)

Cushing, Rev. J.N. 1888. *Elementary Handbook of the Shan Language*. Rangoon: American Baptist Missionary Press.

Dalton, E.T. 1872. *Descriptive Ethnology of Bengal*. Calcutta: Office of the Superintendent of Government Printing.

Damant, G.H. 1880. 'Notes on the Locality and Population of the Tribes dwelling between the Brahmaputra and Ningthi Rivers.' in *Journal of the Royal Asiatic Society* xii:228 ff CHK.

Dasgupta, Bidhu Bhusan. 1993. *Assamese Self-Taught*. Calcutta: Das Gupta Prakashan.

Dhongde, R.V. & Phramaha Wilaisak Kingcom. 1992. 'Standard-Thai & Tai-Phake - a comparison of Phonology, - a comparison of Morphology.' in *Pan Asiatic Linguistics: Proceedings of the Third International Symposium on Language and Linguistics*. Bangkok: Chulalongkorn University Vol II: 735-745.

Diller, Anthony. 1988. 'Tai scripts and Proto-Tai' in *The International Symposium on Language and Linguistics – Faculty of Liberal Arts, Thammasat University 9-11 August 1988*. Bangkok: Thammasat University. 228-247

Diller, Anthony. 1992. 'Tai Languages in Assam: Daughters or Ghosts?' in Carol J. Compton and John F. Hartmann (eds). *Papers on Tai Languages, Linguistics and Literatures*. Northern Illinois Centre for Southeast Asian Studies. Paper #C. 5-43

Diller, Anthony. 1994. 'Tai Languages: Varieties and Subgroup Terms'. Thai-Yunnan Project Newsletter 25:8-17

Diller, Anthony. 2001. 'Grammaticalization and Tai Syntactic Change' in M.R. Kalaya Tingsabadh and Arthur Abramson, *Essays in Tai Linguistics*. Bangkok: Chulalongkorn University Press. 139-175.

Dixon, R.M.W. 1972. *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 1997. *The rise and fall of languages*. Cambridge: Cambridge University Press.

Dixon, R.M.W. 2001a. 'The typology of demonstratives and related items.' Research Centre for Linguistic Typology: Position paper for workshop

Dixon, R.M.W. 2001b. 'Adjective Classes.' Research Centre for Linguistic Typology: Position paper for workshop

Dixon, R.M.W. 2002. 'Copula clauses and verbless clauses.' Research Centre for Linguistic Typology: Position paper for workshop

Duanmu, San. 1992. 'A featural analysis of some onset-vowel interactions." in Ratliff, Martha and Eric Schiller, (eds). *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society 1991*. Arizona State University. 141-158.

Edmondson, Jerold A. and David B. Solnit. 1997a. 'Introduction' in Edmondson, Jerold A. & David B. Solnit (eds.) *Comparative Kadai - The Tai Branch*. Arlington: University of Texas. 1-27.

Edmondson. Jerold A. and David B. Solnit. 1997b. 'Comparative Shan' in Edmondson, Jerold A. & David B. Solnit (eds.) *Comparative Kadai - The Tai Branch*. Arlington: University of Texas. 337-359.

Egerod, Søren., 1957. 'Essentials of Shan phonology and script'. in *Academia Sinica: Bulletin of the Institute of History and Philology*. 29:121 -7.

Egerod, Søren., 1959. 'Essentials of Khün phonology and script'. in *Acta Orientalia* 24:123-146.

Enfield, N.J. 2000. *Linguistic Epidemiology - On the polyfunctionality of 'acquire' in mainland Southeast Asia*. PhD Thesis, Melbourne University.

Fausböll, V. (ed.) 1896. *Jātaka with Commentary*. Oxford: Pali Text Society. Volume 6, republished by the Pali Text Society in 1990.

Fillmore, Charles J. 1968. 'The Case for Case' in Emmon Bach and Robert T. Harms (eds) *Universals in Linguistic Theory*. New York: Holt, Rinehart and Winston. 1-88.

Foley, William A. and Robert D. Van Valin. 1985 'Information packaging in the clause' in Timothy Shopen (ed.) *Language typology and syntactic description, Volume 1: Clause structure*. Cambridge: Cambridge University Press.

Foley, William A. 1986. *The Papuan Languages of New Guinea*. Cambridge: Cambridge University Press.

