Shan-Ni grammar and processes of linguistic change

Carmen Eva Marseille

Research Master Linguistics Thesis
Leiden University
Supervisor: Dr. Mily Crevels
May 2019
For questions and comments regarding this thesis, please get in touch by sending an e-mail to carmenm@live.nl
Abstract

This thesis is the first descriptive work on Shan-Ni, a Tai-Kadai language spoken in Kachin state and Sagaing region of Northern Myanmar. Being a Tai language in longterm close contact with several Tibeto-Burman languages, Shan-Ni has several features that are not common in other Tai languages, but do show similarities with Tibeto-Burman languages. The frequency of disyllabic words, the presence of different grammatical markers including TAM markers, and the variation in word order distinguishes Shan-Ni in particular. This thesis does not only describe these features, but also connects them to their presence in other languages, including both Tai-Kadai and Tibeto-Burman languages, Shan-Ni is in contact with. Some features of Shan-Ni are partially present in other Tai languages, but have developed further or in a different direction from certain points in history, which correspond with periods of migration. Through its grammar, Shan-Ni indicates relations with other Southwestern Tai languages of the Northern tier spoken both at the Myanmar-Chinese and Myanmar-Indian border. The expression of Tibeto-Burman-like constructions is made possible through the addition or different usage of grammatical markers, which nonetheless do have a Tai etymology.
# Table of contents

Abstract 3  
Acknowledgements 7  
Abbreviations 8  
List of figures 9  
List of tables 10  

Preface 11  

## 1 The Shan-Ni and their language 13  
1.1 Sociolinguistic situation 13  
1.2 History 19  
1.3 Genetic affiliation 23  
1.4 Aims of this thesis 26  
1.5 Methods 27  

## 2 Phonology and orthography 32  
2.1 Consonants 32  
2.1.1 Consonant clusters 34  
2.1.2 Final consonants 34  
2.2 Vowels 35  
2.3 Tone 37  
2.4 Syllables 39  
2.5 Allophones and phonological variation 40  

## 3 The shape of words 42  
3.1 Compounding 42  
3.1.1 The formation of class terms 45  
3.1.2 Two words becoming one 47  
3.1.3 Personal pronouns 48  
3.2 Contractions 50  
3.2.1 Determiners 51  
3.2.2 Interrogatives 53  
3.3 Reduplication 54  
3.3.1 Remnants of an old system 55  
3.4 Summary 57  

## 4 Sentence structure and types of sentences 59  
4.1 Representing and omitting referents 59  
4.1.1 Ellipsis 60  
4.1.2 Activating referents wit naj⁵/ne² 61  
4.1.3 Indefinite referents and classifiers constructions 62  
4.2 Copula 65  
4.3 Existential and possessive phrases with yang⁴ 68
Acknowledgements

This research would have been impossible without the help of many amazing people. Research in rural Myanmar can be difficult, but the kindness and support of the people I have met throughout my research made it a lot easier. Of course there are a few names I want to mention in particular. Each in their own way, they have helped me to make this thesis a little better. Of course any mistakes are mine.

First of all, the speakers Nan Ee Pyu, Khin Shwe Nge, Daw Nyint Yee, Daw Shwe, Ma Khaing, Daw Sin, Ayee Thida, U Chit Maung, U Than Sein, Daw Cho Me, Sayadaw Pannya Vamsa, Daw Khin Pyone Yee and Naan Nwe Ni Htun who have worked with me with a lot of patience, and this work would not have been there without them. Sometimes friends or family would come in during interviews and comment on things as well, of course I am also grateful for their contributions.

Outside Indawgyi, I would like to thank the people of the Voice of Shan-Ni newspaper, the Shan-Ni Youth Network, and the teacher trainers from Shan-Ni literacy course in Mingun and U Khin Maung Aye, for their kindness, hospitality and helpfulness.

This is the first work about Shan-Ni in English, but there are others working on Shan-Ni in Burmese and Shan-Ni itself. The authorities in this are Daw Khin Pyone Yee and Naan Nwe Ni Htun, and their willingness to sit down with me and explain things about the language has been very important in developing a deeper understanding of the language. Being able to read their work at home also enabled me to continue studying after coming home.

The idea to do research on Shan-Ni at Indawgyi came from a conversation with Stephen Traina-Dorge. He and Patrick Compton have been very supportive in my preparation and throughout my research, helping out with all the practical matters and introducing me to people in the beginning. Conversations with them, Marshall Kramer and Laur Kiik also helped me to understand more about Kachin state in general. I also thank Wyn Owen for sending me his worlists, and André Müller for helping out with the tones and sending me his questionaires. I also thank Sarah Fairman, Diana Sabogal and Maïra Kaye for proofreading parts of the thesis.

My supervisor, Mily Crevels, has always been ready to help, not only with the academic issues but also with the practical and bureaucratic matters related to fieldwork and research. Her willingness to meet frequently and be responsive to any questions both during my fieldwork as well as in the writing process has been a great support for me in this research.

A big thank you to all of you!
List of figures

Figure 1 Areas where Shan-Ni is spoken (within red circles) and the Indawgyi lake (blue square) 11
Figure 2 Shwe Myitzu pagoda 13
Figure 3 Map of Indawgyi with major villages and nearby cities 14
Figure 4 Visual representation of Indawgyi villages and their names 16
Figure 5 Map of the Nan-Chao empire in the 8th & 9th century 18
Figure 6 Parabaik manuscript 1179 AD 19
Figure 7 The Tai-Kadai language family 21
Figure 8 Migration history of Tai groups around northern Myanmar 24
Figure 9 Tone boxes of different Tai languages of the Southwestern Tai language family 112
Figure 10 Shan-Ni’s relationship with Khantii, Phake and Tai Mao 123
List of tables

Table 1 Selection of Indawgyi village names and their meaning  
Table 2 Speakers and types of contributions  
Table 3 Shan-Ni Consonant inventory  
Table 4 Minimal pairs distinguishing /pʰ/ and /f/  
Table 5 Consonant inventory and transliteration for this study  
Table 6 Consonant clusters  
Table 7 Final consonants  
Table 8 Vowel inventory  
Table 9 Overview of vowels and their symbols for open and closed syllables  
Table 10 Diphtongs  
Table 11 Tones in Shan-Ni (Indawgyi)  
Table 12 Gedney's box (adapted from Gedney 1989: 2002)  
Table 13 Gedney's box with Shan-Ni tonal categories and samples in each box  
Table 14 Tones in Myitkyina Shan-Ni (adapted from Wyn Own 2011)  
Table 15 Consonant inventory and transliteration for this study  
Table 16 Consonant clusters  
Table 17 Disyllabic words with a <Cə¹> structure as initial syllable  
Table 18 Overview of allophones and phonological variation  
Table 19 Examples of compounds  
Table 20 Body part compounds and their equivalents in Shan Gyi  
Table 21 Words with unmarked consonant as initial syllable  
Table 22 Personal and possessive pronouns  
Table 23 Contractions in Tai Mau (Young 1985) represented in Shan-Ni  
Table 24 Interrogatives (and indefinite pronouns)  
Table 25 Shan-Ni reduplication matching examples of affective reduplication in Young (1985: 19)  
Table 26 Adjectives with affective reduplication in Shan-Ni fitting Young's typology  
Table 27 Numeral classifiers in Shan-Ni  
Table 28 Question words and indefinite pronouns  
Table 29 Shan-Ni grammatical markers and their correspondences in other languages  
Table 30 TAM markers in Shan-Ni and Tai Aiton/ Phake, origin and usages  
Table 31 TAM markers and their functions in relation to verbs  
Table 32 Consonant sound changes in Shan-Ni and other Tai-speaking groups  
Table 33 Consonant clusters in Shan-Ni and their cognates in other Tai languages  
Table 34 Tone splits in Shan-Ni and other Tai languages  
Table 35 Interrogatives in Tai languages  
Table 36 Demonstratives in Shan-Ni and other Tai languages  
Table 37 The presence of Shan-Ni TAM markers in other languages  
Table 38 Features in Shan-Ni and other Tai languages
### Abbreviations

#### Glosses

<table>
<thead>
<tr>
<th>gloss</th>
<th>meaning</th>
<th>gloss</th>
<th>meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>first person</td>
<td>N</td>
<td>noun</td>
</tr>
<tr>
<td>2</td>
<td>second person</td>
<td>NEG</td>
<td>negation</td>
</tr>
<tr>
<td>3</td>
<td>third person</td>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>BM</td>
<td>burmese</td>
<td>POL</td>
<td>polite</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
<td>PROH</td>
<td>prohibitive</td>
</tr>
<tr>
<td>CLF</td>
<td>classifier</td>
<td>PRT</td>
<td>particle</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
<td>QP</td>
<td>polar question</td>
</tr>
<tr>
<td>DEF</td>
<td>definite</td>
<td>RDP</td>
<td>reduplication</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphasis</td>
<td>Rel</td>
<td>relative clause</td>
</tr>
<tr>
<td>EXIST</td>
<td>existential</td>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
<td>TAG</td>
<td>tag question</td>
</tr>
<tr>
<td>H</td>
<td>headnoun</td>
<td>V</td>
<td>Verb</td>
</tr>
<tr>
<td>LNK</td>
<td>linker</td>
<td>VPL</td>
<td>verbal plural</td>
</tr>
</tbody>
</table>

### Grammaticalizations

*The capitalized word represents the original word the grammaticalization derived from*

<table>
<thead>
<tr>
<th>gloss</th>
<th>form</th>
<th>function</th>
</tr>
</thead>
<tbody>
<tr>
<td>COME</td>
<td>maa⁴</td>
<td>TAM marker</td>
</tr>
<tr>
<td>FINISHED</td>
<td>yaw⁵</td>
<td>TAM marker</td>
</tr>
<tr>
<td>GIVE</td>
<td>hau⁴</td>
<td>causative</td>
</tr>
<tr>
<td>GO</td>
<td>kaa⁵</td>
<td>TAM marker</td>
</tr>
<tr>
<td>KEEP</td>
<td>waj⁵</td>
<td>TAM marker</td>
</tr>
<tr>
<td>MEET</td>
<td>nyaa⁴</td>
<td>&quot;at&quot;</td>
</tr>
<tr>
<td>OLD</td>
<td>kaw¹</td>
<td>TAM marker</td>
</tr>
<tr>
<td>SHOULD</td>
<td>kɤn⁴</td>
<td>TAM marker</td>
</tr>
<tr>
<td>YET</td>
<td>kɤn⁴</td>
<td>TAM marker</td>
</tr>
<tr>
<td>STAY</td>
<td>u¹</td>
<td>TAM marker</td>
</tr>
</tbody>
</table>
Preface

Shan-Ni is one of the most western languages of the Tai-Kadai language family. Whereas larger and more well-known languages like Thai, Lao and Shan Gyi are spoken in areas adjacent to each other where the speakers themselves are the majority, the Shan-Ni live in an area surrounded by speakers of Tibeto-Burman languages. Speakers of Shan-Ni have been in long-term contact with speakers of Lolo-Burmese (e.g. Lisu, Mru, Burmese), Nungish (e.g. Rawang, Anong) and Kachinic (e.g. Jinghpaw, Kadu) languages in the east and north, Naga languages in the west, and Burmese in the south. To a lesser extent, the Shan-Ni have been in contact with the Austroasiatic languages (Palaung, Wa). It is no wonder then, that Shan-Ni differs in structure from other Tai languages, and has features that might be common among its neighbours, but not among its relatives.

Lexically, Shan-Ni is conservative, with most words being of Tai origin. For these words, consonants and vowels often match, while the tones differ. But in Shan-Ni many of these words have become disyllabic, or acquired grammatical functions. This makes the innovations Shan-Ni has gone through clearly observable. I decided to focus on these innovations, because they are fundamental in understanding how the language works. I also believe these are of scientific importance. A good description of these features provides clues on language contact dynamics in the region, but can also inform literature on language contact in general.

Focussing on these innovations requires space. But despite the importance given to innovations and language contact in this thesis, it remains a descriptive work. There is no prior descriptive source, nor any source in English on Shan-Ni. Therefore, my aim remained to make this thesis as complete as possible, while still shining the light on what makes Shan-Ni special. I found a balance in this by writing some of the more descriptive pieces in a way that illustrates either the implications or the background behind which innovations have taken place.

Like other linguistic minorities in Myanmar, the Shan-Ni have only recently been able to openly organize language teaching and promotion in their own communities. Due to pressure from both the Burmese and Kachin armies, their freedom of cultural and linguistic expression has been limited for decades. Therefore, they have been absent from most linguistic and anthropological studies of the area, or are claimed to all have assimilated and shifted to Burmese (e.g. Takatani 2007). Since in 2013 and 2014 amendments in the law made it possible for schools to host extracurricular language classes, the Shan-Ni have rapidly developed materials and training to reach schools across the region. The sudden freedom to study the language means that this research comes at a time that the community itself is very active and motivated to work with the language, finding ways in which the language can be integrated into education, culture and society again. Therefore the scientific and community interest align, and I hope that this study can be complementary to the existing efforts of people I respect very much.
This thesis focuses on the variety of Shan-Ni spoken at the Indawgyi lake. How representative this is for other regions is not clear to me yet. Speakers consider Indawgyi Shan-Ni to be the most “central” variety, intelligible for speakers from both the west (Chindwin Valley) and the east (Irrawaddy valley).

Chapter I in this thesis describes the sociolinguistic and historical background of the Shan-Ni, the genetic affiliation, and the methods used in this research. Chapter 2 is about the phonology and orthography, and the transliteration used in this thesis. The examples used will have both the local orthography in the unique Shan-Ni script, and the transliteration in Roman script. Chapter 3 is about the shape of words. It shows how monosyllabic words have become disyllabic, by adding class terms in front of the earlier noun, which then lost their transparency making them unrecognizable as independent words. It also shows how Shan-Ni uses contractions also existing in Tai Mao for interrogatives and determiners, but added another process on top of it. It also shows how reduplication processes from Tai Mao are present in a fossilized form in Shan-Ni, and are no longer productive. Chapter 4 is about the structure of sentences. Shan-Ni distinguishes between referents that are new to the listener, referents that are known but inactive, and referents that already clear to the listener, in which case they will omit the whole argument. Because the omission of arguments is very common, and Shan-Ni has many grammatical markers that affect the structure of the sentence, the basic SVO order is hardly applicable in practice. While chapter 4 deals with the most common sentence structures, chapter 5 discusses some of the individual grammatical markers and the type of sentences in which they occur. The eight tense, aspect and mood markers and a few of their combinations are described in chapter 6. Finally, chapter 7 compares these features to the other Tai languages in the region, pointing out how the features described relate to their neighbours and which features are clear exceptions and should be traced back to another origin.
Chapter 1: The Shan-Ni and their language

1.1 Sociolinguistic situation

Shan-Ni is a Tai-Kadai language spoken in Kachin state and Sagaing region in Northern Myanmar. The estimate in the Ethnologue (Eberhard et al. 2019) of 100,000 speakers is very conservative. Some speakers estimate there to be as much as two million speakers, though most agree that this would be unlikely. The red circles in Figure 1 indicate where the Shan-Ni live. They are predominantly agriculturalist and live in valleys and plains; the easternmost group lives in the valley of the Irrawaddy river in Kachin state, while the western group lives in the valleys of the Chindwin and Uru rivers. The central group lives around the Indawgyi lake and the flat stretch of land around the railroad going from Katha to Myitkyina. For this research, most data was collected in the Indawgyi region, which is indicated by the blue square.

The Shan-Ni call themselves Tai Naing or Tai Leng, but because of the allophones /n/ - /l/ and /ai/ - /ɛ/ many prefer the Burmese name Shan-Ni, avoiding the confusion or region-specific connotations. The use of the /n/ is associated with the western region and the /l/ with the eastern region. In the Indawgyi region people use both variants interchangeably or simply refer to the language as khwaam⁴ taaj⁴ 'Tai language'. The names Tai and Shan refer
to their affiliation to the larger Tai group, and Naing/Leng/Ni means ‘red’. According to Hla Maw Maw (2017: 22), the name ‘red’ refers to their original settlement at the Red River (China, near frontier with Vietnam) and their traditional costumes with red waistbands. The name Tai Lieng is sometimes used as well, with lieng meaning ‘red’ in Tai Long or Shan Gyi, the largest Tai language spoken in Myanmar. The newspaper, youth organization and many speakers themselves use the name Shan-Ni to refer to the group as a whole, and following them, so do I in this thesis.

The language vitality of Shan-Ni differs from place to place. In Kachin state, Shan-Ni is spoken among all generations in rural places, and to a lesser extent in the cities. This is mainly due to restrictive language policies during the military regime from 1962 onwards. Due to the conflict in Kachin state, the rigor of these policies was the strongest between 1962 and 1994. During the ceasefire in the 1990s, the Shan-Ni started to recover old manuscripts, create literacy books, and secretly teach the language in summer schools. These efforts were led by Daw Khin Pyone Yee, who in 2011 became Kachin state minister of Shan affairs. When ethnic minority language classes became legal and supported by the state in 2013, she institutionalized her material and trainings with financial support from Unicef and other donors. Currently, about 100 schools in Kachin state have Shan-Ni literacy books and teachers trained by her team, providing classes for kindergarten, grade one and two (Daw Khin Pyone Yee, pers. comm.). In some places, this provides support for children who start school as Shan-Ni monolinguals, elsewhere it provides support for bilingual children who would otherwise only study Burmese. In urban areas it is more a language revitalization program. Classes are taught by twenty-year-old teachers who study and teach the language of their grandparents, even though many parents are not fluent. Although the language is less vital in these areas, the determination of teachers and families, and the comparatively better access to resources and information provides hope for the language to have a bright future.

The return of Shan-Ni into public life is very much driven by the youth. Around the same time language lessons became legal, people began using phones, internet and electricity. Roads improved and motor bikes became popular contributing to the mobility and connectivity of Shan-Ni youth. There have been several updates to the Shan-Ni fonts to type on the computer, and mobile phone apps are being developed. The Shan-Ni organize festivals, dance competitions and pageants. Many speakers have not learned to read to script yet, but the Shan-Ni newspaper publishes three pages every edition in Shan-Ni, written by Naan Nwe Ni Htun and Sai Kyaw Lin. These two writers also published another book, that I use in this study as a source. It is called ‘Basic Spoken Shan-Ni’ (2017) and includes conversations, stories, and lists of specific terminology (kinship, body parts, agricultural tools) in Shan-Ni, Burmese and Shan-Ni transliterated in Burmese script. Within one year 5000 books were

---

1Hla Maw Maw (2017) does use the name Taileng in her anthropological thesis about the cultural heritage of the group, because her study has a strong in-depth focus on the eastern group, who exclusively use this name.
sold\(^2\). For many adults who have never had literacy lessons, the Shan-Ni transliterated in Burmese and the conversational style make this book a true treasure. It allows them to practice reading and writing by themselves, and show their children elements of the language.

Traditionally, the Shan-Ni have always cultivated the lowlands, while trading with the other ethnic groups who lived in the highlands. They have been in long-term contact with Kachin\(^3\) groups, Kadu, Naga and Tai Khamti that live in the region. The shared history with other Tai groups and the Burmese will be discussed under the sections ‘History’ and ‘Genetic affiliation’. Due to the conflict, in the 1960s and 1970s Kachin communities who used to live in the mountains near the Shan-Ni were forced to seek refuge in the lowlands where Shan-Ni communities lived. They built their own settlements next to the Shan-Ni villages, though they largely lead separate lives.

This research took place in the Indawgyi region, a wildlife sanctuary home to Myanmar’s largest lake, the Indawgyi lake. It is an important habitat for hundreds of species of plants, insects, reptiles, and a resting place for species of migrating birds and fish. Burmese visitors tend to come for the *Shwe Myitzu* (ရေမြပုသောဘုရား), a floating pagoda which during the wet season can only be reached by boat (Figure 2). The Shan-Ni villages are built at the edges of the lake. The map below (Figure 3) shows the main villages: Nam Mun, Lone Ton and Nyaung

\(^2\)Since in rural Myanmar it is very normal to copy books, the reach is probably a lot bigger.