Fromkin, Victoria A. 1978. *Tone - A Linguistic Survey*. New York: Academic Press.

Gait, Sir. Edward, 1905. *A History of Assam*. republished in 1992, Guwahati: Lawyers Book Stall.

Gedney, William J. 1972. 'A Checklist for Determining Tones in Tai Dialects' in M. Estelle Smith (ed). *Studies in Linguistics in honor of George L. Trager*. The Hague: Mouton. 423-37.

Givón, T. (ed) 1997. *Grammatical Relations - a Functionalist Perspective*. Amsterdam: John Benjamins.

Glick, Irving I. & Sao Tern Moeng. 1991. *Shan for English Speakers - Dialogues, Readings, Grammar Notes and Vocabulary*. Wheaton, Maryland: Dunwoody Press.

Glick, Irving I. & Sao Tern Moeng. 1996. *Shan Newspaper Reader*. Kensington, Maryland: Dunwoody Press.

Gogoi, Lila. (ed.). 1971. *The Tai Khamtis of the North-East*. New Delhi: Omsons Publications.Gogoi, Nomal Chandra. 1987. *The Assamese English Tai Dictionary*. Tinsukia: published by the author.

Gogoi, Nomal Chandra. 1987. *The Assamese English Tai Dictionary*. Tinsukia: published by the author.

Gogoi, Nomal Chandra. 1994. *Morphological Study of the Tai Phake Language*. PhD. Thesis, Dibrugarh University.

Gogoi, Padmeswar. 1968. *The Tai and the Tai Kingdoms; with a Fuller Treatment of the Tai-Ahom Kingdom in the Brahmaputra Valley*. Gauhati: Department of Publication, Gauhati University.

Gogoi, Padmeswar. 1976. *Tai Ahom Religion and Customs*. Gauhati: Publication Board, Assam.

Gogoi, Puspa. 1993. *Descriptive Catalogue of Tai Manuscripts*. Dhemaji: Ban-Ok Pup-Lik Moung-Tai.

Gohain, Ngowken. 2001. 'The Tai-Phakes of Assam' in *Indian Journal of Tai Studies*. Moranhat, Assam: Institute of Tai Studies and Research, 1:30-33.

Goldberg, Adele. 1996. 'Making One's Way Through the Data', in Shibatani, Masayoshi and Sandra A. Thompson. (eds.) *Grammatical Constructions - Their Form and Meaning*. Oxford: Clarendon Press.

Greenberg, Joseph H. 1990. 'Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements'. in Keith Denning and Suzanne Kemmer (eds) *On language. Selected writings of Joseph H. Greenberg*. Stanford: Stanford University Press. 41-70

Greenberg, Joseph H. (ed). 1966. *Universals of language*. Cambridge (Mass): The M.I.T. Press.

Gregerson, Kenneth J. & Jerold A. Edmondson. 1998. 'Some puzzles in Cao Lan'. in *Proceedings of the International Conference on Tai Studies, July 29-31, 1998*. Bangkok: Mahidol University.

Grierson, Sir George. 1904. *Linguistic Survey of India*, Vol. 2. *Mon-Khmer and Siamese-Chinese Families*. Calcutta: Office of the Superintendent of Government Printing. Republished in 1966, Delhi: Motilal Banarsidass.

Gurdon, P.R. 1895. 'On the Khāmtīs' in *Journal of the Royal Asiatic Society*. 27:157-164.

Haas, Mary R. 1964. *Thai–English Students Dictionary*. Stanford: Stanford University Press.

Harris, Jimmy G. 1976. 'Notes on Khamti Shan'. in *Thai Linguistics in honour of Fang Kuei Li*. Bangkok: Chulalongkorn University. 113-141.

Hartmann, John F. 1984. *Linguistic and Memory Structures in Tai-Lue Oral Narratives*. Pacific Linguistics Series B - No 90. Canberra: Australian National University.

Hazarika, Chow Nogen. 1993. ꠡꠦꠇꠦ *(Heaven)*. Dibrugarh: Ban Ok Pup Lik Mioung Tai.

Hazarika, Chow Nogen. 1996. *We Revive We Survive*. Guwahati: Published by Nang Kalpana Hazarika.

Hazarika, Sanjoy. 1995 (1994). *Strangers of the Mist - Tales of War & Peace from India's Northeast*. New Delhi: Penguin.

Himmelmann, N.P. 1998. 'Documentary and descriptive linguistics'. *Linguistics 36*. 161-195.

Hodgson, B.H. 1850. 'Aborigines of the North-East Frontier. in *Journal of the Asiatic Society of Bengal*. 19:309-316.

Hudak, Thomas John. 1990 (1987) 'Thai' in Bernard Comrie (ed). *The major languages of East and South-east Asia*. London:Routledge. 29-47.

Hunter, Sir W.W. 1868. *A comparative Dictionary of the Languages of India and High Asia*. London: Trubner

Hyman, Larry M. (ed.). 1979. *Aghem Grammatical Structure*. Southern California Occasional Papers in Linguistics, No 7. Los Angeles: University of Southern California.

Indian Atlas. 1871. India: Trigonometrical Survey.

Jagacinski, Ngampit. 1992 '*wa 2a* and complement-taking predicates in Thai.' in Ratliff, Martha and Eric Schiller, (eds). *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society 1991*. Arizona State University, 205-223.

Jones, Sir William. 1799. *The Works of Sir William Jones*. Six vols. London: G.G. and J. Robinson.

Jonsson, Nanna L, D.A. 1991 *Proto Southwestern Tai*. Dissertation for the Degree of Doctor of Arts, State University of New York, Albany.

Judson, A. 1893. *Burmese-English Dictionary* (revised and enlarged by Robert C. Stevenson). Rangoon: Superintendent of Government Printing, Burma. Republished in 1993. New Delhi: Asian Educational Services.

Kanittanan, Wilaiwan. See Wilaiwan Kanittanan.

Kenstowicz, Michael and Charles Kisseberth. 1979. *Generative Phonology, Description and Theory*. New York: Academic Press.

Kettatte, Boonyong, See Boonyong Kettatte.

Kingcom, Phramaha Wilaisak. 1992. *A comparative study of Standard-Thai and Tai-Phake Spoken in Assam (India)*. Doctoral thesis submitted to the Department of Linguistics, Deccan College Postgraduate and Research Institute, Poona.

Kullavanija, Pranee. See Pranee Kullavanija.

Lambert, Dorothy Mack. 1969. *The Semantic Syntax of Metaphor: A Case Grammar Analysis*. Ph.D. dissertation, University of Michigan.

Leach, Edmund. 1972. 'The structure of symbolism' in J.S. La Fontaine (ed.) *The Interpretation of Ritual - Essays in Honour of A.I. Richards*. London:Tavistock Publications.

Lekawatana, Pongsri. See Pongsri Lekawatana.

Leyden, J. 1808. 'On the Language and Literature of the Indo-Chinese Nations. *Asiatick Researches* Vol x:158-289

Li, Fang-kuei. 1977. *A handbook of comparative Tai*. Honolulu: University of Hawaii.

Long Yaohong and Zheng Guoqiao. 199°. *The Dong Language in Guizhou Province, China*. Arlington: Summer Institute of Linguistics and University of Texas at Arlington. (translated by D. Norman Geary).

Luo, Yongxian. 1997. *The Subgroup Structure of the Tai Languages: A Historical-Comparative Study*. Journal of Chinese Linguistics: Monograph Series No. 12.

Luo, Yongxian. 1999. *A Dictionary of Dehong, Southwest China*. Canberra: Pacific Linguistics. C-145.

Manpoong, Chau Khouk. See Chau Khouk Manpoong.

Mithun, Marianne. 1998. 'The significance of diversity in language endangerment and preservation'. in Lenore A. Grenoble and Lindsay J. Whaley (eds) *Endangered languages. Current issues and future perspectives*. Cambridge: Cambridge University Press. 163-191

Morey, Stephen. 1998. *The Verbal system of the Central Victorian Language, the Aboriginal Language of Melbourne – an investigation into the manuscripts of Rev. William Thomas (1793-1867)*. Unpublished Honours Thesis, Monash University.

Morey, Stephen. 1999a. သိက်ဃျးရ်ဂ္ခော3 (ပျိန္ဒ်603/�513၀၁၁) (Book for teaching the Tai Language - Tai Aiton Primer). Dibrugarh: printed by Triograph Press.