\(^3\)Including Zaiwa, Lisu, Rawang, Maru, Lashi and Jinghpaw.

**Figure 2** Shwe Myitzu pagoda (photo by Victoria Milko)
Bin. Lone Ton is the administrative head village, and is home to a hospital, guesthouses for visitors and a military base. Nam Mun and Nyaung Bin are the biggest villages of the area, and function as trade hubs connecting Indawgyi to the cities Hopin (Nam Mun) and Hpakan (Nyaung Bin). Most people are agriculturists or anglers, though gold, amber and jade mines also provide opportunities for both labour and trade.

The language vitality and dominance of Shan-Ni in the Indawgyi region differs on each side of the lake. This has to do with the history and current opportunities. At the north side of the lake, Shan-Ni is the dominant language. The youth is using more Burmese nowadays, but

Figure 3 Map of Indawgyi with major villages and nearby cities (adapted from Google Maps).
many adults are more comfortable with Shan-Ni than they are with Burmese. On the westside, bilingualism is common among all generations. It is a busy road for trade and tourism, hosting important cultural and religious sites. Burmese provides economic opportunities, but Shan-Ni is maintained as the language they use among themselves. People do codeswitch often, but Shan-Ni remains the preferred language. Children and new residents in these villages learn Shan-Ni rather than getting by with Burmese. In the south people use more Burmese, and almost all Shan-Ni speakers are adults over thirty years old. This might be related to the presence of the military base and thus stricter control in the area in the past. The east side used to be a dense forest and a swamp, with only few settlements. In the 1990s, Burmese migrants from central Myanmar came to work in the logging and fishing industries, vastly outnumbering the pre-existing population. This changed the eastern side from a relatively inhabitable area to an attractive area to live with a lot of arable land. Members of earlier Shan-Ni communities married the Burmese migrants. There are still a few older speakers, but often being part of Burmese families, they rarely use Shan-Ni.

Figure 4 shows all the village names in English and Shan-Ni, with the Shan-Ni script. The names have been burmanized, but the original Shan-Ni names are still identifiable. The drawings correspond with the meaning of the village names, and have significance for the history of the lake. They are descriptive of flora and fauna of villages, functions or oral history. A selection of them are listed in Table 1. The villages nyɔng²paang¹ ‘Nyaung Bin’ long⁴kɔng⁵ ‘Lone Kauk’ refer to types of trees, and long⁴tong² ‘Lone Ton’ long⁴caang⁴ ‘Lonsant’ and nam⁵mi³lɔng² ‘Nam Mi Laung’ to places where humans, bears and elephants come down from the mountains and gather. lɔj⁴mon² ‘Loi Mon’ tells the origin story of the lake, which ends in a widow looking back from the mountain after the area had flooded and the lake was created.

<table>
<thead>
<tr>
<th>Name</th>
<th>Shan-Ni script</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>nɔng³long³</td>
<td>မြို့ ကြောင်ကြောင်း</td>
<td>big lake (bm: indawgyi)</td>
</tr>
<tr>
<td>phraa¹nɔng³long³ci¹kham⁴</td>
<td>ကြောင်ကြောင်း မြို့ ကြောင်ကြောင်း</td>
<td>golden pagoda on the big lake</td>
</tr>
<tr>
<td>mɔng⁴ nɔng³</td>
<td>မိုင် ကြောင်ကြောင်း</td>
<td>lake city</td>
</tr>
<tr>
<td>maak¹mong⁴kaj¹</td>
<td>ပျော်ကြောင်ကြောင်း</td>
<td>chicken mango (small mangos)</td>
</tr>
<tr>
<td>long⁴tong²</td>
<td>ကြောင်ကြောင်း</td>
<td>gathering place</td>
</tr>
<tr>
<td>long⁴kɔng⁵</td>
<td>ကြောင်ကြောင်း</td>
<td>poisonous tree</td>
</tr>
<tr>
<td>lɔj⁴mon²</td>
<td>လားခေါင်</td>
<td>look back from the mountain</td>
</tr>
<tr>
<td>nam⁵mi³lɔng²</td>
<td>မြို့ ကြောင်ကြောင်း</td>
<td>water (place) where bears descend</td>
</tr>
<tr>
<td>nyɔng²paang¹</td>
<td>မြို့ ကြောင်ကြောင်း</td>
<td>banyan tree</td>
</tr>
<tr>
<td>lɔj⁴kham⁴</td>
<td>လားခေါင်</td>
<td>gold mountain (bm: shwe taung)</td>
</tr>
<tr>
<td>tong¹caang⁵kham⁴</td>
<td>ကြောင်ကြောင်း ကြောင်ကြောင်း</td>
<td>gathering place of golden elephants</td>
</tr>
<tr>
<td>long⁴caang⁵</td>
<td>ကြောင်ကြောင်း</td>
<td>elephants descend</td>
</tr>
<tr>
<td>nam⁵mo¹kam¹</td>
<td>မြို့ ကြောင်ကြောင်း</td>
<td>brown water well</td>
</tr>
<tr>
<td>haw³paa⁴</td>
<td>မြို့ ကြောင်ကြောင်း</td>
<td>fish louse</td>
</tr>
</tbody>
</table>
Figure 4 Map with symbols of Indawgyi villages and their names (Face of Indawgyi, CM)

Illustrations by Dohee Kwon, names were collected in Ywatit and Maing Naung by Stephen Traina-Dorge and CM and later doublechecked in Pan La by Sayadaw Pannya Vamsa.
1.2 History

The Shan-Ni consider their history and heritage to be that of the four city states Mong Yang 'heron city' (Mohnyin), Mong Kaung 'drum city' (Mogaung), Wun Seu 'tiger state' (Wuntho) and Mong Myit 'dagger city' (Momeik). Nowadays, these territories are divided between different administrative territories. Mohnyin and Mogaung are located in Kachin state, Wuntho in Sagaing Region, and Momeik in Shan State.

Shan-Ni are believed to have been in this region for at least 2000 years (Hla Maw Maw 2017: 74). Their first important settlement, Maing Mao, no longer exists, but ancient ruins found in 2009 south of Kachin state’s capital Myitkyina are thought of as being the remnants of Maing Mao. The first significant empire it was part of, was the Nan-Chao empire. During the Nan-Chao period, around 600-900 AD, Chinese sources refer to Maing Mao as Lishiu 'beautiful water'. The Nan-Chao empire (see Figure 5) existed out of a confederation of several Tai states centred around the old capital of Dali in Southwestern China. Mong Mao was the westernmost state, connecting Nan-Chao to India. It was the place where Nan-Chao got a large part of its gold and salt (Luce 1961: 69, 71). Nan-Chao would also send their elite criminal offenders to wash gold in Lishiu instead of giving them the dead penalty (Luce: 1961 71). Figure 5 shows the capital ‘Lishiu Cheng’ near modern-day Myitkyina, the Ta Erh Kingdom around the Indawgyi lake, and ‘An Hsi Cheng’ where Mong Kaung is now.

According to the Shan-Ni chronicles, three sons of the Maing Mao king left their father's home to establish Mong Kaung (722 CE), Mong Yang (723 CE) and Mong Bann (724)\(^5\). The fourth son stayed home to rule the empire. When the ruler of Mong Kaung died, his brother Sao Sam\(^6\) from Mong Bann, took over Mong Kaung. From there, he conquered many other areas, including the Hukhawn region, Katha, Assam and Manipur (Hla Maw Maw 2017: 76). After the tenth century the Nanchao empire gradually became less powerful, until it finally fell under the Mongol invasion led by Kublai Khan in 1253 (Sai Aung Tun 2009: 10). In the meantime, a tributary to Nan-Chao, Mong Mao, had grown more prosperous. Tai elites from Nan-Chao moved to Mong Mao, becoming a new center of power in the region (Sai Aung Tun 2009: 14). The Mong Mao Empire or ‘Mong Mao Long’, was ruled from the Shweli river valley, with its capital near the current location of the Chinese border city Ruili. While the king himself stayed in his own capital, his brothers went out and established themselves in Mong Kaung (1215) and Assam (1218). This was the start of the Ahom Empire in northeastern India. The brother who went to Mong Kaung, Sao Sam Lung Hpa, made Mong Kaung the new military capital of Mong Mao. Mong Kaung became powerful again, ruling over at least 99 mongs or (city)states (Sai Aung Tun 2009: 18).

\(^5\)Current location of the Indawgyi lake and the Hpakant jade mines.

\(^6\)Sao ‘leader’ and sam ‘third son’ refers to Sao Sam being a king and the third son in his family, hence there are many Shan leaders with the same name.
Mong Kaung preexisted the Sao Sam Lung Hpa and Mong Mao Lung, but it is very likely that

Figure 5 Map of the Nan-Chao empire in the 8th & 9th century (copied from Luce 1961)
Sao Sam Lung Hpa brought a lot of people with him into the area. According to Sai Aung Tun, the people that came with Sao Sam Lung Hpa were the Shan-Ni (Sai Aung Tun 2001: 5). They mixed with the people who were already living there, though they outnumbered them greatly (Sai Kam Mong 2004: 15). Large-scale immigration into the region by a welcome group is also discussed in this parabaik (palm leaf) inscription from 1179 AD (Figure 6), in which is written;

Notification to Mayor,
Maing Mao Minister.
Through Royal's order the villages of Maing Mao town must be united and properly ruled and unfair ruling is strictly prohibited. There may be no disputes with neighbouring towns and villages. Peace must be ensured. Maing Mao city must be divided into 4 regions and 8 villages through royal order. Some villages such as Wakhaung, Oolaik, Moat Loi, Kazu and Kat Cho/Khat Cho in addition to Maing Mao Township are to be united through rule of law. The increased population must be properly provided for. This order was received on the 3rd waning moon day of Tagu, 541 ME.

(Palm-leaf manuscript 1179 AD, adapted from Hla Maw Maw 2017: 85)

Figure 6 Parabaik manuscript 1179 AD (Hla Maw Maw 2017:85)

The manuscript demonstrates that (1) there was an established administration, (2) there was an increased population, causing tension and requiring one administrative unit to be split up into multiple smaller units and (3) the rulers were actively supporting the newcomers, by ordering the established community to provide for them.

In 1526, a group of armies from different Shan states led by a Shan-Ni prince from Mong

---

7 Although the years do not exactly add up, it must be taken into account that these are subject to different calendars. The Tai, Burmese, Chinese and western calendars are all different, so confusion easily arises. The point is here that different sources discuss a large migration into the area around 1200, which influences the linguistic history of the Shan-Ni.

8 The Shan states were independent Shan princedoms. According to their own interest, they would unite or fight each other and neighbouring groups. Most of them were located in what is now Shan State in eastern Myanmar. The places where Shan-Ni live were part of the ‘northern Shan states’.
Yang sacked the Burmese court of Ava (Fernquist 2005). The occupation did not last for very long. When the Burmese regained strength under king Bayinnaung, he took revenge and occupied the Shan states that had attacked Ava before. By 1557 all Shan-Ni territories were in Bayinnaung’s hands. This merely meant local kings had to pay tribute to the Burmese crown. This changed in 1768-1769, when Mong Kaung and other Shan allies supported the Chinese in a battle against the Burmese and lost (Sai Kam Mong 2004: 33). From that moment onwards, the Burmese court would hold a tighter grip on the northern Shan states, sending their own kings and imposing harsh military rule (Sai Kam Mong 2004: 44). The then important city of Waimaw (south of Myitkyina) was destroyed by the Burmese in 1810. Gradually the Kachin also started to become more powerful, capturing Mong Kaung and devastating the Indawgyi valley in 1883 (Hunter 1908b: 137).

The British captured Mandalay and announced the annexation of Upper Burma in 1885. The ruler of Wuntho, Sao Aung Myat, refused to accept, and continued to fight the British until 1891. The final crackdown by the British was incredibly violent, leading Sao Aung Myat to flee to China and the other survivors to migrate to other parts of Myanmar (Hla Maw Maw 2017: 106, Hunter 1908a: 155). Wuntho, once was a powerful Shan state, features in the 1908 Imperial Gazetteer of India as a township under Katha district, with only villages (no cities) and a mention of the population being “almost exclusively Burmese” (Hunter 1908a: 156).

Having been defeated by the Burmese, Kachin and British, the Shan-Ni who stayed in the area became minorities in their own land. Burmese, Kachin, Indian, Nepali and other people came into the area to build the railroad, connecting the urban areas of Katha, Mohnyin, Hopin, Mogaung and Myitkyina to Mandalay. Cochrane (1915:24) reports that in this period the Shan had assimilated to Burmese customs and speech outside the house, but continued to maintain Shan customs and speech inside the house. Several people I spoke to in Hopin and Mohnyin, told me that this continued until the 1960s, when people stopped speaking Shan inside the house, out of fear to be overheard by Burmese or Kachin soldiers.

After the Second World War the Union of Burma was created, which was supposed to be a federal state in which ethnic states would have a large amount of autonomy over their own territories. The Shan-Ni agreed to support the inclusion of their land into Kachin state, which itself was promised possible independence in ten years time. These agreements were made under the Burmese general Aung San. He and his cabinet were assassinated before they could assume office. His successors decided to install more centralist policies, infuriating ethnic leaders who agreed to be part of the Union under the promise of federalism. This worsened throughout the 1950s, with a larger emphasis on Buddhism and the Burmese language, completely banning the teaching of local languages in 1962. All over the country, ethnic groups started to form their own armies, resisting the Burmese state. During British times, the majority of the army existed out of Kachin soldiers, but during the 1950s these were gradually replaced with Burmese soldiers. In 1962 the Kachin Independence Army (KIA) was formed, and the Shan-Ni got caught in a conflict between the Burmese and Kachin armies (Hla Maw Maw 2017: 105). Both the Kachin Independence Organization (KIO) and the Burmese
government claimed to be the legitimate government of Kachin state, and used violence, taxes and forced recruitment accordingly. Eventually, the Burmese set up their military bases in the lowlands where the Shan-Ni lived, while the fighting continued in the surrounding hills. The Kachin villages in the hills were under constant attack in an effort to cut the KIA off from access to food, information, recruits, and other resources. As a result, many Kachin were displaced and moved into Shan-Ni villages. Around the Indawgyi lake, there are large Kachin communities in Nam Mun, Lonton and Maing Naung. During the ceasefire from 1994-2011 there was increased freedom of movement and expression. When the ceasefire ended in 2011, Myanmar was already going through a democratic transition, giving military personal less control over people’s daily life.

1.3 Genetic affiliation

Shan-Ni is part of the southwestern branch of the Tai-Kadai language family. Figure 7 shows the Tai-Kadai family tree according to Diller et al. (2008: 7). Shan-Ni is not present in his overview, but is discussed later in the same book by Edmondson as part of the southwestern branch (Edmondson 2008: 192). It shares a large number of cognates and grammatical features with other languages in the southwestern branch, including some well-known languages such as Thai and Lao, as well as smaller Tai languages present in geographical proximity of Shan-Ni (Khamti, Phake, Aiton). The cognates Shan-Ni shares with these languages oftentimes only differ in tone but have the same consonants and vowels, and have sometimes changed their meaning or have gained different grammatical functions. The biggest Tai language of Myanmar is in this overview referred to as ‘Shan’. The word Shan represents all Shan groups in Myanmar, but the language meant here is the variety spoken by the dominant Tai group living in Shan State of Myanmar, which is otherwise known as Tai Yai, Tai Long or Shan Gyi (big/major Shan in Burmese). I will refer to this language throughout this thesis as Shan Gyi. Within the southwestern branch, Shan Gyi is less related to Shan-Ni than the Tai languages spoken in the Northern Myanmar – Northeastern India – Southwestern China border region, but due to its political and cultural prominence, it does have influence on Shan-Ni.

Edmondson (2008: 203) distinguishes three groups of Shan: Southern Shan, Northern Shan and Khamti. The northern tier includes the languages of groups that had settled in Northern Myanmar before the 6th century AD (Edmondson 2008: 184). He shortly discusses the settlement history and phonology of the different groups, and places Shan-Ni in the northern group. The Shan-Ni in Kachin state themselves are familiar with four other ‘Tai’ groups: Tai Khamti, Tai Nua (Lue), Tai Long (Shan Gyi) and Tai Sa. Of these, they understand most of Tai Khamti, followed by Tai Nua. They consider the Khamti their closest relatives. Depending on the person, Shan-Ni and Khamti might be considered mutually intelligible, since they understand most of each other’s words (including tones) but the word order and grammar differs. Some Shan-Ni see the Khamti as a representation of their own past, since the Khamti
live in more isolated areas and have preserved traditions which the Shan-Ni only know about from the stories of their grandparents. People who have had more exposure to Shan Gyi understand that as well, though people experience the language as quite distant. The Shan-Ni do not understand Tai Sa, but that is probably because it is not a Tai language but a variety of the Tibeto-Burman Ngochang. However, the Tai Sa are seen as Tai because they are culturally Tai (Edmondson 2008: 193).

Looking at the different Tai languages in the broader region, Shan-Ni shares a lot more similarities to Tai Phake than to any other language. The Tai Phake migrated from Mong Kaung to Assam around 1775 (Morey 2005a: 19), which means they were at the center of the Shan-Ni court, and their ancestors probably spoke the exact same language as the ancestors of the Shan-Ni 250 years ago. The Khamti on the other hand, have lived separately from the Shan-Ni from at least the Nan-Chao period onwards (±1400 years). They are included in the Nan-Chao map (Figure 5) as “little Brahmins of the north” referring to their affiliation with Tibet at the time (Sai Kam Mong 2004: 15). Throughout history, they have paid tribute to the

![Figure 7](image-url)  
**Figure 7** The Tai-Kadai language family (copied from Diller et al. 2008: 7)
rulers of Mong Kaung, but their territory was too isolated for them to maintain intensive contact with the Shan-Ni.

In the history described above, a large group of other Tai people moved to the Mong Kaung/Myitkyina together with Sao Sam Lung Hpa around 1215 CE. These people most likely came from the Shweli river valley. Around the same time, his brother and a large group of followers started the Ahom Empire in Assam. This area is currently the border of Ruili (China) and Namkham (Myanmar). The languages spoken there are Tai Nua and Tai Mau. Tai Mau is restricted to a small area around Namkham. There is a short description of Tai Mau by Young (1985), whose book also includes various narratives and an extensive Tai Mau wordlists. Tai Lue or Tai Nua is spoken in a large area in China, Lao, Thailand and Myanmar, and has a lot of internal variety.

Chantanaroj (2007) compares the phonology of different dialects of Tai Nua, and includes Tai Mau, Tai Long and Shan-Ni as well. In her study, Shan-Ni is included under the name Tai Lai, and includes data from six different locations. Her study showed that the western dialects of Tai Nua were closer to Shan-Ni than they are to the eastern dialects of Tai Nua, with the data collected in Ruili, Zhefang and Namkham (Tai Mau) in the Shweli river valley showing the most similarity to Shan-Ni (Chantanaroj 2007: 78). This is the area where Sao Sam Lung Hpa came from around 1200, so similarities between Tai Nua spoken in that area and Shan-Ni support the theory that a large group of people migrated with Sao Sam Lung Hpa into the area. Edmondson (2008: 203) who also compared the phonology different varieties of Tai languages in the borderlands, found innovations in the Shweli river valley that he finds more likely to be the result of considerable in-migration from the east rather than innovation within the sedentary population.

Figure 8 shows the probable migration history of the different groups related to the Shan-Ni. I have added geographical references to show which groups are currently living in the same location. Based on this we can expect indeed Shan-Ni shares most features with Tai Phake, followed by Tai Aiton. A large group from Ruili/Namkham joined the pre-existing Tai group in Mong Kaung in around 1215, while at the same time a group moved to Assam (Ahom). The people that stayed behind, nowadays speak Tai Mau and Tai Lue. There may thus be features all these group share with Shan-Ni, having a shared history until 1200. The pre-existing population was related to the Tai Khamti, and maintained contact throughout the centuries, though they lived in different places. Currently, there are Khamti in the Chindwin region living alongside Shan-Ni, and Khamti in Assam alongside the Phake and Aiton. The Shan-Ni in the Indawgyi and Irrawaddy region do not live alongside a large population of another Tai group. All different Tai groups are in contact with different Tibeto-Burman speaking groups in the region, but due to the differences in agricultural and lifestyle practices, this contact is not necessarily very intense.