Morey, Stephen. 1999b. သိက်ဝ်ဂ္ဂိ603/�513၀၁၁ (Tai Aiton History). Dibrugarh: printed by Triograph Press.

Morey, Stephen. 1999c. သိက်ဃျးရ်ဂ္ခော3 (ပျိန္ဒ်603/၅၇}၇၇) (Book for teaching the Tai Language - Tai Phake Primer). Dibrugarh: printed by Triograph Press.

Morey, Stephen. 2001a. သိက်ဒ္ဂ်ဃျ၁ရ် (Book of calling the *Khon*). Dibrugarh: printed by Triograph Press.

Morey, Stephen. 2001b. သိက်ဒ္ဂ်ဃျ၁ရ်လ၁ရ် (Grandfather teaches Grandchildren). Dibrugarh: printed by Triograph Press.

Morey, Stephen. 2001c 'The literature of the Tai of Assam' in *Melbourne Papers in Linguistics and Applied Linguistics*. Volume 1, Issue 1.

Morey, Stephen. 2002a. 'Tai languages of Assam, a progress report – Does anything remain of the Tai Ahom language?' in David and Maya Bradley, (eds). *Language Maintenance for Endangered Languages: An Active Approach*. London: Curzon Press. 98-113

Morey, Stephen. 2002b. 'The study and revival of the Ahom Language' in *Indian Journal of Tai Studies*, 2:89-103

Mühlhäusler, P. 1990. '"Reducing" Pacific languages to writing', in J. Joseph and T. Taylor (eds.) *Ideologies of Language*. London: Routledge, 189-205.

Nartsupha, Chatthip and Ranoo Wichasin. 1998. 'Tai cultural revival in Assam' in Shalardchai Ramitanondh, Virada Somswasdi and Ranoo Wichasin (ฉลาดชาย

รมิตานนท์, วระดา สมสวัสดิ์ และ เรณู วิชาศิลป์) ใน ๗: *Tai*. Chiang Mai: Chiang Mai University. 171-194.

Nartsupha, Chatthip; Ranoo Wichasin, Ranoo Arthamesara and Nongnut Chantarabhaya (ฉัตรทิพย์ นาถสถภา, เรณู วิชาศิลป์, เรณู อรรฐาเมศร์ และ นงนุช จันทราภัย). 1991 (๒๕๓๔). *Muang Nun Sun Kham* (เมืองนุนสุนคำ) Bangkok: Sangsarga Printing House

Navavan Bandhumedha. 2002. *The life of Professor Dr. Banchob Bandhumedha*. Bangkok: Centre for Social Development.

Needham, J.F. 1894. *Outline Grammar of the ဝဒ (Khâmtî) Language - as spoken by the Khâmtis residing in the neighbourhood of Sadiya*. Rangoon: Superintendent of Government Printing, Burma .

Nichols, Johanna. 1986. 'Head-marking and dependent-marking grammar'. *Language* 62:56-119

Ngi Kheng Chakap and Ai Che Let Hailung (ငိခ္ဂိင္ဒ်ဃျ၀ရ် ၁၁ဃျ၁ရ် ၁၁ျ၁�773 ကိဒ/လ၁ဝ8/ရ် ၃ဝ၃္ဒ်). 2001. (၁၃၆]). *The Good Way of Teaching*. (ဃျး၄ဝ၁၄/ရ်) Naharkatia: published by Dr. N.W. Gohain with computing by Dipangkar Baruah.

Noss, Richard B. 1964. *Thai Reference Grammar*. Washington DC: Foreign Service Institute.

Panjok, Muhi Chandra Shyam. (ed.) 1981. 'History of Tai Khamyang Group of Great Tai Race' - paper read and submitted to the *International Conference of Tai Studies*, New Delhi, February 1981.

Panupong, Vichin. See Vichin Panupong.

Payne, Thomas E. 1997. *Describing Morphosyntax - a guide for field linguists*. Cambridge: Cambridge University Press.

Phalung, Nabin Shyam. 1998. 'A Brief Survey on Tai Aitons', in Shalardchai Ramitanondh, Virada Somswasdi and Ranoo Wichasin (ฉลาดชาย รมิตานนท์, วระดา สมสวัสดิ์ และ เรณู วิชาศิลป์) ใน ๗: *Tai*. Chiang Mai: Chiang Mai University. 114-125.

Phraya Upakit Silapasarn (พระยาอุปกิตศิลปสาร). 1918-1937. *The basis of the Thai Language* (หลักภาษาไทย) republished in 1995 in Bangkok: Thai Wattana Panij. (in Thai).

Phukan, J.N. & Buragohain, P.C. 1966. 'A note on Tai-Ahom Couplets'. in *Lik Phän Tai (Journal of the Tai Historical and Cultural Association of Assam)* 1:54-59.

Phukan, J.N. 1966. 'The Tai-Ahom Language'. in *Lik Phän Tai (Journal of the Tai Historical and Cultural Association of Assam)* 1:2-23.

Phukan, J.N. 1998. 'Language and script of the Ahom in the thirteenth century.' Paper presented at the Seminar in Chiang Mai, Thailand on 26th August 1998 on the occasion of the Celebration of the 770th Anniversary of Chao-lung Siu-Ka-Pha.

Phukan, Punaram Mohan. (ফুকন, পোনাৰাম মহন) 1998. (১৯৯৮). *Tai Ahom Vocabulary* (তাই আহোম শব্দকোষ), Dibrugarh: Professor Girin Phukan (in Assamese).

Pongs. ... atana. 1970. *Verb Phrases in Thai: A Study in Deep Case Relationships*. Ph.D. dissertation, University of Michigan

Pranee Kullavanija et al (eds.) 1984, *Six Tai Lexicons*. Bangkok: Centre of Thai Language and Literature, Chulalongkorn University. (based on *Tai Yu Jian Tsu* by Yu Sui Yong and Lo Mei Jen, Peking, Min Ju Shu Pan Tse, 1979)

Ramat, Anna Giacalone and Paul Hopper (eds.). 1998. *The Limits of Grammaticalization*. Amsterdam: John Benjamins.

Rangkupan, Suda. 1997. *An Invesigation of hây Complex Constructi...s in Thai.* available on the internet at http://wings.buffalo.edu/linguistics/rrg

Ranoo Wichasin (เรณู วิชาศิลป์). 1986 (๒๕๒๙). *The writing system of the Tai-Ahoms* (ระบบการเขียนของไทอาหม). MA Thesis: Chiang Mai University.

Ranoo Wichasin (เรณู วิชาศิลป์). 1996 (๒๕๓๙). *Ahom Buranji* (พงศาวดารไทอาหม). Bangkok: Amarin Printing & Publishing. (in Thai).

Ranoo Wichasin. (เรณู วิชาศิลป์) forthcoming. 'Alphabets of the Shan' (อักษรกลุ่มไทใหญ่) (in Thai)

Ranoo Wichasin. (เรณู วิชาศิลป์) forthcoming. 'The State of Studies on Shan (Tai Yai) and Ahom Manuscripts'. (สถานภาพการศึกษาเอกสารไทใหญ่ และไทอาหม) (in Thai)

Reesink, Ger P. 1999. *A Grammar of Hatam, Bird's Heade Peninsula, Irian Jaya.* Canberra: Pacific Linguistics. C-146.

Rhydwen, Mari. 1996. *Writing on the backs of the blacks: voice, literacy and community in Kriol fieldwork.* St Lucia: University of Queensland Press.

Robinson, W. 1849. 'Notes on the Languages spoken by the various Tribes inhabiting the Valley of Asam and its mountain confines'. in *Journal of the Asiatic Society of Bengal.* 18 (Part 1):311-318, 342-349.

Rose, David. 2001. *The Western Desert Code - An Australian Cryptogrammar.* Canberra: Pacific Linguistics.

Sandefur, J. R. 1986. *A Language Coming of Age: Kriol of North Australia. SIL-AAB Series A Vol 10.* Darwin: Summer Institute of Linguistics.

Sao Tern Moeng. 1995. *Shan - English Dictionary.* Kensington, MD, USA: Dunwoody Press.

Saul, Janice E. and Nancy Frieberger Wilson. 1980. *Nung Grammar.* Arlington: Summer Institute of Linguistics and The University of Texas at Arlington.

Sharma Thakur, G.C. 1982. *The Tai Phakes of Assam.* Delhi: B.R. Publishing Corporation.

Terwiel, B.J. n.d. 'Ahom Script: Its Age and Provenance'. Draft article (copy supplied by Atul Borgohain).