In terms of descriptive work on languages related to Shan-Ni, the most relevant is the work of Morey (2005a, 2005b, 2008). His description includes the Tai languages spoken in Assam, India (Phake, Aiton, Khamyang and Khamti), focusing mostly on Aiton and Phake. Earlier
work on the varieties spoken in India by Diller (1992) is less extensive but provides useful insights as well. Khamti is included in many 19th and 20th century sources (e.g. Robinson 1849, Needham 1894, Grierson 1904). Descriptive work specifically focusing on Khamti includes a phonology and wordlist by Weidert (1977) and a PhD thesis by Inglis (2014). Rikker Dockum has also been working on Khamti in recent years and has given various conference presentations relevant to Shan-Ni, though, to my knowledge, he has not published on Khamti yet. For Tai Mau I will refer to Young (1985). Lao (Enfield 2007) is less closely related to Shan-Ni, but Enfield’s work includes many theoretical considerations that are relevant for Shan-Ni as well. When comparing words from different languages, I have retrieved data from SEALang (a.o. for Lao, Shan Gyi, Ahom and Phake). The SEALang is a database of digitized dictionaries and wordlists of several Southeast Asian languages. For each entry SEALang includes the source, which is often written by the same authors I cite elsewhere (e.g. Enfield for Lao, Morey for Phake and Ahom), controlling the variation between lexical items from the database and from other sources.

1.4 Aims of this thesis

The absence of research on Shan-Ni means that there is a large gap in understanding the developments within the Tai language family, as well as the language contact situation in the India-Myanmar-China border region. With at least 100,000 speakers spread over a large area and a history of ruling over an even larger region, it is a missing puzzle piece in many ways. Understanding more about Shan-Ni thus goes beyond Shan-Ni itself. But before any conclusions can be drawn regarding Shan-Ni’s place in the language family and contact with
other languages, Shan-Ni needs a solid description first. This thesis aims to lay a foundation for this in two ways. First by describing the main features of the language, and second, by connecting these findings to other Tai and non-Tai languages in region.

Shan-Ni is conservative in the sense that it most words do have a clear Tai origin, but innovative in way these words are used. This makes Shan-Ni unintelligible for most other Tai speakers in northern Myanmar. There are two main ways in which Shan-Ni stands out. (1) the use of disyllabic words where other Tai languages have monosyllabic words (2) the presence of grammatical markers allowing constructions not present or common in other Tai languages. This creates a unique situation in which the innovations that distinguish Shan-Ni from other Tai languages, are very overt and are fundamental in understanding how the language is structured. The presence of an extra syllable that lacks for the same word in all other Tai languages, or tense marking unique to Shan-Ni are strange phenomena that deserve their own chapters.

The structure of this thesis is centred around these innovations. These require more attention than features that are more common in Tai languages. However, having no prior descriptive work on Shan-Ni, it is important to give enough descriptive information to understand the environment in which these innovations operate. The result is that the structure differs from what one might expect from a descriptive thesis.

After the phonology, chapter 3 describes the shape of words, centred around the question of how disyllabic words are structured. While studying these different processes, the structures of nouns, pronouns, demonstratives and interrogatives are described. The structure of sentences in Shan-Ni also differs from other Tai languages, because it is often influenced by the omission of arguments and the presence of grammatical markers. Chapter 4 thus discusses the sentence structure by describing the omission of arguments, and different types of sentences. Chapter 5 describes grammatical markers, which each have their own function in the language as well. Chapter 6 is dedicated to the TAM markers, perhaps the strangest phenomenon for a Tai language. Even in closely related languages in which these markers are present is, the system is less elaborate than in Shan-Ni. Chapter 7 discussed how these innovations are represented in related languages, and what the comparisons might teach us about the position of Shan-Ni within the Tai language family and the history of the language contact.

1.5 Methods

For this research I spent five months in Myanmar from April – August 2018. Most of that time I spent at the Indawgyi lake, and I also interviewed people in Myitkyina, Hopin, Mogaung, Mohnyin, and Mandalay and attended a teacher training in Mingon. I also followed an intensive Burmese course in Yangon for two weeks in the beginning of my fieldwork. The people I worked with were bilingual in Shan-Ni and Burmese, and did not speak English. In the first few weeks before my course an English-Burmese interpreter joined me, who
normally worked for Face of Indawgyi. Stephen and Patrick from Face of Indawgyi also introduced me to several people in the community, which helped me to get started with my research. Other people I met in the train, or cafés, or through other friends. After the Burmese course, I conducted my research independently in Burmese and Shan-Ni. I collected six types of data:

1. word lists (Swadesh)
2. narratives – traditional stories people choose to tell
3. narratives based on stimuli (Frog story, Hunting story, Pear story)
4. elicitation based on questionnaires (lists with sentences in Burmese)
5. elicitation (questions I asked)
6. The Basic Spoken Shan-Ni book (to be discussed)

I asked people to choose whether they would like to tell a traditional story, a personal story or a story based on stimuli. Most people preferred traditional stories, especially the women, who are very skilled and passionate storytellers. I planned to use the stimuli as a back-up, but since most people had (traditional) stories to tell, I ended up using the stimuli only on a few occasions. In the beginning I used the Frog story as a stimulus, because I was less familiar with both Shan-Ni and Burmese, and knowing what kind of words to expect made it easier to start with. Later I asked specifically for narratives based on stimuli, because the usage of certain grammatical features such as TAM markers is different for traditional stories (events in the past, citing conversations, etc.) or more dynamic stimuli (talking about what is going on the present). The Hunting story (created by Marine Vuillermet) is specifically designed to elicit motion events.

Before recording narratives, I always asked for verbal consent. I asked speakers if they agreed to being recorded, and I explained them that if at any point they did not want to continue or start again, or if at a later moment after the recording they would change their minds and preferred not to participate, that this would be no problem. I also asked them whether they would like their names to be mentioned or whether they would prefer to be anonymous, and all wanted their names to be included. I always let them listen to the recording afterwards, and asked them if they were happy with the recording, if they wanted me to send it to their phone (if possible), and if they wanted to continue to write out this version. One person got insecure before the recording and decided not to participate. Another person stopped halfway and wanted to start again. Another person was initially very enthusiastic about the recording, but a family member came in and criticized her language use, after which she did not want to continue. In those cases, I did not use the recordings they did not want me to use.

An issue that I encountered was that for some stories, there was not enough time to write out a full narrative. I tried to prevent this by asking people to tell me a story of around three minutes, which usually would end up to be at least five minutes. But some people would continue speaking for longer anyway, and I did not cut them off. I chose to accept this, because
I wanted people to tell stories they enjoyed telling in the way they themselves preferred. With a story of five minutes, I could start writing out the story directly after the recording, and make one more appointment to continue, and then finish. If it was longer, it could happen that it was either not possible to meet more times, or the person got bored of the old story and suggested to tell a better story. Because my Burmese as meta language was developing over time, writing out stories went slower and it was harder for me to control the situation. As a consequence, I have several stories of which I have only written out the first part.

The Burmese elicitation questionnaires have been designed by André Müller, who uses them to compare several languages in northern Myanmar. They contain sentences aimed to elicit different ways of expressing emphasis, causation, and new situations. The questionnaires allowed me to elicit more complex sentences that I would not have been able to formulate as clearly and accurately myself in Burmese. The disadvantage here is the influence of Burmese, which especially in long sentences may affect the word order and the use of grammatical markers. This made the ‘emphasis’ list most effective, because the sentences were shorter and included more things people say in daily life. I elicited the emphasis list in three villages (Nammilaung, Ywatit and Pan La), all lists in Nammilaung, and parts of the other lists in Ywatit and Pan La. Asking the same sentence in multiple villages helped me to confirm whether variation was local, free or dependent on the context. For example, ang¹ ‘for’ and caa⁴ ‘for’ are always used in the same sentences, but preference differs per village/person. There is no distinction in meaning between the different pronunciations of the linker si³/yi³/i³/ni³, but the presence of either of them is obligatory. Without having compared these lists, these findings would have been a lot harder to uncover.

An overview of all the people I worked with and the meaning of codes that are used with the examples in this thesis can be found in Table 2. In the examples, numbers refer to the sentence in the story, except for the elicitation and fieldnote items (DS_LT and MK_LT), where numbers refer to the month it was written down.

The Basic Spoken Shanni book (2017) was a very valuable source I used to fill gaps in my own data when I had already returned to the Netherlands. Especially when I had few instances of a certain phenomenon in my own data of which I did not fully understand what the function was, I could search for it in the Basic Spoken Shan-Ni book. This makes this thesis consirably more substantive than it would otherwise be. The book is written by Naan Nwe Ni Htun and Sai Kyaw Lin. Both write Shan-Ni language content for the newspaper ‘The Voice of Shan-Ni’. I recorded a few stories with Naan Nwe, and I have had many discussions about the language with her as well. She is from Ailithu, a small village north of Indawgyi lake. She grew up speaking Shan-Ni, and learned Burmese later in school. The book contains 130 pages of dialogues, narratives and wordlists from specific domains. It is called ‘Basic Spoken Shan-Ni’ because it contains examples of the spoken language people use in their daily lives. Whenever I use this book as a source, the reference is (BSS_page number). It is popular among adults and children alike, as it features a Burmese translation and transliteration,
making it accessible for adults who are fluent but have never learned how to read in Shan-Ni, and for children who are studying the language in school. To avoid misinterpretations because of the Burmese translation, I have cross-checked most Shan-Ni terms I used from the book with speakers in person, and when that was not possible I checked their Phake and Shan Gyi cognates in the online SEAlang database, to come as close to the original meaning as possible.

The examples of other data I collected follow the orthography of the Basic Spoken Shanni book as much as possible. This is possible due to the completeness of the book. I also used it to check the tones. When I could not find a word in the book, I tried to check the spelling and tone with Naan Nwe. The tones were very difficult for me, because I had never worked with tones before, I was not sure how many tones there were, and most people I worked with were not very aware of which tone they used in which word either. Eventually I learned how to use Gedney’s box (see section 2.3), figured out which tones there were, and learned to recognize them. Since the writers already had a system and mark tones in the script, I could easily check the book or with writers or teachers which tone was used, not to rely only on my own hearing.

For each section, I tried as much as possible to use a mix of examples from both spoken data and the book. This was not always possible, because even if grammatical particles were used by speakers, I could not verify this with them when I had already returned to the Netherlands.
Due to phonological reduction, grammatical particles may sound similar; it is not always clear which particle was meant by speakers. For example, the TAM markers $kaa^5$ and $k\eta n^4$ and the verbal plural particle $kan^4$ can all be pronounced as $/\varnothing n/$ in fast speech, especially when followed by a word starting with a nasal. Because these are quite complex grammatical functions, when I asked people about it during the interviews, they would often say it was just a sound.

There are no secondary sources in English on Shan-Ni itself, but there are sources on related languages I cite in my work. These include many cognates that have retained a very similar shape to Shan-Ni. Although the comparison is important, I have decided not to adapt the transliterations chosen by the original authors. This is because the Tai languages discussed have a large vowel and tone inventory, making the transliteration very precise work. Tai languages in the region have four, five or six tones, which may partially correspond with each other, but never match exactly. Vowels merge and split; sometimes there might be two different vowels in another language representing one vowel in Shan-Ni or the other way around. Hence, a tone represented in Shan-Ni as $<^5>$ corresponds with tone $<^4>$ in Phake (Morey 2002), though Shan-Ni tone $<^2>$ in Phake can be either $<^3>$ or $<^5>$. This will be discussed in more detail in Chapter 7, where in the first part the phonology of the different Tai languages will be compared.
Chapter 2: Phonology and Orthography

Shan-Ni is written in its own variety of lik tai or ‘tai script’. It shares most similarities with the scripts of Tai Khamti and the Tai languages of Assam in northeastern India\(^9\). The script is syllabic; vowels are attached to consonant graphemes. In the 1990s, the Shan-Ni added tones to their script. Many of the consonants correspond with the Burmese script, and those that do not correspond in form often do not have an exact Burmese equivalent. Some of these differences with Burmese are shared with Tai Yai or Shan Gyi, the main Shan language in Myanmar, though there are more similarities with the script of the Tai languages of Northeastern India. Since the script itself is the best fit for the phonology of Shan-Ni, I will include the script in my discussion of the phonology and throughout my thesis. The phonology and orthography of Shan-Ni are also discussed in a powerpoint by Wyn Owen (2013) based on wordlists from Myitkyina, northeast of where my data was collected. With a few exceptions, there are quite some similarities between his analysis and mine in the vowels and consonants, though it differs with regards to the tones. He also uses slightly different graphemes for the consonants. In some cases that can be attributed to the font used, with some graphemes belonging to Shan Gyi rather than Shan-Ni. The font I am using is the newest available edition of the same font used by the newspaper and the Basic Spoken Shan-Ni book, the 2018 version of the Shan-Ni font developed by the Mandalay Art House.

2.1 Consonants

The consonant inventory of Shan-Ni exists of 19 consonants. There is no distinction between voiced and voiceless consonants, though there is an alveolar voiced stop /d/ which is an allophone of the nasal /n/ and the lateral approximant /l/. Many words in Shan-Ni which start with the lateral /l/ in the eastern region, are pronounced with the nasal /n/ in the west. In the Indawgyi region where my data was collected, words of both varieties are accepted and used interchangeably, sometimes even within the same sentence. I thus treat these as

<table>
<thead>
<tr>
<th>Table 3 Shan-Ni consonant inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Plosive</td>
</tr>
<tr>
<td>Plosive (asp.)</td>
</tr>
<tr>
<td>Nasal</td>
</tr>
<tr>
<td>Tapp</td>
</tr>
<tr>
<td>Affricate</td>
</tr>
<tr>
<td>Fricative</td>
</tr>
<tr>
<td>Approximant</td>
</tr>
<tr>
<td>Lateral Aprx.</td>
</tr>
</tbody>
</table>

\(^9\) For more on the Tai scripts of Assam, see Morey 2005a (179-205). For the development and history of Shan scripts in general, see Sai Kam Mong (2004)
allophones of each other, though since they are of equal status so they both get a place in the consonant inventory. They both are occasionally pronounced as a voiced alveolar stop /d/ or an alveolar tap /ɾ/. The alveolar tap is marked in cursive, because as a phoneme by itself the /ɾ/ only occurs in Buddhist terminology, borrowed from Pali\textsuperscript{10}. The velar stop /k/ is often pronounced a voiced velar fricative /ɣ/ and the aspirated velar stop /kʰ/ is often pronounced as the voiceless velar fricative /x/. The affricate /ts/ is a sharp sound, contrasting with the alveolar fricative /s/ which may sometimes be slightly aspirated. The labial fricative /f/ and the aspirated labial stop /pʰ/ are treated as a single phoneme by some speakers, though there are enough minimal pairs to recognise the distinction, as shown in Table 4.

<table>
<thead>
<tr>
<th>Table 4 Minimal pairs distinguishing /pʰ/ and /f/</th>
</tr>
</thead>
<tbody>
<tr>
<td>ṭaːa³</td>
</tr>
<tr>
<td>ṭaːa⁵</td>
</tr>
<tr>
<td>ṭaːj⁴</td>
</tr>
<tr>
<td>ṭaːaj⁴</td>
</tr>
</tbody>
</table>

The transliteration of the consonants in this study is straightforward. Aspirated plosives are represented by the relevant plosive followed by an <h>, the affricate /ts/ is represented by a <c>. The palatal approximant /j/ is represented by a <y>, and the palatal nasal /ɲ/ as <ny>.

<table>
<thead>
<tr>
<th>Table 5 Consonant inventory and transliteration for this study</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPA</td>
</tr>
<tr>
<td>k</td>
</tr>
<tr>
<td>kʰ</td>
</tr>
<tr>
<td>ɲ</td>
</tr>
<tr>
<td>ts</td>
</tr>
<tr>
<td>s</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>t</td>
</tr>
<tr>
<td>tʰ</td>
</tr>
<tr>
<td>n</td>
</tr>
<tr>
<td>?</td>
</tr>
</tbody>
</table>

\textsuperscript{10} See Owen (2013) for a complete overview of Shan-Ni graphemes of Pali consonants. I have included only the /ɾ/, since this consonant is present in commonly used terminology, such as \textipa{ɾa'hɑ̂ːント⁴ 'monk'}.  

33
2.1.1 Consonant clusters

The first consonant is followed by either a vowel or an approximant, in that case forming a consonant cluster. The velar approximant /w/ can follow the velar stops /k/ and /kʰ/, and the palatal /j/ can follow the stops /k, kʰ, t, tʰ, p, pʰ/ and the labial nasal /m/. In clusters where the palatal /j/ follows the velar /k/ or /kʰ/, it modifies the sound of the cluster to an alveolar palatal /tɕ/ or /tɕʰ/. The palatal approximants is represented in the transliteration by a <y>, hence for example /pj/ is represented as <py>. The velar approximant /w/ is represented by a <w>, hence /kw/ is represented as <kw>.

These consonant clusters occur in both native words and Burmese loans. If they are native words, they have been through a sound change unique to Tai varieties of northern Myanmar (see section 7.1). The palatal approximant /j/ in native words occurs between consonants and the vowels /e/, /i/ and /ɛ/. Similarly, the clusters <kw> and <khw> are always followed by /a/ or /aː/. Since most cognates in Tai languages have retained the same vowels and consonants, the addition of approximants in Shan-Ni stands out. In other Tai languages, these consonant clusters only exist in loan words.

The alveolar tap /ɾ/ is also as second part of the cluster only used for religious terms of Pali origin. As part of the cluster many speakers tend to pronounce as a lateral approximant with an epenthetic vowel preceding it, making the pronunciation of phraa⁴ more like phə¹laa⁴.

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Orthography</th>
<th>Transliteration</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>kw</td>
<td>kw</td>
<td>kwaa¹</td>
<td>‘go’</td>
</tr>
<tr>
<td>kʰw</td>
<td>kw</td>
<td>khwam⁴</td>
<td>‘language’</td>
</tr>
<tr>
<td>tɕ</td>
<td>ky</td>
<td>kyin⁴</td>
<td>‘eat’</td>
</tr>
<tr>
<td>tɕʰ</td>
<td>khy</td>
<td>khyem³</td>
<td>‘needle’</td>
</tr>
<tr>
<td>pj</td>
<td>py</td>
<td>pyet¹</td>
<td>‘duck’</td>
</tr>
<tr>
<td>pʰj</td>
<td>phy</td>
<td>phyit¹</td>
<td>‘to quarrel’</td>
</tr>
<tr>
<td>mj</td>
<td>my</td>
<td>myit⁵</td>
<td>‘to slice’</td>
</tr>
<tr>
<td>pʰr</td>
<td>phr</td>
<td>phraa⁴</td>
<td>‘temple’</td>
</tr>
<tr>
<td>tr</td>
<td>tra</td>
<td>traa⁴</td>
<td>‘Buddhist law’</td>
</tr>
</tbody>
</table>

2.1.2 Final consonants

Syllables can either be open (ending in a vowel) or closed (ending in a consonant). When syllables are closed, vowels can be followed by any unaspirated plosive, nasal or approximant. Below is an overview of all different possibilities in combination with an initial consonant /k/ [k], the unmarked vowel /a/ and unmarked tone /1/. In the script they are represented...
with their regular grapheme with a little crescent on top, marking the difference between an initial consonant /k/ [κ] and a final consonant /k/ [ɔ],. The pronunciation of the final consonant can be more or less explicit, with stops sometimes being quite silent, approaching a glottal. The approximant /j/ is represented in the script in two ways, depending on the vowel preceding it. Following the vowels /a/, /a:/ and /o/ it is represented by a high comma, illustrated the table 7 by kaj¹ ‘chicken’ [ɔ]. With the vowels /u/ and /o/ the symbol for <ny> is used, illustrated the table 7 by koj² banana [ŋ]. The final [ŋ] is sometimes aspirated.

<table>
<thead>
<tr>
<th>Table 7 Final consonants</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPA</td>
</tr>
<tr>
<td>kap</td>
</tr>
<tr>
<td>kat</td>
</tr>
<tr>
<td>kak</td>
</tr>
<tr>
<td>kaw</td>
</tr>
</tbody>
</table>

2.2 Vowels

Shan-Ni has an elaborate vowel system, with an especially large group of back vowels, including diphthongs. An overview is given in Table 8. There is distinction in vowel length between /a/ and /a:/, but this does not apply for other vowels. In speech, vowels are sometimes reduced to an /ɤ/ or /ə/, but while /ɤ/ is also a phoneme by itself, /ə/ is not, although it is common to be used in a non-stressed position instead of /a/, especially in disyllabic words.

<table>
<thead>
<tr>
<th>Table 8 Vowel inventory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front unrounded</td>
</tr>
<tr>
<td>Close</td>
</tr>
<tr>
<td>Close-Mid</td>
</tr>
<tr>
<td>Open-Mid</td>
</tr>
</tbody>
</table>

Since Shan-Ni has many vowels, I use a few IPA vowel symbols in my transliteration, with the exception of /a:/, which is represented as <aa>. Some vowels are represented differently in the orthography when they are followed by a consonant, so both variants are included in table 8. The independent symbols include the symbol for the “inherent vowel”, the /a/.

35
inherent vowel does not have a diacritic in this script, hence any consonant without diacritic or crescent (closing a syllable), is automatically followed by an /a/. Often the /a/ is not stressed and has a level tone. In that case, it is often pronounced as a /a/. Hence the symbol for /m/ <  noreferrer > in the Table below representing ma¹ ‘NEG’, and the the symbol for /k/ <  noreferrer > in  noreferrer in kat¹ ‘cold’ represents the sound ka.

<table>
<thead>
<tr>
<th>vowel</th>
<th>Independent symbol</th>
<th>Diacritics in open syllables</th>
<th>Diacritics in closed syllables</th>
</tr>
</thead>
<tbody>
<tr>
<td>/a/</td>
<td>noreferrer</td>
<td>ma¹</td>
<td>‘NEG’</td>
</tr>
<tr>
<td>/a:/</td>
<td>noreferrer</td>
<td>waa²</td>
<td>‘speak’</td>
</tr>
<tr>
<td>/i/</td>
<td>noreferrer</td>
<td>si¹</td>
<td>‘four’</td>
</tr>
<tr>
<td>/e/</td>
<td>noreferrer</td>
<td>me⁴</td>
<td>‘wife’</td>
</tr>
<tr>
<td>/ɛ/</td>
<td>noreferrer</td>
<td>te⁴</td>
<td>‘build’</td>
</tr>
<tr>
<td>/ɔ/</td>
<td>noreferrer</td>
<td>mɔ²</td>
<td>‘pot’</td>
</tr>
<tr>
<td>/u/</td>
<td>noreferrer</td>
<td>ngu⁴</td>
<td>‘snake’</td>
</tr>
<tr>
<td>/o/</td>
<td>noreferrer</td>
<td>ngo⁴</td>
<td>‘cow’</td>
</tr>
<tr>
<td>/ɯ/</td>
<td>noreferrer</td>
<td>mɯ²</td>
<td>‘hand’</td>
</tr>
<tr>
<td>/ɤ/</td>
<td>noreferrer</td>
<td>mɤ²</td>
<td>‘time’</td>
</tr>
</tbody>
</table>

In open syllables the vowels end the syllable, in closed syllables they are followed by a consonant, which can be an approximant /j/ or /w/. When they are followed by an approximant, this modifies the vowel itself. I consider the modified sounds that emerge diphthongs, as they are pronounced as a single phoneme. They cannot be followed by a consonant, which full vowels would. The vowel  noreferrer can be pronounced as either /aɯ/ or /ɛɯ/.

Some words can be pronounced in both ways, while others are more likely be pronounced as either /au/ or /ɛu/. The example  noreferrer ‘1SG’ can be pronounced as either kau⁴ or kɛu⁴, but the interrogatives are always pronounced with /ɛu/ (ɛu⁴ ‘which’) and the word for ‘new’  noreferrer as mau¹. In the transliteration I will use the variant that corresponds with the utterance.

The /oj/ is sometimes pronounced as  noreferrer making for example  noreferrer ‘glasspot’ sound like  noreferrer and  noreferrer ‘look’ sound like  noreferrer. This variant also exists in a few other Tai languages (Edmondson 2008: 198).

<table>
<thead>
<tr>
<th>Diphthongs</th>
</tr>
</thead>
<tbody>
<tr>
<td>kaj¹</td>
</tr>
<tr>
<td>paaj²</td>
</tr>
<tr>
<td>soj¹</td>
</tr>
<tr>
<td>kaj¹</td>
</tr>
<tr>
<td>kaj¹</td>
</tr>
<tr>
<td>kau⁴</td>
</tr>
</tbody>
</table>

Looking at the languages in the region, this might be an ongoing sound change in Shan-Ni. Cognates in different languages have sounds similar to either /au/ or /ɛu/. However, this variety does not seem to be common within languages (as least this is not described).
2.3 Tone

For tone there is a slight distinction between the Myitkyina (eastern) variant of Shan-Ni and the Indawgyi (central) variant of Shan-Ni. The Myitkyina variant is used in literacy books for school children across the region, and was also the variant Wyn Owen (2013) based his phonological overview on. In that variant, there are six tones. My data is from the Indawgyi region, and has five tones. However, sometimes the sixth tone is represented in the written language for words that are written very frequently, for example the word for ‘year’ pi⁶, which is included in documents with a date and posters and merchandise for new year celebrations. Therefore, I have included this tone in the overview as a symbol, but not as a phoneme. An overview of the tones is given in Table (11). The tones differ in pitch and contour, and can have additional features such as creakiness or a ‘longer’ tone. The order of the numbers corresponds with the order they are presented in in Shan-Ni literature and teacher trainings. Tone 1 is the unmarked level tone, Tone 2 the low dash representing the low falling creaky tone, Tone 3 the high dash representing the high level tone, Tone 4 the colon representing the midfalling ‘long tone’ and Tone 5 the low dot the high falling creaky tone. Words that have the the sixth tone in Myitkyina, at Indawgyi either have the second (low falling) or fourth (mid slightly falling) tone. This development will be discussion further in Chapter 3.

<table>
<thead>
<tr>
<th>Pitch</th>
<th>Description</th>
<th>Additional features</th>
<th>Orthography</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>ka³³</td>
<td>midlevel</td>
<td>-</td>
<td>(unmarked)</td>
<td>ho³ma¹</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>မိုး ‘shoulder’</td>
</tr>
<tr>
<td>ka²¹ˀ</td>
<td>low falling</td>
<td>glottalized/creaky</td>
<td>-</td>
<td>maa²</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>မ ‘mad’</td>
</tr>
<tr>
<td>ka³⁵</td>
<td>high rising</td>
<td>-</td>
<td>-</td>
<td>maa³</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>က ‘dog’</td>
</tr>
<tr>
<td>ka⁴⁴²</td>
<td>mid falling</td>
<td>‘long’ tone</td>
<td>:</td>
<td>maa⁴</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ယ ‘come’</td>
</tr>
<tr>
<td>ka⁴¹ˀ</td>
<td>high falling</td>
<td>glottalized/creaky</td>
<td>-</td>
<td>maa⁵</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>န ‘horse’</td>
</tr>
<tr>
<td>Same as tone 2 or tone 4</td>
<td>-</td>
<td></td>
<td>pi⁶</td>
<td>တ ‘year’</td>
</tr>
</tbody>
</table>

The development of tones in Tai languages is very systematic, and can be studied through ‘Gedney’s box’ (Gedney 1989). Depending on the original phonology of the words, they can be categorized in a “boxes” (Table 12). The vertical axis categorizes words based on the original consonant in Proto-Tai, and the horizontal axis categorizes words based on whether they ended in a certain tone or a long or short dead syllable. Hence, all words that in Proto-Tai would have a voiced consonant and “Tone A” are together in one box (A4). Though currently tones differ among the Tai languages, all words that belong to a certain box (for example A4) are likely to have the same tone. Based on this, Gedney compiled a word list, connecting sets of words to each box. When studying a Tai language, this can be used to check
the tonal categories. Usually words belonging to the same box will have the same tone, and based on that different boxes with the same tones can be grouped together. For Shan-Ni this works very well, most words belonging to the same box indeed share the same tone. When grouping boxes with the same tones together, this results in the overview shown in Table 10.

I have given one example in each box, and gave them a colour to illustrate which boxes share the same tone. In this overview, the mid-level tone 1=B123-DS123-DL123, low falling (creaky) tone 2 = B4-C123, high rising tone 3=A1, mid falling (long) tone 4=A234 and high falling (creaky) tone 5=C4-DS4-DL4. This system is designed for monosyllabic Tai languages, but in Shan-Ni disyllabic words are common. As this method is based on of proto-Tai, it thus applies to the syllables which are cognates to the old proto-Tai lexemes targeted. In box B3, ma derives from proto-Tai, while the modern Shan-Ni word for ‘shoulder’ is ho³maa¹. In box DL3, mɔk¹ derives from proto-Tai while mɔk¹ya² is the Shan-Ni word for ‘flower’. The syllables of disyllabic words in the table that correspond with the targeted Proto-Tai cognates are therefore underlined.

Table 12 Gedney’s box (adapted from Gedney (1989: 2002))

<table>
<thead>
<tr>
<th>Tone A</th>
<th>Tone B</th>
<th>Tone C</th>
<th>D-long</th>
<th>D-short</th>
</tr>
</thead>
<tbody>
<tr>
<td>Voiceless friction sounds *s, hm, ph, etc.</td>
<td>A1</td>
<td>B1</td>
<td>C1</td>
<td>DL1</td>
</tr>
<tr>
<td>Voiceless unaspirated stops *p, t, k etc.</td>
<td>A2</td>
<td>B2</td>
<td>C2</td>
<td>DL2</td>
</tr>
<tr>
<td>Glottal *ʔ, ʔb etc.</td>
<td>A3</td>
<td>B3</td>
<td>C3</td>
<td>DL3</td>
</tr>
<tr>
<td>Voiced *b, m, l, z , etc.</td>
<td>A4</td>
<td>B4</td>
<td>C4</td>
<td>DL4</td>
</tr>
</tbody>
</table>

Table 13 Gedney’s box with Shan-Ni tonal categories and samples in each box

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>C</th>
<th>DS</th>
<th>DL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3 high rising</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>hu³ ka³</td>
<td>khaj¹ ʔa³</td>
<td>mɔk¹</td>
<td>khat¹ ʔa³</td>
</tr>
<tr>
<td></td>
<td>‘ear’</td>
<td>‘egg’</td>
<td>‘pot’</td>
<td>‘broken’</td>
</tr>
<tr>
<td>2</td>
<td>4 mid falling (long)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>kyin⁴ ʔo³</td>
<td>kaj¹ ʔo³</td>
<td>kaw² ʔo³</td>
<td>khyet¹ ʔo³</td>
</tr>
<tr>
<td></td>
<td>‘chicken’</td>
<td>‘nine’</td>
<td>‘frog’</td>
<td>‘lungs’</td>
</tr>
<tr>
<td>3</td>
<td>lin⁴ o³</td>
<td>ho³ma⁴ ʔo³</td>
<td>maan² ʔo³</td>
<td>naa²k¹ ʔo³</td>
</tr>
<tr>
<td></td>
<td>‘earth’</td>
<td>‘shoulder’</td>
<td>‘village’</td>
<td>‘chest’</td>
</tr>
<tr>
<td>4</td>
<td>mu⁴ o³</td>
<td>2 low falling</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mu⁴ o³</td>
<td>(creaky)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘hand’</td>
<td></td>
<td>nam⁵k¹</td>
<td>lɤt⁵</td>
</tr>
<tr>
<td></td>
<td>‘father’</td>
<td></td>
<td>‘water’</td>
<td>‘blood’</td>
</tr>
<tr>
<td>5</td>
<td>5 high falling</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>mɔk¹ya²</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>‘flower’</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The word *pi⁶* - ‘year’ belongs to box A2, hence at Indawgyi is pronounced as *pi⁴*. Table 14 shows the tone box for Myitkyina Shan-Ni (adapted from Wyn Owen 2013). The sixth tone in Myitkyina is used in words that fall under the boxes A2, A3 and B4. In Indawgyi Shan-Ni A2 and A3 are categorized as Tone 4 (mid falling long), while B4 is categorized as Tone 2 (low falling creaky). In Myitkyina Shan-Ni both Tone 4 (mid long) and Tone six (mid falling) are similar to Tone 4 (mid long falling) at Indawgyi, but the distinction is significant in that variety, and Tone 6 is also used for box B4 which in Shan-Ni is Tone 2 (low falling creaky), which is quite distant.

| Table 14 Tones in Myitkyina Shan-Ni (adapted from Wyn Owen 2013) |
|-------------------|--------------|--------------|--------------|-------------------------|
|                   | A            | B            | C            | DS              |
| 1                 | 3 high rising | 1 mid level  | 2 low creaky | 1 mid level    |
| 2                 | 6 mid falling |              |              |                |
| 3                 |              |              |              |                |
| 4                 | 4 mid long   | 6 mid falling |              | 5 high falling creaky |

### 2.4 Syllables

A Shan-Ni syllable exists of at least a consonant, a vowel, and a tone. If the initial consonant is a plosive, it can form a cluster with an approximant. The syllable can be open or closed in by a consonant. If the syllable is closed, the vowel can be followed by a nasal or a stop. These can be alveolar, velar or palatal. In summary, the structure can thus be displayed as C(A)V(C)T, hence allowing the following variations;

<table>
<thead>
<tr>
<th>Table 15 Syllable structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>CVT</td>
</tr>
<tr>
<td>CAVT</td>
</tr>
<tr>
<td>CVCT</td>
</tr>
<tr>
<td>CAVCT</td>
</tr>
</tbody>
</table>

The majority of Shan-Ni words are monosyllabic, although in comparison with other Tai languages it has a relatively large disyllabic group in its lexicon. Some of these are compounds of two monosyllabic words, which both exist in their separate forms in Shan-Ni as well. I have not found any phonological differences between the same word in isolation and as part of a
compound yet$^{12}$.

A full syllable cannot have a $<\text{C}_{a^1}>$ structure with an /a/ and a level tone /¹/. There are two types of syllables that can have a $<\text{C}_{a^1}>$ structure: prefixes and initial syllables of disyllabic words. In both cases, this is probably the result of these syllables being un unstressed, and have gradually developed a reduced phonology. The prefixes are listed in Table 16, with $ma^1$ being the negator, $te^1$ the future marker, and $kha^1$ the deseritative. Though $te^1$ is not represented by a lone consonant like the others, the pronunciation is the same.

Table 16 Prefixes with reduced phonology

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Vowel</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ma^1$</td>
<td>ə</td>
<td>NEG</td>
</tr>
<tr>
<td>$te^1$</td>
<td>əəə</td>
<td>FUT</td>
</tr>
<tr>
<td>$kha^1$</td>
<td>ə</td>
<td>DES, full form $khai^2$</td>
</tr>
</tbody>
</table>

Within disyllabic words the $<\text{C}_{a^1}>$ only occurs in initial syllables (Table 17). These are reduced forms of previously meaningful initial syllables of compounds, which can be shown by a number of words in transition that accept two forms, one that is phonologically reduced and one that is not. In their reduced forms they are not transparent. Because full syllables cannot have the syllable structures these reduced variants have, they are referred to as minor syllables, and words that start with a minor syllable as sesquisyllables or ‘one-and-half syllables’. This is described in more detail in Chapter 3. For the phonology, it matters that a $<\text{C}_{a^1}>$ is never independent, but dependent on the word following it.

Table 17 Disyllabic words with a $<\text{C}_{a^1}>$ structure as initial syllable

<table>
<thead>
<tr>
<th>Word</th>
<th>Vowel</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>$ma^1su^3$</td>
<td>əə</td>
<td>‘they’</td>
</tr>
<tr>
<td>$a^4sang^4$</td>
<td>əəəə</td>
<td>‘now’ / HESIT</td>
</tr>
<tr>
<td>$a^4saak^5$</td>
<td>əəəə</td>
<td>‘age’</td>
</tr>
<tr>
<td>$na^4hu^4$</td>
<td>əəə</td>
<td>‘how’</td>
</tr>
<tr>
<td>$pha^1b^3$</td>
<td>əəə</td>
<td>‘onion’</td>
</tr>
<tr>
<td>$kha^1lau^4$</td>
<td>əəə</td>
<td>‘how much’</td>
</tr>
</tbody>
</table>

2.5 Allophones and phonological variation

A few sounds discussed in the sections above are allophones and allow variation. These are listed in Table 18. Sometimes this depends on the position of the phoneme within the word, sometimes this differs per word, sometimes on the function of the lexeme. The phonemes /k/, /kʰ/, /a/, /a:/ are often pronounced in their alternative pronunciation, though this might be considered less polite. The /a/ is usually pronounced as /ə/ when it is unstressed and/or not followed by a coda. The /a:/ is usually pronounced as /ə/ when it is a grammaticalized

$^{12}$ This might be something to look into for future research, especially since the reduction to a sesquisyllable does change both the vowel and the tone, one may expect to find this as at a less advanced stage in compounds as well.
TAM marker. For example maa⁴ ‘come’ is pronounced as /maː/ when it is the regular verb ‘come’, but is often reduced to /ma/ when it is used as TAM marker. The /l/ and /n/ are equal in status, and usually allow variation between each other and sometimes to /d/ and /ɾ/, though this is less common. There are some words that are usually pronounced with either /l/ or /n/, but in most words both varieites are acceptable. The only clear exception is the word for ‘sister’ naang⁴, which people agree should not be pronounced as *laang⁴. This is the same for /au/ and /ɛu/, which usually can be used in both ways. They are represented by the same symbol <ŋ>. The linker si³ allows variation to yi³ and i³. I think this is free variation, but because I am not sure yet, I always include the original pronunciation in the transliteration.

Table 18: Overview of allophones and phonological variation

<table>
<thead>
<tr>
<th>Phoneme</th>
<th>Alternative pronunciation</th>
<th>Symbol</th>
<th>Transliteration</th>
<th>Variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>/k/</td>
<td>/x/</td>
<td>k</td>
<td>word initial often /x/</td>
<td></td>
</tr>
<tr>
<td>/kʰ/</td>
<td>/ɣ/</td>
<td>kh</td>
<td>word initial often /ɣ/</td>
<td></td>
</tr>
<tr>
<td>/l/</td>
<td>/n/, /d/, /ɾ/</td>
<td>l, n, d, r</td>
<td>usually both /l/ and /n/ accepted, /d/ and /ɾ/ rare</td>
<td></td>
</tr>
<tr>
<td>/n/</td>
<td>/l/, /d/, /ɾ/</td>
<td>n, l, d, r</td>
<td>usually both /n/ and /l/ accepted, /d/ and /ɾ/ rare</td>
<td></td>
</tr>
<tr>
<td>/s/</td>
<td>/ʃ/, /ɔ/</td>
<td>s, y, ø</td>
<td>free variation in linker si³</td>
<td></td>
</tr>
<tr>
<td>/a/</td>
<td>/ə/</td>
<td>a</td>
<td>Unstressed or without coda often ə</td>
<td></td>
</tr>
<tr>
<td>/aː/</td>
<td>/ə/</td>
<td>aa</td>
<td>unstressed ə in TAM markers</td>
<td></td>
</tr>
<tr>
<td>/au/, /ɛu/</td>
<td>/au/, /ɛu/</td>
<td>au, ɛu</td>
<td>Word specific, either both or one of them accepted</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 3: The shape of words

With the exception of a few Burmese and perhaps Pali loans, the large majority of words in Shan-Ni have a clear Tai origin. Yet, the shape of Shan-Ni words distinguishes Shan-Ni from other Tai languages. Nouns are often disyllabic, words have shifted to another word class and gained grammatical functions, and otherwise independent words have contracted into new lexemes. These processes are often connected; a lexeme may become part of new disyllabic noun, while its old shape becomes a grammatical marker. While specific grammatical functions will mostly be discussed in other chapters (4, 5 and 6), this chapter is concerned with the development of the shape of words, and the different processes involved in shaping particular domains in the Shan-Ni lexicon.