Terwiel, B.J. n.d. 'The Rotating Naga'. Draft article (copy supplied by Atul Borgohain).

Terwiel, B.J. 1978. 'The origin of the Ta'i peoples reconsidered'. in *Oriens Extremus.* No 25 (2):239-258.

Terwiel, B.J. 1980. *The Tai of Assam and Ancient Tai Ritual, Volume I, Life Style Ceremonies,* Gaya: Centre for South East Asian Studies.

Terwiel, B.J. 1981. *The Tai of Assam and Ancient Tai Ritual, Volume II, Sacrifices and Time-reckoning,* Gaya: Centre for South East Asian Studies.

Terwiel, B.J. 1988. 'Reading a dead language: Tai Ahom and the Dictionaries.' in Bradley, D, E.J.A. Henderson & M. Mazaudon (eds.). *Prosodic analysis and Asian lingustics - to honour R.K. Spriggs.* Canberra: Australian National University.

Terwiel, B.J. 1989. ' Neo-Ahom and the Parable of the Prodigal Son' in *Bijdragen tot de Taal-, Land- en Volkenkunde.* 145-1:125-145.

Terwiel, B.J. 1995. 'Funerals and the Ritual Tug-Of-War in Southeast Asia - an essay on comparability.' in *International Institute for Asian Studies Yearbook 1995.* 72-82.

Terwiel, B.J. 1996. 'Recreating the Past: Revivalism in Northeastern India' in *Bijdragen - Journal of the Royal Institute of Linguistics and Anthropology,* (Leiden) 152:275-292.

Terwiel, B.J.and Ranoo Wichasin (eds. and transls.). 1992. *Tai Ahoms and the Stars; Three Ritual Texts to Ward off Danger.* Ithaca: Cornell University SEAP .

Trask, Robert Lawrence. 1999. *A Dictionary of Grammatical Terms in Linguistics.* London: Routledge.

Turung, Sri Lecham Chandra. 1986. 'A Short History of the Thai Turung of Assam' in *Silapakorn (ศิลปากร)* Vol 29, No. 6:64 -66.

Van Valin Jr., Robert D. 2001. *Introduction to Syntax.* Cambridge: Cambridge University Press.

Van Valin, Robert D. and David P. Wilkins. 1996. 'The Case for 'Effector': Case Roles, Agents and Agency Revisited', in Shibatani, Masayoshi and Sandra A. Thompson. (eds.) *Grammatical Constructions - Their Form and Meaning.* Oxford: Clarendon Press.

Vichin Panupong (nee Chantavibulya). 1970. *Inter-Sentence Relations in Modern Conversational Thai.* Bangkok: The Siam Society.

Weidert, Alfons. 1977. *Tai-Khamti Phonology and Vocabulary.* Franz Steiner Verlag, Wiesbaden.

Weidert, Alfons. 1979. 'The Reconstruction of the Ahom Tone-system.' in *Zeitschrift der Deutschen Morgeländischen Gesellschaft* (Journal of the German Oriental Academy) 1979 . (in German).

Wichasin, Ranoo. See Ranoo Wichasin.

Wilaiwan Khanittanan (วิลัยวรรณ์ ขนิษฐานันท). 1983. (๒๔๒๖) 'Language of the Tai Phakes' (ภาษาไทพาเก) in *Linguistics along Historical Lines - the Evolution of Thai and English. (ภาษาศาสตร์เชิงประวัติ - วิวั... การภาษาไทยและ ภาษาอังกฤษ).* Bangkok: Thammasat University (in Thai).

Wilaiwan Kanittanan. 1986. 'Kamti Tai: from an SVO to an SOV Language' in B.H. Krishnamurti. (ed). *South Asian Linguistics - Structure, Convergence and Diglossia.* Delhi: Motilal Barnarsidas. 174-178.

Wong-opasi, Uthaiwan. 1992. 'The interplay between tone, stress and syllabification in Thai" in Ratliff, Martha and Eric Schiller, (eds). *Papers from the First Annual Meeting of the Southeast Asian Linguistics Society 1991.* Arizona State University. p 441-481.

Wright, Pamela Sue. 1996. *A Lao Grammar for Language Learners.* Bangkok: Thammasat University.

Young, Linda Wai Ling 1985. *Shan Chrestomathy.* Berkeley: Centre for South and Southeast Asian Studies, University of California.