A recurrent theme in this respect are the “layered” innovations. Certain innovations in Tai Mau described by Young (1985), appear to be the foundation on which Shan-Ni has built further. This concerns, for example, the determiners and interrogatives, which in Tai Mau have been ‘contracted’. Shan-Ni integrates these same contracted forms in its own innovations. The same goes for reduplication. Young (1985) describes a productive process for reduplication, and Shan-Ni is using fossilized elements of the same system, complemented by words from other word classes. Through these innovations, the shape of words in Shan-Ni in comparison to other Tai languages can reveal details about historical developments in the language and the broader language family.

This chapter will start by discussing the dynamics of compounding processes, followed by an analysis of the transition from compounds to sesquisyllables (1,5 syllables), disyllabic words of which the the first syllable is reduced and non-transparant. This is followed by an illustration of the same process in pronouns, demonstratives and interrogatives. Demonstratives and interrogatives include the same contractions described by Young (1985) for Tai Mau, but have expanded beyond these forms. The final section (3.3) shows how parts of an old system of reduplication are still present in Shan-Ni, though no longer productive. All these different developments together show layers of innovations, developments that have been interrupted at some point in history, and continued in another direction.

3.1 Compounding

Most words in Shan-Ni are monosyllabic; among nouns, however, the share of disyllabic words is a lot larger. Some of these disyllabic words are transparent compounds, with two items that can both be traced back to two separate lexemes. Other disyllabic words are not transparent; when the two syllables are split, the meaning of at least one of them cannot be traced back to an earlier monosyllabic form. The presence of non-reduced compounds allows us to understand the way compounds are put together, the presence of words in transition allows us to study the process of how words lose their transparency.

Compound nouns in Shan-Ni are formed through a regular process, in which the first word refers to a ‘type’ of concept. In Lao there is a similar process, which Enfield refers to as class
terms, following the typology of Grinevald (2000). This is a specific group of words, which can form the initial part of a compound with a more general meaning, while also occurring as independent words. These initial parts categorize the whole compound (not only the element they are attached to) and are phonologically dependent and lexically specified. For Lao, Enfield (2007:147) describes a partial overlap with numeral classifiers, with class terms having a broader lexicon. In Shan-Ni there are class terms that also exist as classifiers, though I cannot say whether all classifiers can be used as class terms. In Shan-Ni class terms are lexically specific and obligatory elements which become part of a word, so they cannot be freely attached to words. While classifiers are used to count objects, class terms become part of the compound itself. I think it is likely that transparent classifiers qualify as a possible source for class terms, and that these can become obligatory class terms when used frequently by speakers. Morey (2005a: 229) refers to the equivalent of class terms in Phake and Aiton as noun class markers, but he does not enter too much in detail and it does not seem to be as common as in Shan-Ni.

What Enfield describes for class terms in Lao also applies to Shan-Ni, though the usage in Shan-Ni is more extensive, and, as will be shown, has advanced to the stage of phonological reduction and loosing transparency. A major difference is that while in Lao this only applies to objects and people (Enfield 2007:147), in Shan-Ni it is also applied for ‘intangible’ nouns referring to e.g. time, feelings and colours (see Table 19).

In both Lao and Shan-Ni, the same class term could be followed by a subcategory of the object the class term refers to, or it could be followed by a description of the object. For Lao, Enfield (2007: 146) gives the example of paa³ ‘fish’. The usage of paa³ is regardless of whether the class term paa³ ‘fish’ is followed by the name of a species (e.g. mackerel) or a description which needs paa³ ‘fish’ to form this particular meaning (e.g. ‘ink’ to form ‘inkfish’ or ‘gold’ to form ‘goldfish’).

While a compound with the name of a species (e.g. mackerel) might be classified as exocentric, and a compound with a descriptive feature (e.g. goldfish) as endocentric, what

<table>
<thead>
<tr>
<th>Table 19 Examples of compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class term</td>
</tr>
<tr>
<td>mɤ² ‘time’</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>phaa² ‘cloth’</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>si¹ ‘colour’</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>caɯ⁴ ‘heart’</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>
matters in Lao (and Shan-Ni) is that all fall within the category fish, and therefore start with phaa³. The same process applies to the formation of compounds in Shan-Ni. The first word is a more categorical term, which could be considered a class term. Table 19 illustrates how these class terms work in Shan-Ni, by showing different class terms as part of different compounds. For example, the class term phaa² 'cloth', can refer to all kinds of cloths, and can take different kinds of word (classes) to form new words. In combination with the verb sɤ¹ 'spread', it creates the concept 'bedsheet'. With the word ho³ 'head', phaa² ho³ 'turban' is formed. With taj⁴, the Tai/Shan autonym, the new meaning phaa² taj⁴ 'traditional clothes' is created.

Adding mɤ² 'time' to phok⁵ 'tomorrow' or cau⁴ 'heart' to on² 'happy' merely categorizes them as types of periods and feelings. However, adding mɤ² 'time' or cau⁴ 'heart' to maj² 'hot', both create new meanings; mɤ² maj² becomes 'hot season' and cau⁴ maj² becomes 'worried'. The result is that regardless of whether the second syllable has a different meaning on its own, all types of feelings start with cau⁴, and all types of temporal references start with mɤ². In Shan-Ni the first component is not necessarily phonologically reduced, and hence can be seen as a type of compound.

For colour terms, the class term si¹ 'colour' marks a difference in word class. As a noun, si¹ is an obligatory element of the colour terms. When the colour term functions as an adjective, it comes behind the noun, without si¹. For example, in (1), a 'red turban' is phaa² ho³ nɛng⁴, without si¹. This reflects the Burmese usage of colour terms, which as adjectives are also 'bare' colour terms, but as nouns are placed between the nominalizer ø- and -jaun 'colour'.

(1) a. ဗ ခ န င ် း ပ ါ င ် း b. ဗ ခ ါ င ် း ဗ ီ င ် း း ေ ပ ါ င ် း
    si¹  nɛng⁴  phaa² ho³  nɛng⁴
    colour  red  cloth  head  red
    'red'            'red turban'
    \[Burmese\]

In Lao, there is a distinction between 'basic' colour terms and derived colour terms (e.g. sky for blue, lotus for purple), with the second category sometimes being obligatory marked with sii3, while this is not obligatory for the category of basic colour terms (Enfield 2007: 148). In Shan-Ni there is no distinction between types of colours, only between colours in different word classes.
### 3.1.1 The formation of class terms

The formation or addition of class terms and hence the transition from monosyllabic words to disyllabic compounds, can be studied by looking at body parts. A number of body parts are listed in Table 20. Most body parts in Shan-Ni are disyllabic, and contain one syllable that corresponds with other Tai languages. Usually this is the second syllable (not the class term), though this is not always the case. To illustrate this, I compare Shan-Ni body parts with their Shan Gyi counterparts. Shan Gyi has a few compounds in body parts, but is predominantly monosyllabic. For example, the word ‘eye’ in Shan Gyi is *taa⁴*, which in Shan-Ni is the second part of word for *hoj¹taa⁴* ‘eye’. The class term *hoj¹* is also a classifier for round objects, referring here to the shape of the eye.

Class terms being more general, means that they are often used to refer to the shape or

<table>
<thead>
<tr>
<th>Body part</th>
<th>Compound</th>
<th>Etymology</th>
<th>Shan Gyi (SEALang)</th>
</tr>
</thead>
<tbody>
<tr>
<td>knee</td>
<td>ṭo⁴kha²</td>
<td>head + knee /kʰaw²/</td>
<td></td>
</tr>
<tr>
<td>shoulder</td>
<td>ṭo⁴maa¹</td>
<td>head + shoulder /maa²/</td>
<td></td>
</tr>
<tr>
<td>hair</td>
<td>ṭo⁴maa²</td>
<td>hair + head /kʰon¹/</td>
<td></td>
</tr>
<tr>
<td>eyebrows</td>
<td>ṭo²taa⁴</td>
<td>hair + eye /kʰon¹taa⁴/</td>
<td></td>
</tr>
<tr>
<td>heart</td>
<td>maa¹kʰon¹ceu⁴</td>
<td>fruit + head + feeling /ho¹tsau¹/</td>
<td></td>
</tr>
<tr>
<td>shin</td>
<td>naa²kʰyɛng²</td>
<td>shin (including face) /naa³ kʰɛŋ³/</td>
<td></td>
</tr>
<tr>
<td>chest (m)</td>
<td>naa²ok¹</td>
<td>face + chest /ʔɯk⁴/</td>
<td></td>
</tr>
<tr>
<td>chest (f)</td>
<td>naa²ceu⁴</td>
<td>face + feeling /ʔok⁴/</td>
<td></td>
</tr>
<tr>
<td>eye</td>
<td>hoj¹taa⁴</td>
<td>round + eye /taa¹/</td>
<td></td>
</tr>
<tr>
<td>hand palm</td>
<td>phaa¹mu⁴</td>
<td>palm/sole + hand /pʰaa²mu⁴/</td>
<td></td>
</tr>
<tr>
<td>cheek</td>
<td>phaa¹kʰyɛm²</td>
<td>palm/sole + cheek /kɛm³/</td>
<td></td>
</tr>
<tr>
<td>neck</td>
<td>ʔɔk⁴</td>
<td>neck/throat /kʰɔ⁴/</td>
<td></td>
</tr>
<tr>
<td>throat</td>
<td>kʰɔ⁴</td>
<td>hole + neck/throat /kʰɔ⁴/</td>
<td></td>
</tr>
<tr>
<td>mouth</td>
<td>sop¹</td>
<td>mouth/lips /sʰop⁴/</td>
<td></td>
</tr>
<tr>
<td>lips</td>
<td>ping³sop¹</td>
<td>edge + lips /sʰop⁴/</td>
<td></td>
</tr>
<tr>
<td>moustache</td>
<td>moj³sop¹</td>
<td>short body hair + mouth /nɔt²/</td>
<td></td>
</tr>
<tr>
<td>waist</td>
<td>hɛŋ²kʰɛŋ²</td>
<td>shape + waist /ʔɛŋ³/</td>
<td></td>
</tr>
<tr>
<td>brain</td>
<td>ʔɔk⁴ek¹</td>
<td>brain/marrow + yoke /ʔɔk⁴/</td>
<td></td>
</tr>
</tbody>
</table>

13 I have chosen Shan Gyi (or Tai Long) because it is the most spoken Shan language of Myanmar, and therefore has a useful and extensive database on the SEALang website allowing me to find cognates. It also has not gone through any of the processes described in this chapter, unlike some other Tai languages (to be discussed in Chapter 4).
structure of the object, while the second part of the compound is more specific. Hence depending on the position within the compound, the same lexeme can have a different function. The word ho³ ‘head’, for instance, can be added both as class term (first syllable) and as a second syllable, but the function is not the same. The class term ho³ ‘head’ is used for body parts that are ‘on top’ or ‘point to the outside’, like ho³khaw² ‘knee’ and ho³maa¹ ‘shoulder. When ho³ is added as a second syllable, it refers to something literally located on the head, for example khon³ho³ ‘hair’. In khôn³ho³ ‘hair (on the head)’ khon³ ‘hair’ is the class term, which is used in Shan-Ni for other types of hair as well, such as khon³taa⁴ ‘eyebrows’ which is also a compound in Shan Gyi. Shan-Ni also uses ho³ in maak¹ho³cew⁴ ‘heart’. The word Shan Gyi uses for heart, ho¹tsaw¹, also exists in Shan-Ni, but refers exclusively to the heart in the emotional sense. To refer to the physical heart, it needs to add maak¹ ‘fruit’, a distinction Shan Gyi does not make.

Shan-Ni also has a tendency to eliminate ambiguities. The word khɔ⁴ in Shan-Ni exclusively refers to ‘neck’, whereas the Shan Gyi cognates refers to both ‘neck’ and ‘throat’. To distinguish between ‘neck’ and ‘throat’, Shan-Ni includes kɔng⁴ ‘hole’ in the compound kɔng⁴khɔ⁴ ‘throat’. The same goes for sop⁴ which can mean both ‘mouth’ and ‘lips’ in Shan Gyi. In Shan-Ni, sop⁴ is only ‘mouth’, whereas ping¹sop¹, literally ‘the edge of the mouth’, refers to lips. In that same way, sop⁴ is used in mɔj³sop¹ ‘moustache’, with mɔj³ referring to ‘short body hair’. In these compounds, Shan Gyi uses a different monosyllabic word.

Another process is the grammaticalization of parts of compounds into class terms. For example, phaad¹ ‘sole/palm’ is used in other Tai languages in combination with ‘hand’ for ‘palm’ phaad¹mu⁴ and with ‘foot’ phaad¹tin⁴. Shan-Ni extends the usage to ‘cheek’, adding it to the Shan Gyi cognate kyɛm², forming disyllabic phaad¹kyɛm². The same goes for naa² ‘face/front’ which is part of the word for ‘shin’ naa²hyɛm², which is the front part of the leg. Shan-Ni goes beyond this and adds naa² also to other words, such as the male and female varieties of ‘chest’. But while the other part of the compound for ‘male chest’ is the same in Shan Gyi, for ‘female chest’ Shan-Ni uses naa²cew⁴ literally ‘the face of feeling’. This probably refers to the chest being in front of the heart, connecting it to (maak¹ho³cew⁴ ‘heart’ described above.

Through different processes, Shan-Ni is moving towards more disyllabic nouns. This allows the language to be more specific with concepts, eliminating ambiguities and describing features of the object referred to. For some words, this looks like a classifier function. However, the typical features marked by the initial syllable do affect the meaning, determining information about the concept as a whole, not just providing a category like classifiers do. They are not used for counting, like classifiers are. They become part of a new word, which sometimes in meaning may be similar to the older, monosyllabic variant, although there is often a slight semantic change, or a more specific distinction (e.g. kɔng⁴khɔ⁴ ‘throat’ rather than only khɔ⁴ ‘neck/throat’. At the stage discussed above, both parts of the compound are still transparent and can be traced back to two meaningful lexical items. The
next section will specify how they merge towards disyllabic words with less transparent initial syllables.

3.1.2 Two words becoming one

The addition of class terms to form compounds is a first step in a process that for some words is already at a more advanced stage. At this more advanced stage, the class terms have reduced their shape to the initial consonant +/ə/ and a level tone, represented as <a¹>. This type of unstressed vowel <a¹> exclusively occurs in prefixes and the first syllable of disyllabic words. In both cases, they are unstressed and followed by a full, stressed syllable. This is an areal feature, present in many Austroasiatic and Tibeto-Burman languages in the region, including Burmese, Jinghpaw and Palaung. Because these syllables are reduced to a shape that regular, full syllables never have, they are referred to as minor syllables or parts of sesquisyllables ‘one and a half syllables’ (Butler 2014: 11). The term sesquisyllables was first proposed by Matisoff (1973) for Austroasiatic. Butler (2014) dedicated her thesis to deconstructing the Southeast Asian sesquisyllable by studying it in Burmese, Khmer and Bunong. She lists the following properties for sesquisyllables (Butler 2014: 11):

a. Prosodic prominence is word-final
b. Non-final syllables are phonologically reduced
   - Non-final syllables have a reduced segmental inventory
   - Non-final syllables have a reduced syllable shape
   - Non-final syllables are light
     o Non-final syllables do not constitute well-formed prosodic words on their own.

These properties are also part of the sesquisyllables in Shan-Ni. The prosodic prominence is word-final, and the non-final syllables are phonologically reduced, limited to a /ə/ or /ɤ/. Their shape is reduced to a form which full words never have, and the non-final syllables are never well-formed prosodic words on their own. In Shan-Ni this phonological reduction is an active process. Several words are still in transition, and thus have two acceptable forms. A few of them are listed in Table 21.

<table>
<thead>
<tr>
<th>Table 21</th>
<th>Words in transition from a disyllabic to a sesquisyllabic shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>concept</td>
<td>Disyllabic shape</td>
</tr>
<tr>
<td>lips</td>
<td>luk⁵ɔn¹ (lit: offspring small)</td>
</tr>
<tr>
<td>mango</td>
<td>ping¹sop¹ (lit: edge mouth)</td>
</tr>
<tr>
<td>child</td>
<td>maak¹mong² (lit: fruit mango)</td>
</tr>
</tbody>
</table>

14 Burmese is a Tibeto-Burman language, Khmer and Bunong are Austroasiatic languages.
While in their reduced form the minor syllables are not transparent, in their shape as constituents of full compounds they are clearly class terms. The word ‘lips’ presented in Section 3.1.2 in its complete form ping¹sop¹ is often pronounced as pa¹sop¹. The fruit species ‘mango’ can be referred to as both ma¹mong² and maak‘mong², showing clearly that ma¹ is a reduced form of maak‘fruit’. The word luk⁵n¹ ‘child’ is in transition to la¹wn¹. Although luk⁵ is clearly the word for ‘offspring’ and ping¹ for edge, in their alternative forms la¹ and pa¹ they are not recognizable as such. The words in Table 22 are words that have already been through the process. Their first syllables have lost their transparency, and can no longer be traced back to an independent morpheme. For people, pa¹ is a common prefix, explaining the pa¹ in pa¹ying⁴ ‘girl’. However, pa¹ is also part of pa¹laa³ ‘almost’, where this would not fit.

The words pha¹lɔ³ ‘onion’ and pha¹maaw¹ ‘bachelor’ have the same initial syllable, but are also unlikely to have derived from the same source. It is thus impossible to derive the origin from the shape of the initial syllable itself.

| Table 22 Words with unmarked consonant as initial syllable |
|----------------|----------------|
| la⁴ling⁴       | ma⁴ding⁴      | ‘monkey’    |
| a’saak³        | ma⁴saak⁴      | ‘age’       |
| pa⁴ying⁴      | ma⁴ding⁴      | ‘girl’      |
| pha⁴lɔ³        | ma⁴ding⁴      | ‘onion’     |
| pha⁴maaw¹      | ma⁴ding⁴      | ‘bachelor’  |
| pa⁴laa³        | ma⁴ding⁴      | ‘almost’    |

3.1.3 Personal pronouns

The formation of disyllabic nouns on the one hand, and the grammaticalization of the monosyllabic variants of the same origin on the other can be illustrated by looking at the personal and possessive pronouns. The second and third person personal pronouns have become disyllabic, while the monosyllabic variants are still used as possessive pronouns. The third person plural possessive khaw³ also has acquired a few other function, which will be discussed in this section as well.

<table>
<thead>
<tr>
<th>Table 23 Personal and possessive pronouns</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Personal pronouns</strong></td>
</tr>
<tr>
<td>1SG kau¹</td>
</tr>
<tr>
<td>2SG mɛɯ⁴</td>
</tr>
<tr>
<td>3SG mu⁴/man⁴</td>
</tr>
<tr>
<td>1PL haw⁴</td>
</tr>
<tr>
<td>2PL ma²su³</td>
</tr>
<tr>
<td>3PL ma¹haw³</td>
</tr>
</tbody>
</table>
Table 23 shows the personal and possessive pronouns in Shan-Ni. The personal and possessive pronouns for the first, second and third person singular and the first person plural are identical, while for the second and third person plural the personal and possessive pronouns differ. When a personal pronoun follows a noun, it is a possessive pronoun. For the singular pronouns and first person plural haw⁴ the same pronoun is used, hence the interlocutor has to interpret from the position of a pronoun whether it is meant to be possessive or not. In the example below, kau⁴ ‘1sg’ follows the noun luk⁵ ‘son’, hence it is a possessive forming the clause ‘my son’. This is followed by a new clause, starting with meu⁴ ‘2sg’ without a noun preceding it, hence it refers to the personal pronoun ‘you’.

(2) luk⁵ kau⁴ meu⁴ kha¹ laaj² sang⁴
child 1sg 2sg DES get what
‘My son, what do you want?’  
(BSS_29)

For the second and third person plural, there is a difference between the possessive and non-possessive personal pronouns. The possessive pronouns are the same as the personal pronouns of other Tai languages in the region, but the personal pronouns have developed into disyllabic lexemes. They have both acquired the prefix ma¹, which may have been derived from any word starting with the consonant /m/. It is likely that this /m/ has been taken from the second and third person singular, meu⁴ and mun⁴ respectively. In addition, the initial consonant /kh/ of khaw³ is reduced to /h/ becoming ma¹haw³. I have encountered ma¹haw³ as well in the newspaper, but not in spoken language. Since the tone is the same, it is likely that ma¹haw³ derived from khaw³. That both forms are used by the same speakers in different positions to express different functions, is illustrated in the examples below. Both sentences (3) and (4) are part of the same story, uttered by the same character (stepmother) addressing the same two daughters (Yee and Ee). In (3) ma’su³ ‘2pl’ is used as a personal pronoun, while in (4) the possessive pronoun su³ ‘2pl’ follows pɔ² ‘father’, forming the nominal phrase ‘your father’.

(3) ma’su³ pi²nɔng⁵ lɔng² an² kau⁴ cau⁴ waj⁵ naa⁵
2pl sisters work thing 1sg order KEEP TAG
‘You sisters, the work I ordered you...’  
(BSS_210)

(4) pɔ² su³ pok⁵ maa⁴ maa¹ mɔk¹
father 2pl return come neg announce
‘Your father has returned unannounced.’  
(BSS_216)
khaw³ has gained a range of functions. In example (5) it functions as a conjunction connecting ‘Yee’ and ‘Ee’. In (6) and (7) it is a type of plural marker, which following two nouns denotes a category of people, for example ‘biological parents’ (6) or grandparents (7). If this literally refers to two people, it requires to be followed by the classifier construction sɔng³ kɔ⁵. As a plural marker, khaw³ can only be used for humans. Plurality is discussed in more detail Chapter 5.

(5) ɓu kw i³
    ye² khaw³ i¹
    yee 3PL ee
‘Yee and Ee’ (first and second daughter) (BSS_126)

(6) ɓu kw i³ kw i³ kw i³ kw i³
    po² leng⁵ me² leng⁵ khaw³
father give.birth mother give.birth 3PL
‘biological parents’ (BSS_80)

(7) ɓu kw i³ kw i³ kw i³ kw i³
    a¹ pu¹ a¹ yaa² khaw³
for grandfather for grandmother 3PL
‘for grandparents/elderly people’ (BSS_78)

(8) pu¹ ta¹ yaa² sɔng³ kɔ⁵ yang⁴ a¹
    grandfather and grandmother two CLF.people EXIST PRT
‘There were a grandfather and a grandmother.’ (NN_ITL_3)

3.2 Contractions

Besides nouns, determiners and interrogatives can also be disyllabic in Shan-Ni. In these words, the shape of the second part often corresponds with monosyllabic determiners and interrogatives in Tai Mau, but not necessarily with other Tai languages. According to Young (1985), these determiners and interrogatives have derived from two different morphemes themselves, and have contracted together into a new (usually monosyllabic) shape. Disyllabic determiners and interrogatives in Shan-Ni thus have gone through two processes. The first process (shared with Tai Mau) resulted in contraction, and the second process (exclusively in Shan-Ni) formed disyllabic determiners and interrogatives. Both processes will be discussed in this section.
Young (1985:24) lists a number of Tai Mau words\(^{15}\) that emerged from contractions of two words that were still present in the language, but merged together to create a new meaning. For example, the words *ti* ‘place’ and *näi* ‘this’ have merged together to form *thài* ‘here’. The same *ti* ‘place’ also merged with *lai* ‘which’, forming the new word *thài* ‘where’. Some of the words she includes in her list are also present in Shan-Ni\(^{16}\), though for most of them Shan-Ni is going a step further. I have summarized the contractions she discusses for Tai Mau and their innovations in Shan-Ni in Table 24. I assume that part of this is a shared development between Shan-Ni and Tai Mau, which at a later stage has developed into two different ways for both languages. The Tai Mau contractions listed below, are present in different forms in Shan-Ni. The interrogatives *theu*\(^4\) ‘where’ and *phaeu*\(^4\) ‘who’ are the same in both languages; The Tai Mau determiners *thài* and *thàn* are preceded by an extra word in Shan-Ni, resulting in *ti²thài\(^2\) ‘here’ and *ti²/pùn\(^4\) than\(^2\) ‘(over) there’, respectively. The interrogative *kha’leu*\(^4\) ‘how much’ is a reduced form of the Tai Mau *khaî* ‘how much’ with an additional *leu*\(^4\) ‘which’. In addition, there are also other processes within Shan-Ni that have affected the shape of determiners and interrogatives. The following subsections will discuss the processes in these two domains.

<table>
<thead>
<tr>
<th>Table 24 Contractions in Tai Mau (Young 1985) represented in Shan-Ni</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original morphemes in Tai Mau</td>
</tr>
<tr>
<td><em>ti</em> ‘place’ + <em>näi</em> ‘this’</td>
</tr>
<tr>
<td><em>ti</em> ‘place’ + <em>nân</em> ‘that’</td>
</tr>
<tr>
<td><em>ti</em> ‘place’ + <em>lai</em> ‘which’</td>
</tr>
<tr>
<td><em>phù</em> ‘person’ + <em>lai</em> ‘which’</td>
</tr>
<tr>
<td><em>kaa</em> ‘price’ + <em>hî</em> ‘or not’</td>
</tr>
</tbody>
</table>

### 3.2.1 Determiners

While *ti²* ‘place’ is already included in the contractions *thài\(^2\) and *than\(^2\) (see Table 24), Shan-Ni requires an extra *ti²* to form *ti²thài\(^2\) ‘here’ and *ti²than\(^2\) ‘there’. In addition, *pùn*\(^4\) ‘yonder’ can be used instead of *ti²* in *pùn*\(^4\)than\(^2\) ‘over there’ to mark a difference between something which is relatively near (*ti² than\(^2\)*) and something that is further away (*pùn*\(^4\)than\(^2\)*).

\[\text{(9) a. } \text{ti²thài\(^2\)} \quad \text{b. } \text{ti²than\(^2\)} \quad \text{c. } \text{pùn*than\(^2\)}\]

\[\text{‘here’} \quad \text{‘there’ (near)} \quad \text{‘over there’ (distant)}\]

\[(DS_LT_08)\]

\(^{15}\) *phaî* ‘who’ is not included in that list, but is used throughout Young (1985).

\(^{16}\) I do not exclude the possibility of other words from her list also having a presence in Shan-Ni, but in that case they have been through phonological processes and need more evidence than I can provide at this stage.
The words for ‘this’ and ‘that’ are also not expressed by a single syllable anymore. Though the cognates of nâi ‘this’ and nân ‘that’ in Tai Mau are present in Shan-Ni, they have developed additional functions. For the demonstrative ‘that’, Shan-Ni uses the same pun⁴ ‘yonder’ used in pun⁴ than² ‘over there’ to form pun⁴ nan⁵ ‘that (thing over there)’. pun⁴ nan⁵ always refers to an object in the distance, either unspecified without noun (10a), or preceded by a noun (10b).

\[(10) \quad \begin{array}{ll}
\text{a.} & \text{pun⁴ nan⁵} \quad \text{b.} & \text{pún⁴ puun⁴ nan⁵} \\
\text{that} & \text{house that} \\
\text{‘that (thing over there)’} & \text{‘that house’}
\end{array} \]

For nearby objects, Shan-Ni has added nang¹ to ne²/naj⁵, creating demonstrative nang¹ ne². As described above for pun⁴ nan⁵ ‘that (thing over there)’, nang¹ ne² can be used by itself (11a) or behind a noun (11b). When only ne²/naj⁵ is used behind a noun, it is a definite marker (11c) rather than a demonstrative.

\[(11) \quad \begin{array}{llll}
\text{a.} & \text{nang¹ ne²} & \text{b.} & \text{waan¹ nang¹ ne²} & \text{c.} & \text{waan¹ ne²} \\
\text{this} & \text{cup this} & \text{cup} & \text{DEF} \\
\text{‘this’} & \text{‘this cup’} & \text{‘the cup’}
\end{array} \]

By itself, ne²/naj⁵ has a range of functions, of which the most common can be analyzed as a definite marker or a topicalizer. The difference between using ne² and nang¹ ne² is illustrated below. With only ne² the speaker makes clear s/he is referring to ‘the water’ mentioned earlier in the story (12). When using nang¹ ne² the speaker implies there are perhaps multiple waterbodies, but only this specific one is getting big (13).

\[(12) \quad \begin{array}{l}
\text{nam² ne² yau¹ maa⁴} \\
\text{water DEF big COME} \\
\text{‘The water got big’}
\end{array} \]

\[(13) \quad \begin{array}{l}
\text{nam² nang¹ ne² yau¹ maa⁴} \\
\text{water this big COME} \\
\text{‘This water got big’}
\end{array} \]
3.3.1 Interrogatives

<table>
<thead>
<tr>
<th>Interrogatives</th>
<th>Contractions</th>
<th>Compounds</th>
</tr>
</thead>
<tbody>
<tr>
<td>lɛɯ⁴ 'which'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>theɯ⁴ 'where'</td>
<td>ti 'place'+ lai 'which'</td>
<td>-</td>
</tr>
<tr>
<td>phɛɯ⁴ 'who'</td>
<td>phù 'person'+ lai 'which'</td>
<td>-</td>
</tr>
<tr>
<td>mɯ² lɛɯ⁴</td>
<td>mɯ² 'time' + lɛɯ⁴ 'which'</td>
<td>-</td>
</tr>
<tr>
<td>mɯ² thɛɯ⁴</td>
<td>mɯ² 'time' + thɛɯ⁴ 'where'</td>
<td>-</td>
</tr>
<tr>
<td>kha¹lɛɯ⁴ 'how much'</td>
<td>kaa 'price' + h ̌ 'or not'</td>
<td>khaɨ + lɛɯ⁴ 'which'</td>
</tr>
<tr>
<td>kha¹lɛɯ⁴+CLF</td>
<td>kaa 'price' + h ̌ 'or not'</td>
<td>khaɨ + lɛɯ⁴ 'which'</td>
</tr>
<tr>
<td>na¹hɯ⁴ 'how'</td>
<td>-</td>
<td>nang¹ 'this' + hɯ⁴ 'how'</td>
</tr>
<tr>
<td>kɔp⁵ sang⁴ 'why'</td>
<td>-</td>
<td>kɔp⁵ 'reason' + sang⁴ 'what'</td>
</tr>
<tr>
<td>sang⁴ 'what'</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>le⁴/le² 'yes or no?'</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Shan-Ni uses the same interrogatives for ‘where’ and ‘who’ mentioned by Young (1985) as contractions in Tai Mau, but has applied its own processes to the interrogatives as well. In the overview given in Table 25, the ‘contractions’ refer to processes also present in Tai Mau, and ‘compounds’ refer to processes specific to Shan-Ni. Using the word mɯ² ‘time’ in combination with either lɛɯ⁴ ‘which’ or theɯ⁴ ‘where’ forms the interrogative ‘when’. While the other interrogatives have joined two words into a new shape, mɯ² lɛɯ⁴ and mɯ² theɯ⁴ have not gone through such a process, and in both cases mɯ² ‘time’ is still very much a separate lexeme from lɛɯ⁴ ‘which’ or theɯ⁴ ‘where’. The interrogative hɯ⁴ follows nang¹ ‘this’ in Shan-Ni, though this is often reduced to na¹hɯ⁴. The interrogative kha¹lɛɯ⁴ ‘how much’ underwent both the contracting and the compounding process. First it became a contraction described for Tai Mau as existing of kaa ‘price’ and h ̌ ‘or not’ with khaɨ as a result. In Shan-Ni, lɛɯ⁴ ‘which’ was added and khaɨ was reduced to kha¹ through the process initial syllables undergo, described earlier in this chapter. Shan-Ni uses kha¹lɛɯ⁴ both for ‘how much’ and ‘how many’, with ‘how many’ requiring an extra word specifying the type of thing asked about. This is described in more detail in section 4.4.
3.3 Reduplication

Shan-Ni reduplicates adjectives, adverbs and verbs. Reduplicated adjectives modify nouns they follow, and reduplicated adverbs modify verbs they follow. This form of reduplication is obligatory; a non-reduplicated adjective functions as a verb in an intransitive predicate, and many adverbs simply do not exist in a non-reduplicated form. The difference between the absence and presence of reduplication is shown in the following sentences, which each start with a disyllabic body part noun. Behind *kɔng²kaang⁴ 'chin', *yaaw⁴ 'long' is only used once (14), hence it functions as a verb, forming the phrase ‘the chin is long’. When *yaaw⁴ is reduplicated, it becomes an adjective, hence with *khon⁵ho³ 'hair', *yaaw⁴ *yaaw⁴ 'long.RDP' forms the phrase 'long hair' (15).

(14) ဗြဲနောင် အတာ
kɔng²kaang⁴ *yaaw⁴
chin   long
’[The] chin is long.’ (BSS_87)

(15) ဗြဲနောင် အတာအတွင်း
khon⁵ho³ *yaaw⁴*yaaw⁴
hair   long.RDP
‘long hair’ (BSS_87)

(16) ဗြဲ  စှာ  အတာ
kon⁴ tin⁴ *yaaw⁴
person  leg   long
’a long-legged person’
(Lit: '(S)he is a person of whom the legs are long’) (BSS_87)

In (16) it might seem like long-legged is an adjective modifying ‘person’, but actually these are two verb phrases built into each other. *tin⁴ *yaaw⁴ means ‘the legs are long’, which together functions as a verb in relation to *kon⁴ ‘person’. This is outlined below.

(17) a. [tin⁴] [ *yaaw⁴]
 N    V
 ’N (tin⁴) is V (yaaw⁴)’

b. [kon⁴] [tin⁴ *yaaw⁴]
 N    V
 ’N (kon⁴) is V (tin⁴ yaaw⁴)’

The constructions above are not part of the same process described in section 3.1, used to form disyllabic nouns. In nominal compounds the class term specifying the category comes first, followed by a specifying item. Adjectives and verbs follow the noun. Hence, when as in
(18a) huj¹’round’ precedes taa⁴, it is a class term part of the nominal compound. When pom² ‘round’ comes behind laa² ‘face’, it is a separate word, in this case a verb, creating the phrase ‘[the] face is round’.

Reduplicated adverbs modify the verbs they follow, in the same way adjectives modify the nouns they follow. In (19) the reduplication kyaang¹ kyaang¹ ‘fast’ modifies kwaa¹ ‘go’, in (20) cong³ cong³ ‘a little’ modifies sən³ ‘study’, in (21) lj⁴ lj⁴ ‘slow’ modifies kwaa¹ ‘go’ and in (22) nam³ nam³ ‘many’ modifies taan² khwaam⁴ ‘talk’. The adverbs in the examples below are not used without reduplication, unless they are preceded by the intensifier aaj⁴ ‘very’.

3.3.1 Remnants of an old system

Adjectives and certain intransitive verbs can be modified by reduplicated dummy words that have no meaning by themselves. Young (1985:19) calls this “affective reduplication” and describes how in Tai Mau there are pairs of these “nonsense items” with good and bad
connotations. These pairs usually exist of the same consonants, but have different vowels. Items with a good connotation use the vowels /ɔ, ē, i, e/ and items with a bad connotation use the vowels /u, o, a/. For example, to change ʔūn ‘warm’ to ‘very warm’, ʔūn tɔ̂t-tɔ̂t would indicate that it’s a pleasant kind of warmth, while ʔūn tût-tût would be the unpleasant variety. Similarly, kɛ̄n kɛ̌k kɛ̌k ‘very solid’ might refer to a solid construction that won’t fall apart anytime soon, kɛ̄n káak káak ‘very solid’ might refer to something that is supposed to be more smooth or flexible (Young 1985:21).

<table>
<thead>
<tr>
<th>Concept</th>
<th>Good Connotation</th>
<th>Bad Connotation</th>
</tr>
</thead>
<tbody>
<tr>
<td>‘warm’ ʔūn</td>
<td>ʔūn tɔ̂t-tɔ̂t</td>
<td>ʔūn tût-tût</td>
</tr>
<tr>
<td>‘solid’ kɛ̄n</td>
<td>kɛ̄n kɛ̌k kɛ̌k</td>
<td>kɛ̄n káak káak</td>
</tr>
</tbody>
</table>

(Young 1985: 20, 21)

This system is not productive in Shan-Ni, but remnants of this system are present. Several items listed by Young (1985) also feature in the Basic Spoken Shan-Ni book (Table 25).

| Table 25 Shan-Ni reduplication matching examples of affective reduplication in Young (1985: 19) |
|---------------------------------|-----------------|-----------------|
| Concept            | Reduplicated Form | Connotation    |
| ‘cold’ kat¹            | ʔōn tɔ̂t-tɔ̂t     | good            |
| ‘hot’ maj²            | maj²waat¹waat¹   | bad             |
| ‘red’ nɛng⁴           | nɛng⁴khying¹khying¹ | good           |
| ‘stiff’ khyɛng³       | khyɛng³khɔng¹khɔng¹ | good         |
| ‘sweet’ waan³         | waan³ta¹ɛm²ɛm²   | bad            |

However, there is only one variant of each in Shan-Ni, hence the distinction between ‘good’ and ‘bad’ cannot be verified for Shan-Ni. In Table 26 I give an overview of the pairs I found in both the Basic Spoken Shan-Ni book and in Young (1985), and the connotation they should have according to her. Of the other examples I found, the first three (hɛŋ²‘dry’, let¹ ‘sunny’ and yɔm³‘thin’) are concepts that are not in Young’s list but fit the pattern she describes. Others are partially included, but do not match completely. For example un¹ ‘warm’ is included in Young’s list, but not with mun²mun². Instead it is included as un tɔ̂t tɔ̂t (positive) – un tût tût (negative) (Young 1985: 21).

Tai Mau uses yot tsük tsük for ‘to trickle down’, of which the last part corresponds with yuuk⁵ cух⁵ cух⁵, which Shan-Ni uses for ‘wet’. Both are used when talking about rain, but not in the same way. The other variant Shan-Ni uses for ‘wet’, yam⁴, does have two variants, but they
do not necessarily represent a pair with a good and bad connotation, and start with another consonant. For ‘mushy’ Shan-Ni uses $\varepsilon^2paap^5paap^5$ while Tai Mau uses $yaam \tilde{\varepsilon}t \tilde{\varepsilon}t - yaam \tilde{\varepsilon}at \tilde{\varepsilon}at$.

### Table 26 Adjectives with affective reduplication in Shan-Ni fitting Young's typology

<table>
<thead>
<tr>
<th>'dry' $heng^2$</th>
<th>$\varepsilon^2yaam^4$</th>
<th>'warm' $un^1$</th>
<th>$\varepsilon^2yaam^4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
</tr>
<tr>
<td>'sunny' $let^1$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>'wet' $yu$k$^5$</td>
<td>$\varepsilon^2yaam^4$</td>
</tr>
<tr>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
</tr>
<tr>
<td>'thin' $yom^3$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>'wet' $yam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
</tr>
<tr>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
<td>$\varepsilon^2yaam^4$</td>
</tr>
</tbody>
</table>

In addition to the ones listed in tables 25 and 26, there are quite a lot of different reduplication patterns that resemble the system as described by Young, though they do not seem to follow the same rules. In view of the absence of the good/bad connotation, I expect that affective reduplication is no longer a productive process in Shan-Ni, but that the fossilized forms are present in the language. Shan-Ni takes in new words that are either verbs, adverbs or adjectives themselves. Besides the structure described above with a single item followed by two of the same items (XYY), another common pattern is two times two words (XXYY). These are more common patterns in Burmese as well, and might be gradually replacing affective reduplication in Shan-Ni.

\[
(23)\quad \text{mɤ² hɤ² naj⁵ kɔ⁵ phun³ tuk¹ sɔk⁵sɔk⁵ si³ yik⁵yik⁵ cɤk⁵cɤk⁵}
\quad \text{time monsoon DEF PRT rain fall trickle.RDP LNK moist.RDP trickle.RDP}
\quad \text{In the monsoon season it rains a lot.}
\quad \text{(BSS_109)}
\]

In the example above (23), $sɔk^5sɔk^5$ and $cɤk^5cɤk^5$ are connected to ‘trickle’ in Tai Mau ($jɔt\ tɔs\ tɔs$ (good), $jɔt\ tɔs\ tɔs$ (bad)), though they are used together in a way that they complement each other rather than being opposites expressing good and bad connotations. They are used together with $yik^5yik^5$, which refers to ‘moist’, a related concept connected to rain, though not a “nonsense item” or synonym to $cɤk^5cɤk^5$.

### 3.4 Summary

The shape of words in Shan-Ni illustrate the different innovation processes words have been through, sometimes in combination with each other. Shan-Ni shares parts of these developments with other languages, but other parts not. Class terms and similar phenomena exist in other Tai languages, but they do not complement the majority of nouns like they do
in Shan-Ni. The next step in the process, reducing the first syllable/constituent of the compound and turning the class terms into a minor syllable without a meaning of its own, is a process common in non-Tai languages in the area. The old monosyllabic form, is no longer a regular noun but gains a grammatical function. A good example to demonstrate this process are the personal pronouns. While in their monosyllabic shape the former personal pronouns are still used as possessive pronouns, the second and third person plural personal pronouns have acquired class terms.

The shape of the question words and determiners also reveal participation in two different processes. First the contractions, similar to Tai Mau, and second, the addition of class terms. Through the combination of these processes, Shan-Ni has developed complex forms to ask questions and talk about space, that do not correspond exactly with any other language in the region. This process of layering innovations is applied in a similar way to reduplication. Shan-Ni is using fossilized parts of what is a productive system in Tai Mau. Shan-Ni only uses one half of pairs, lacking the good and bad connotation distinction, and is also uses reduplication without the “nonsense items”.

The sesquisyllables, interrogatives and determiners will be compared in more depth to other Tai languages in Chapter 7, alongside other features that will be discussed in the coming chapters. The unique contribution the shape of words have to this, is that they illustrate the layeredness of developments, which might be less clear in other features, though it certainly affects the way these other features should be looked at.
Chapter 4: Sentence structure and types of sentences

In a sentence with an overt subject, object and verb without TAM or other grammatical markers, Shan-Ni has a subject-verb-object order, as in the sentences below. In (24) the dog scares a child, and in (25) someone calls a frog.

(24) maa³ khyauk⁵ luk⁵ɔn¹
    dog   scare.BM child
    ‘The dog scares the child.’

(25) man⁴ hong⁵ taaj⁴ khyet¹
    3SG  call    scream frog
    ‘He calls the frog.’

Sentences like the ones above are rare. Most sentences do not have an overt subject, object and verb without TAM or other grammatical markers. The subject, and sometimes the object, may be omitted through ellipsis. Grammatical markers are present in almost every sentence, influencing the word order and sentence structure. To understand how sentences work in Shan-Ni, this chapter will thus focus on different types of sentences and the grammatical processes that govern them. First information structure and ellipsis will be discussed, the expression of arguments within the sentence. Then different types of sentences with copula, existentials, questions and passives will be discussed. The other frequent grammatical markers that are not present in these constructions will be explained in chapter 5, and the usage of TAM markers in chapter 6.

4.1 Representing and omitting referents

The way referents are included in sentences in Shan-Ni depends on the information required in order for the speaker to convey the intended message. In order to make an accessible message with the least amount of processing effort, different types of coding are employed (Van Valin & La Polla 1997: 201). New referents need an introduction, inactive referents need to be activated, and active referent can be left out. When it is already clear to the listener who the referent is, the referent is ‘active’. In sentences with active referents, Shan-Ni omits the referent. This process is called ellipsis. To introduce a new referent an indefinite NP is used, and to activate an accessible but inactive referent a definite NP is used (Van Valin & La Polla 1997: 201).
Ellipsis in Shan-Ni can apply to one or more noun phrases that are omitted, because the speaker does not consider it necessary to mention them. This is either because they are clear within the context or because the speaker already mentioned them before. For example, sentences like (26), (27), and (28) are completely normal:

(26) pi⁴ han³ aw⁴ pɔk⁵ maa⁴ kɤn⁴
    if see take return come SHOULD
    ‘If [you] see [it], [you] should bring it back.’

(BSS_23)

(27) kwaa¹ kaa⁵ taaj⁴ waj⁵
    go go die keep
    ‘[he] went [into the fire] and died’

(EP_LPL_38)

(28) en² tok¹ ji³ te¹ kwaa¹
    play fall LNK FUT go
    ‘[he] falls and goes [down to the water]’

(KSN_FS_43)

Sentence (26) does not specify 1) who should be returning something 2) what is supposed to be returned. Similarly, (27) does not specify 1) who went (and died) and 2) where he went. (27) and (28) almost exclusively contain verbs and TAM markers. Hence, in a regular conversation between two people about themselves and their daily activities, a large component of their conversation may consist almost exclusively of verbs and TAM markers, because they can drop all noun phrases referring to themselves and everything obvious to themselves. This means that, for example, a verb like khaaj³ ‘sell’ which is clearly (di)transitive because it is used to refer to selling ‘something’, often is not accompanied by an argument specifying what is being sold, because everybody knows what is being sold.

The usage of ellipsis thus requires an alternative way to study sentence structures and transitivity. For Lao, Enfield (2008: 88) suggests that the best way to look at transitivity is to distinguish between phrases where the argument is present but ellipsed, and those in which there is no argument at all. In practice, this means that it is often impossible to verify the distinction. Rather than classifying verbs as ‘transitive’, ‘intransitive’ or ‘ambitransitive’, he suggests looking at a variety of argument structure constructions and the accessibility of verbs to those constructions. When looking into different argument structure constructions in Shan-Ni, the difference in word order seems to depend on the presence of certain
grammatical markers. Therefore, throughout this chapter and the next, sentence structure will be discussed in relation to the different grammatical markers.

4.1.2 Activating referents with *naj⁵/ne²*

The definite marker *naj⁵* (often reduced to *ne²*) originally derived from the meaning 'this'. The most common usage of *naj⁵* is as a definite marker, marking inactive or accessible referents. These are referents that either been mentioned before or are present in the context, though not active at the moment of speech yet. It is common for definite markers to assume this role (Van Valin & LaPolla 1997: 201). Because ellipsis in Shan-Ni is very common, active referents that are already the topic of conversation are usually left out. The usage of *naj⁵* allows the speaker to switch between different referents present within the context. In (29) this is illustrated by marking both 'father' and 'mother'. The preceding conversation is about family, and now we learn where the parents live; the father lives in Hpakant, and the mother lives in Yangon. In (30), the story is about two women, Me Long and Me On, hence *ne²* is used to clarify that this part is about what Me Long does. In (31) the speaker is talking about the age she will have. In (32) *ne²* marks a longer argument 'last month'.

(29) pɔ² naj⁵ u¹ ti² phaa³ kaan⁵ me² naj⁵ u¹ ti² taa¹koong¹
father DEF stay place stone.plates break.off mother DEF stay place yangon
'Father stays in Hpakant, mother stays in Yangon.' (BSS_30)

(30) mɛ⁴ long³ ne² suk⁵ kaa⁵ kan⁴
wife big DEF rinse GO VPL
'Me Long rinsed [the vegetables].' (CM_ML_8)

(31) kɛw⁴ naj⁵ saam³ sip¹ te¹ thon² yaw⁵
1SG DEF three ten FUT reach FINISHED
'I will be thirty years old then.' (BSS_32)

(32) pon⁵ mɤ² mɤ² lɤn⁴ ne² su¹ waj⁵
last time time month DEF buy KEEP
'It was bought last month.' (NY_NMLe_17)

In conversations and spoken narratives *man⁴* '3sg' is often used in combination with *naj⁵*. This would be the equivalent of English 'as for X ...'. It is the same as shown above in (29), where the mother and father are both marked in their own clause to inform where each of
them lives. In the following examples, U Than Sein is describing a family scene from a picture. Since there are multiple people on the picture, he points out the activities of different participants in the scene. For example, in (33) with pho³long³ man⁴ naj⁵ ‘as for the husband’ followed by a description of his activity het¹ lɔng² ‘does work’. After describing the other things happening in the picture, the ‘husband’ is reintroduced as ‘father’ with pɔ² man⁴ naj⁵ ‘as for the father’ followed by what he is about to do; te¹ kwaa¹ laɯ⁴ tɤn² ‘will go into the forest’ (34). The father is alone in the forest, and he remains the active referent in the story. Because it is clear that the referent is the father, the argument is dropped from that moment onwards (35).

(33) pho³long³ man⁴ naj⁵ het¹ lɔng²
husband 3SG DEF do work
‘The husband is working.’

(34) pɔ² man⁴ naj⁵ te¹ kwaa¹ laɯ⁴ tɤn²
father 3SG DEF FUT go in forest
‘The father will go into the forest.’

(35) te¹ kwaa¹ tɤn² te¹ kwaa¹ het¹ lɔng²
FUT go forest FUT go do work
‘He is going into the forest, he is going to work.’

General concepts such as money or rice, are directly understood by all interlocutors and require no introduction. Therefore, they are already treated as accessible referents (Van Valin & LaPolla 1997: 200). In (36) ngɯn⁴ ‘money’ is marked by ne² in the general statement ‘money doesn’t make people happy’.

(36) ngɯn⁴ ne² caa⁴ kon⁴ ma¹ hau⁴ com⁴ cau⁴
money DEF for person NEG give happy heart
‘Money doesn’t make people happy.’
Lit: Money does not give people a happy heart

4.1.3 Indefinite referents and classifier constructions

Indefinite constructions can be used to introduce new referents into the context (Van Valin & LaPolla 1997: 201). In (37) the classifier constructions fulfil that role. First classifier
constructions introduce a boy (la⁴n¹ kɔ⁵ nɤng²) and a dog (maa³ to⁴ nɤng²) who are new to the hearer. Besides these two referents there is also a frog (introduced with the classifier construction khyet¹to⁴ nɤng²) which was caught by ‘the boy’. Because this referent has already been introduced, la⁴n¹ ne² ‘the boy’ has a definite marker.

(37) la¹ɔn¹ kɔ⁵     nɤng² maa³ to⁴      nɤng² yaa¹  sɔj¹ khyet¹ to⁴
child CLF.human one dog CLF.animal one place pot frog CLF.animal

nym² ne² yang⁴ an² naj⁵ la¹ɔn¹ ne² waam⁵ aaj⁵
one DEF EXIST thing DEF child DEFcatch PRT

‘There are a boy and a dog, and a frog, which was caught by the boy.’

(KSN_FS_01)

When using nouns in combination with numerals, Shan-Ni uses numeral classifiers, with both the classifier and numeral following the noun. In case of a single object, the numeral nɤng² ‘one’ follows the classifier (38a), while with two or more objects the numeral precedes the classifier (38b). When following a classifier, the first consonant of nɤng² ‘one’ is often pronounced differently depending on what precedes it, making it sound like ɣɤng² or ʔɤng².

(38) a. ngo⁴ to⁴ nɤng²         ngo⁴ song³ to⁴
cow CLF.animal-one      cow two CLF.animal
‘one cow’            ‘two cows’

In a sentence, a classifier construction functions as a normal noun phrase. In (39) the person referred to has one shirt, expressed by sx² phuⁿ³ nɤng² with phuⁿ³ ‘cloth’ referring to the material of the shirt, and nɤng² ‘one’ to in the number of shirts. In (40) the speaker has three oranges, expressed by maak¹ cɔk¹ saam³ hoj¹ with saam³ referring to the number ‘three’, and hoj¹ to the shape ‘round’. The classifier hoj¹ is also used for fruits that are not necessarily round¹⁷, for example bananas. In (41) two different arguments that are classified with hoj¹ are listed by the speaker. The two classifier constructions remain separate with hoj¹ marking each of them individually, but they do follow each other without any linking particle.

(39) ya¹ mɯn⁴ sx² phuⁿ³ nɤng² yang⁴ u¹
at 3SG shirt CLF.cloth one EXIST STAY
‘She has one shirt.’

(BSS_18)

¹⁷ And other objects that are not necessarily round, e.g. stones.
‘I have three oranges.’

‘I gave them two bananas and a coconut to make sweets.’

Table 27  Numeral classifiers in Shan-Ni

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Used for</th>
</tr>
</thead>
<tbody>
<tr>
<td>ᵅə₄</td>
<td>an²  general</td>
</tr>
<tr>
<td>ƙɔ⁵</td>
<td>persons</td>
</tr>
<tr>
<td>ᵄɔ⁴</td>
<td>to⁴  animals</td>
</tr>
<tr>
<td>ᵁ³</td>
<td>hoj¹  round, solid objects, e.g. stones, oranges</td>
</tr>
<tr>
<td>ƙħɛp¹</td>
<td>flat, hard objects</td>
</tr>
<tr>
<td>ƙa₃</td>
<td>waan¹  round, open objects, e.g. bowls, cups</td>
</tr>
<tr>
<td>ƙhu₃</td>
<td>phuⁿ³ flat, flexible objects, e.g. blankets or leaves</td>
</tr>
<tr>
<td>ƙth₃</td>
<td>thun³  bags</td>
</tr>
<tr>
<td>ƙu⁴</td>
<td>hu⁴  holes, drops, dots</td>
</tr>
<tr>
<td>ƙm₂</td>
<td>mɔk¹  flowers</td>
</tr>
<tr>
<td>ƙɔ⁵</td>
<td>kɔk¹  cups</td>
</tr>
<tr>
<td>ƙp⁵</td>
<td>op⁵ books</td>
</tr>
<tr>
<td>ƙl₃</td>
<td>lem³  long, slender objects, e.g. boats, pencils, fans</td>
</tr>
<tr>
<td>ƙ⁴</td>
<td>mau⁴  spoons</td>
</tr>
<tr>
<td>ƙɔu²</td>
<td>tau²  bottles</td>
</tr>
<tr>
<td>ƙ²</td>
<td>ku²  pairs, e.g. shoes</td>
</tr>
<tr>
<td>ƙ⁴</td>
<td>thyŋ²  trees, plants</td>
</tr>
<tr>
<td>ƙ</td>
<td>paa²  side</td>
</tr>
</tbody>
</table>
An overview of the classifiers in Shan-Ni is given in Table 27, though this overview is far from complete. Other Tai languages such as Shan Gyi, Aiton or Phake have around 50 classifiers, and the same can be expected for Shan-Ni. It might also be an open word class, accepting new members. For example $mɔk¹$ the classifier for flower, exists as the noun ‘flower’ in other Tai languages, but not as a classifier. The noun ‘flower’ itself in Shan-Ni is the disyllabic noun $mɔk'yaə²$. This process may be influenced by the possibility in Shan-Ni to classify nouns with $an¹$ ‘thing’, or the noun itself. For example, the noun $hɤn⁴$ ‘house’ can be classified in three ways (42). In (a) $hɤn⁴$ ‘house’ is classified by $hɤn⁴$ itself, in (b) by the general classifier $an²$ ‘thing’ and in (c) by $lɛm²$, the classifier for long objects. To my knowledge, there is no pragmatic difference between these three variations.

(42) a. $hɤn⁴ \ hɤn⁴ \ nɤng² \ hɤn⁴ \ an² \ nɤng²$
   house clf.house one       house clf.thing one
   ‘one house’

   b. $hɤn⁴ \ an² \ nɤng²$
   house clf.thing one
   ‘one house’

c. $hɤn⁴ \ lɛm² \ nɤng²$
   house clf.long one
   ‘one house’

(MK_LT_08)

4.2 Copula

Like other Tai languages, Shan-Ni has a copula $pen⁴/ pyin⁴/ pyen⁴$, but its usage is quite limited. It is used in specific constructions like the superlative (see the section 5.1 on $an²$) and otherwise as the verb ‘happen’. Cognates of $pen⁴$ in other Tai languages usually mean ‘be/become’ and are used for equation and proper inclusion. In this kind of sentences Shan-Ni omits the copula and the subject is juxtaposed with it predicate. For example in (43), the personal pronoun $kau¹$ ‘1sg’ and the name ‘Sai Aik’ suffice to make the phrase ‘I’m Sai Aik’. In (44) $kɔng²kaang⁴$ ‘chin’ and $yaaw⁴$, ‘long’ are also juxtaposed without a copula, with $yaaw⁴$ functioning as an attributive verb. In (45) $yaaw⁴yaaw⁴$ ‘long’ is reduplicated, and functions as an adjective instead. The distinction between $yaaw⁴$ ‘long’ as an attributive verb and $yaaw⁴yaaw⁴$ is thus made by republication, and does not require a copula.

(43) $kau¹ \ caai⁴aai²$
   1sg     sai  aik
   ‘I’m Sai Aik’ (first-born son)

(BSS_14)
When talking about someone’s occupation, Shan-Ni uses the phrase \( \text{het}^1 \text{ long}^2 X \) ‘do work X’. Both (46) and (47) are answers to the same question \( \text{meu}^4 \ \text{hit}^1 \text{ long}^2 \text{ sang}^4 \) ‘what kind of work do you do?’. Though normally \( \text{het}^1 \) ‘do’ is a verb, and \( \text{long}^2 \) ‘work’ is a noun, in this construction we could say \( \text{het}^1 \text{ long}^2 \) ‘do work’ as a single verb. In (46) it is followed by another verb \( t^4 \) ‘weave’ and a noun \( \text{huk}^1 \) ‘loom’, making the sentence literally ‘I do work weave loom’. (47) omits the subject, and includes two jobs: farmer and driver. For the farmwork, only \( \text{naa}^4 \) ‘land’ is placed behind \( \text{het}^1 \text{ long}^2 \). In the second part of the sentence the noun \( \text{kaa}^4 \) ‘car’ is marked by \( \text{kɔ}^5 \) ‘also’, thus in this case the verb \( \text{hɔ}^2 \) ‘drive’ comes behind the noun.

When talking about someone’s occupation, Shan-Ni uses the phrase \( \text{het}^1 \text{ long}^2 X \) ‘do work X’. Both (46) and (47) are answers to the same question \( \text{meu}^4 \ \text{hit}^1 \text{ long}^2 \text{ sang}^4 \) ‘what kind of work do you do?’. Though normally \( \text{het}^1 \) ‘do’ is a verb, and \( \text{long}^2 \) ‘work’ is a noun, in this construction we could say \( \text{het}^1 \text{ long}^2 \) ‘do work’ as a single verb. In (46) it is followed by another verb \( t^4 \) ‘weave’ and a noun \( \text{huk}^1 \) ‘loom’, making the sentence literally ‘I do work weave loom’. (47) omits the subject, and includes two jobs: farmer and driver. For the farmwork, only \( \text{naa}^4 \) ‘land’ is placed behind \( \text{het}^1 \text{ long}^2 \). In the second part of the sentence the noun \( \text{kaa}^4 \) ‘car’ is marked by \( \text{kɔ}^5 \) ‘also’, thus in this case the verb \( \text{hɔ}^2 \) ‘drive’ comes behind the noun.

In some contexts, the copula \( \text{pen}^4 \) is still used. In (48) \( \text{pen}^4 \) is used to say that the water got very high. This is what Me Mai sees at the end of the story, when she looks back from the mountain of Lwemon. It is used in combination with the TAM markers \( \text{kaa}^5 \) and \( \text{waj}^5 \), with \( \text{kaa}^5 \) marking that the action is finished, and \( \text{waj}^5 \) marking that this is something that was set in motion, hence the water had finished getting very high. Earlier in the story, (49) shows that \( \text{pen}^4 \) is not used when the water is still rising. To express the concept ‘become’ Shan-Ni uses the TAM marker \( \text{maa}^4 \), which is explained in Chapter 6.
Below are three sentences with *pyin⁴*, which follow each other as part of the origin story of how the Shan-Ni came to their current homeland. In (50), *pyin⁴* is used as 'happen', with the literal translation of the sentence being 'fever happened'. The following two sentences (51) and (52) start with *pyin⁴ sang⁴ (waj⁵)* 'what happened was that...', introducing an explanation into the story. When *sang⁴* follows a verb, the sentence turns into a question, hence *pyin⁴ sang⁴* means 'what happened?'. When the linker *si³/yi³* 'that because' follows *pyin⁴ sang⁴*, this becomes 'what happened was that...'.

(50) KHAI² NAU³ PYIN⁴ MAA⁴ KAA⁵
sick cold COP COME GO
'He got a fever'

(51) PYIN⁴ SANG⁴ WAJ⁵ YI³ ON² MAA⁴ TAA² MAA⁴ TAA²
happen what KEEP LNK weak COME PRT COME PRT
'What happened was that he got weaker and weaker.'

(52) PYIN⁴ SANG⁴ SI³ NYAA⁴ CAW² PUNG¹NAI⁴
happen what LNK meet leader brahmin
'What happened was that he met an astrologer.'

The combination with *sang⁴* 'what' is by far the most common usage of *pen⁴* in Shan-Ni. By itself, it is a question *pen⁴ sang⁴* 'what happened?' which could be answered by *sang⁴ ma³ pen⁴* 'nothing happened'. As shown above, followed by *si³ LNK*, the phrase can be used in a declarative sense, following an explanation of what happened. In some constructions, *si³* is not necessary, for example when followed by a negated clause initiated by *ma¹*. In (53) this
is illustrated with the sentence 'I cannot say what happened to his slipper'. In (54) a husband comes home, finds his wife's blood on the floor, which makes him wonder what could have happened. Here \textit{pyen}^{4}\textit{sang}^{4} is preceded by \textit{te}^{1}\textit{laj}^{2}, resulting in the question 'what could have happened'.

\begin{align*}
\text{(53)} & \text{yaa}^{1}\text{ khọ}^{2}\text{tìn}^{4}\text{mun}^{4}\text{pen}^{4}\text{\textit{sang}^{4}}\text{ ma}^{1}\text{ waa}^{2}\text{ pe}^{5} \\
& \text{place slipper 3SG COP what NEG say able} \\
& \text{`}I cannot say what happened to his slipper.' \text{ (BSS\textunderscore 90)}
\end{align*}

\begin{align*}
\text{(54)} & \text{lot}^{5}\text{ pa}^{1}\text{ying}^{4}\text{ kẹu}^{4}\text{ te}^{1}\text{ \textit{laj}^{2} \textit{pyen}^{4} \textit{sang}^{4}}^{4}\text{ i}^{3} \\
& \text{blood wife 1SG PUT GET COP what LNK} \\
& \text{`}That's my wife's blood – what could have happened?‘ \text{ (EP\textunderscore NU\textunderscore 31)}
\end{align*}

\subsection*{4.3 Existential and possessive predicates with \textit{yang}\textsuperscript{4}}

The existential verb in Shan-Ni is \textit{yang}\textsuperscript{4}. As an existential verb, it functions as a normal verb, and does not require any specific particles. It can be used to refer to things that are physically present in a place, like jade inside a mountain (55), or to say that there were a grandmother a grandfather (56), or that the situation referred to was a long time ago (57). All TAM markers can be used with \textit{yang}\textsuperscript{4}.

\begin{align*}
\text{(55)} & \text{sẹn}^{3}\text{ khẹw}^{3}\text{naj}^{5}\text{ \textit{yang}^{4}}\text{ sẹn}^{3}\text{ khẹw}^{3}\text{naj}^{5}\text{ khaa}^{3} \\
& \text{gem green DEF EXIST gem green DEF search} \\
& \text{`}There is jade, search for the jade.' \text{ (NN\textunderscore TH\textunderscore 20)}
\end{align*}

\begin{align*}
\text{(56)} & \text{pu}^{1}\text{ ta}^{1}\text{ yaa}^{2}\text{ sẹn}^{3}\text{ kọ}^{5}\text{ \textit{yang}^{4}}^{4}\text{ a}^{1} \\
& \text{grandfather and grandmother two CLF people EXIST PRT} \\
& \text{`}There were a grandfather and a grandmother.' \text{ (NN\textunderscore ITL\textunderscore 3)}
\end{align*}

\begin{align*}
\text{(57)} & \text{mọ}^{2}\text{ on}^{4}\text{taang}^{4}\text{taang}^{4}\text{ nang}^{4}\text{tẹ}^{5}\text{tẹ}^{5}\text{ \textit{yang}^{4}}^{4}\text{ maa}^{4} \\
& \text{time long ago really EXIST COME} \\
& \text{`}A really long time ago...‘ \text{ (NN\textunderscore TH\textunderscore 01)}
\end{align*}

An existential predicate can be initiated with \textit{yaa}^{1}\textit{ ti}^{2}\textit{ X} ‘at place X’ to refer the place where the subject ‘exists’. The same construction can also be used to express possession. In the two
sentences below the same construction is used, but in (58) $yaa^1 \ t^2...yang^4$ is used to describe
the existence of a tree at a house, and in (59) $yaa^1 \ t^2...yang^4$ is used to refer to the teacher
having two flowers.

(58) $yaa^1 \ t^2\ hɤn^4\ ton^2\ ma^1kyɛng^4\ yang^4\ khaa^2\ c^4$
    at place house tree tamarind EXIST POL PRT
    ‘At home there’s a tamarind tree.’

(59) $yaa^1\ ti^2\ mɔ³sɔn³mɔk¹yaa²\ sɔng³\ mɔk\ yang⁴\ u¹$
    at place teacher flower two CLF.flower EXIST STAY
    ‘The teacher has two flowers’

Instead of $yaa^1 \ t^2$, the predicate can also start with only $t^2$ ‘place’ (60), only $yaa^1$ ‘at’ (61),
(62) or $nyaa^4$ ‘MEET’ (63), which all have the same function in this position. These variants are
used interchangeably depending on personal preference, but villages or cities are usually
preceded by $nyaa^4$. This usage of $nyaa^4$ ‘MEET’ is striking because it is used as ‘place’ but it is
a verb, and verbs are strange in this position. It is does not have a similar function in other
Tai languages. In Burmese ‘place’ is $neya$, which is pronounced in a similar way. This might
be influencing the usage of the $nyaa^4$ in this position in Shan-Ni. Whenever $nyaa^4$ is used to
refer to a place rather than used as a verb, it will be glossed in small caps.

(60) $ti^2\ məhɤ³pu¹ta¹ya²\ sɔng³\ kɔ⁵\ lɛɯ⁴\ luk⁵sao³\ sip¹\ kɔ⁵\ yang⁴\ nɛ²$
    at 3PL grandparents two CLF.people at child.F seven CLF.person EXIST DEF
    ‘The grandparents had seven daughters.’

(61) $yaa^1\ mɯn⁴\ sɤ²\ phɯn³\ nɤng²\ yang⁴\ u¹$
    at 3SG shirt CLF.cloth one EXIST STAY
    ‘She has one shirt.’

(62) $yaa^1\ kaw⁴\ maak¹\ cɔk¹\ saam³\ hoj¹\ yang⁴\ u¹$
    at 1SG fruit orange 3 CLF.round EXIST STAY
    ‘I have three oranges.’
MEET 1SG time yesterday three hundred EXIST now DEF five thousand EXIST
'I had 300 yesterday, now I have 5000.'

For Phake, Morey (2005a: 310) describes only the option of having ti²...yang⁴ or only yang⁴ by itself. According to him, the difference in Phake is that ti² ‘place’ is used for alienable possession, while in predicates with inalienable possession ti² is not used. For Shan-Ni that does not work, because in (60) ti² does occur in front of the grandparents who have seven daughters, but in (63) only the first part of the sentence is between nyaa⁴...yang⁴, the while second predicate has yang⁴ without ti²/nyaa⁴/ya¹/ya¹ ti². The difference is not alienability, but whether the subject (possessor) is dropped or not. Since ti²/nyaa⁴/ya¹/ya¹ ti² precedes a subject, it will not be there either when there is no subject.

For things that are not there, yang⁴ is used in combination with ma¹ ‘NEG’. In (67) it refers to the lack of money, in (68) the frog not being there anymore, in (69) there not being any food, and in (70) it is used within the construction mx² me² ma¹ yang⁴ kaa² naj⁵ ‘when the mother is not there anymore’.

(BSS_36)

(BSS_30)

(BSS_60)
nɤng² su¹ mau¹ hrat² maa⁴
one buy new again COME

‘Look at him, he said he has no money but he bought a new car again.’

(phv_emp_18)

phaa⁵ leng⁴ maa⁴ kaa⁵ man⁴ luk⁵ maa⁴ kwaa¹ tɔj⁴ khyet¹ ma¹ yang⁴ kaa⁵
sky shine COME GO 3SG get up come go see frog NEG EXIST GO

‘When the sun is up he gets up and checks, the frog is not there anymore.’

(KSN_FS_07)

caa⁴ kyin⁴ caa⁴ yam⁵ kɔ⁵ ma¹ yang⁴ kaa⁵ yaap¹ maa⁴
for eat for chew PRT NEG EXIST GO difficult COME

‘There was nothing to eat or to chew anymore, it became difficult.’

(NN_ITL_11)

mɤ² mɛ² ma¹ yang⁴ kaa⁵ naj⁵ pɔ² naj⁵ mɛ² kɔ⁵ nɤng² jaw⁵
time mother NEG EXIST GO DEF father DEF mother CLF human one FINISH

‘When there is no mother anymore, the father is already a mother.’

(BSS_78)

In most of these examples yang⁴ is followed by the TAM marker kaa⁵ ‘go’. When this TAM marker is used the meaning becomes ‘not anymore’, hence referring to things that have been there, but are gone now. The difference is illustrated by the pair below (71). Without kaa⁵, the phrase sang⁴ ma¹ yang⁴ means ‘there’s nothing’. With kaa⁵, the phrase sang⁴ ma¹ yang⁴ kaa⁵ means ‘there isn’t anything anymore’, which is used for example when a restaurant runs out of a favorite dish at the end of the day.

(71)

a. sang⁴ ma¹ yang⁴
b. sang⁴ ma¹ yang⁴ kaa⁵
what NEG EXIST what NEG EXIST GO
‘There’s nothing.’ ‘There isn’t anything anymore.’

(MK_LT_08)

4.4 Questions and indefinite pronouns

Question words used in content questions and indefinite pronouns are expressed with the same words in different positions. Content questions are formed by placing the question word at the end of the sentence, possibly followed by a TAM marker. These same words at the beginning of a predicate function as indefinite pronouns. Table (28) provides an overview of all the question words and indefinite pronouns in Shan-Ni.
Table 28 Question words and indefinite pronouns

<table>
<thead>
<tr>
<th>Shan-Ni word</th>
<th>Question word</th>
<th>Indefinite pronoun</th>
<th>Indefinite pron. + NEG ma¹</th>
</tr>
</thead>
<tbody>
<tr>
<td>mɤ² le²⁴</td>
<td>‘yes or no?’</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>sang⁴</td>
<td>‘what’</td>
<td>‘something’</td>
<td>‘nothing’</td>
</tr>
<tr>
<td>phew⁴</td>
<td>‘who’</td>
<td>‘whoever’</td>
<td>‘nobody’</td>
</tr>
<tr>
<td>thɛɯ⁴</td>
<td>‘where’</td>
<td>‘anywhere’</td>
<td>‘nowhere’</td>
</tr>
<tr>
<td>phew⁴</td>
<td>‘when’</td>
<td>‘whenever’</td>
<td>‘never’</td>
</tr>
<tr>
<td>mɤ² thɛɯ⁴</td>
<td>‘when’</td>
<td>‘whenever’</td>
<td>‘never’</td>
</tr>
<tr>
<td>nahu⁴</td>
<td>‘how’</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>kɔp⁵sang⁴</td>
<td>‘why’</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>kha¹lɛɯ⁴</td>
<td>‘how much’</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>lɛɯ⁴</td>
<td>‘which’</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

The question words sang⁴, phew⁴ and thɛɯ⁴ always come at the end of the question. In (72) the question phɯn³lik⁵ phew⁴ ‘whose book?’ only exists of the word phɯn³lik⁵ ‘book’ and the question word phew⁴ ‘who’. In (73) thɛɯ⁴ ‘where’ follows the serial verb construction kwaa¹ laj² khaa³ ‘go get search’ with preverbal TAM marking te¹. In (74) the noun phak¹ ‘curry’ precedes sang⁴ ‘what’ to ask about the type of curry that will be cooked.

(72) phɯn³lik⁵ phew⁴
    book who
    ‘Whose book is that?’

(BSS_63)

(73) te¹ kwaa¹ laj² khaa³ thɛɯ⁴
    WILL go get searchwhere
    ‘Where can [I] go look for it?’

(BSS_34)

(74) mɛɯ⁴ te¹ tang² phak¹ sang⁴
    2SG FUT cook curry what
    ‘What curry will you cook?’

(BSS_116)

When these same words are used at the beginning of the sentence, they are indefinite pronouns. In (75), phew⁴ ‘at the beginning on the sentence refers to ‘anybody’, and in (76) thew⁴ refers to ‘anywhere’, which at the end of the sentence would be ‘who’ and ‘where’ respectively. Sometimes instead of sang⁴ ‘something’, ka¹sang⁴ (77) is used, though sang⁴
itself is also used (78). In all these sentences, the indefinite pronoun comes at the beginning of the sentence, usually followed by the particle \( kɔ^5 \). Usually, whether \( kɔ^5 \) is included or not makes no difference because the location already clarifies the meaning of the indefinite pronoun versus the interrogative. For \( mx^2 \ lexu^4 \) 'anytime' it does matter whether \( kɔ^5 \) is included, because \( mx^2 \ lexu^4 \) 'when' can also be an interrogative at the beginning of the sentence, and \( kɔ^5 \) removes this ambiguity. In (79) \( mx^2 \ lexu^4 \) with \( kɔ^5 \) thus refers to 'anytime', while later in (82) \( mx^2 \ lexu^4 \) appears without \( kɔ^5 \) at the beginning of the sentence, meaning 'when'.

\[
(75) \text{phɛɯ⁴ kɔ⁵ ma¹ het¹ leh¹} \\
\text{anybody PRT NEG do also.BM} \\
\text{'Nobody cares.'} \\
\text{(MK_IH_14)}
\]

\[
(76) \text{thɛɯ⁴ ma¹ kwaa¹} \\
\text{anywhere NEG go} \\
\text{'I'm not going anywhere.'} \\
\text{(MK_LT_08)}
\]

\[
(77) \text{ka'sang⁴ kɔ⁵ ma¹ ko⁴} \\
\text{anything PRT NEG afraid} \\
\text{'not afraid at all'} \\
\text{(PV_EMP_12)}
\]

\[
(78) \text{sang⁴ ma¹ yang⁴ kaa⁵} \\
\text{anything NEG EXIST GO} \\
\text{‘There isn’t anything anymore.’} \\
\text{(MK_LT_08)}
\]

\[
(79) \text{mx⁴ lexu⁴ kɔ⁵ pi¹ maa⁴} \\
\text{time INDF PRT PROH come} \\
\text{‘Never come.’} \\
\text{(BSS_24)}
\]

The post-verbal TAM markers \( kaa^5 \) and \( maa^4 \) both come after the question word. (80) and (81) contain almost the same questions, only in (80) \( maa^4 \) indicates that somebody started to cook something in the past and \( kaa^5 \) that somebody finished cooking in the past. For further discussion on \( kaa^5 \) and \( maa^4 \) Chapter 6.
The question words *kha¹leu⁴* ‘how much’ and *mɤ² thɛw⁴/lɛɯ⁴* ‘when’ do not need to come at the end of a sentence. In this case *leu⁴* already follows what it modifies, for example *mɤ²* ‘time’, making *mɤ² leu⁴* literally ‘which time’. In (82) *mɤ² leu⁴* comes at the beginning of the sentence, while in (83) it comes at the end. In both cases the other way around would be acceptable as well, the emphasis would just be different. In (84) the verb *thɤng³* ‘arrive’ and the TAM marker *maa⁴* follow the question word *mɤ² thɛw⁴*, because the TAM marker *maa⁴* should come at the end, and this way the verb and the TAM marker do not have to be split up.

A similar situation in which *leu⁴* already modifies what it follows applies to *kha¹leu⁴* ‘how much’. The syllable *kha¹* is already fossilized and attached to *leu⁴* ‘which’ in *kha¹leu⁴*, but from Tai Mau (Young 1985: 27) we can learn that originally *kha¹* is a reduction from *khai* (‘how much’ in Tai Mau). This is the result of a contraction in Tai Mau is merger from *kaa*. 
'price' and *hi* 'or not', making *kha¹leu⁴* 'which price'. For *kha¹leu⁴* 'how much' it is to have a verb such as *yang⁴* 'have' (85) or *laj²* 'get' (86) following it.

(85) a’saaƙ⁵ mou⁴ kha¹leu⁴ yang⁴ khaa⁵
   age 2SG how.much have POL
   ‘How old are you?’

(86) kaa² man⁴ kha¹leu⁴ laj²
   salary 3SG how.much get
   ‘How much salary does he get?’

Shan-Ni uses *kha¹leu⁴* both for 'how much' and 'how many', with 'how many' requiring an extra word specifying the type of thing asked about. In Khamti, Phake and Aiton only the interrogative for 'how much' is similar to Shan-Ni's *kha¹leu⁴*, for 'how many' they use a word similar to *kii*. In Shan-Ni, when the referent itself is included in the question, the word to specify the type of object asked about is often a classifier, such as *kɔ⁵* 'CLF. human' (87) or *hu⁴* 'CLF. hole' (89) in the examples below. But in (88), the question is not about *khɯn²* 'attend' or the dropped subject 'school', but about the time, hence *mɔng⁴* 'hour' is used. If instead of *mɔng⁴* a classifier for school would be used, the question would be 'how many schools attend?' (e.g. in a sports tournament). In the same way, in (89) the question is not about the quantity of stairs, it is about the quantity of places for stairs. This reflects the fact that stairs are not necessarily fixed, so the quantity of 'stairs' and 'places for stairs' do not have to be the same.

(87) pi²mɔng⁵ yang⁴ kha¹leu⁴ kɔ⁵
   siblings EXIST how.many people
   ‘How many siblings do you have?’

(88) khɯn² kha¹leu⁴ mɔng⁴
   attend how.many hour
   ‘What time does [school] start?’

(89) hung⁴kha¹laaj³ kha¹leu⁴ ti² yang⁴
   stairs how.many place EXIST
   ‘How many stairs (places) are there?’
For polar questions, Shan-Ni uses the interrogative particle ละ le⁴ (or ละ lɛ²) at the end of the phrase. It is unique to Shan-Ni, and does not seem to have an equivalent cognate in other Tai languages. It is possible that it is cognate with the question word nai¹ in Khamti, though in Khamti it is also used in combination with content questions (Inglis 2014: 112). It would not be strange, since /n/ and /l/ are allophones in Shan-Ni for words that often have an initial /n/ in Khamti. In these questions the speaker aims at finding out whether something is true or not, for example in (91) whether the interlocutor is free tomorrow, or in (92) whether the item is bought already or in (93) whether the interlocutor would like to drink sweet tea.

(91) မော်ပါက လူများ လီများ
      mɤ² phɔk⁵    laap⁵ le⁴
time tomorrow free QP
‘Are you free tomorrow?’

(92) စောင် ဝယ် လာ
      sɯ⁵ yaw⁵  le²
buy FINISHED QP
‘Did you buy it already?’

(93) မော်မတွေ့ အရှည် အရှည် ငါလာ
      nam⁵ leng² waan³ te¹  sot⁵  le⁴
water tea sweet FUT drink QP
‘Would you like some (sweet) tea?’

Alternatively, le⁴ or le² can be replaced by နား nɔ³ or နား naa³ when the answer is already implied. This is the same in Aiton and Phake (Morey 2005a: 350). နား (95) and နား (96) are tag questions, meaning something similar to ‘isn’t it?’ or ‘right?’.

(94) မြင်လုံ ပါ ဟာကြား
      haw⁴ te¹ khyi¹ hɔ⁵ men⁴ naa⁵
1PL FUT ride boat fly TAG
‘We’re going by airplane, isn’t it?’
Month face Fut come meet Tag Fut meet talk VPL Stay

‘Next month you will come meet [me] right? [We] will speak to each other.’

(95)

(96)

(97)

4.5 Passive

Shan-Ni can make passive-like sentences with laj² khaam² ‘get crossed’, literally indicating that a subject got “crossed” by some outside force. In (96) the subject was bitten by a dog, and in (97) a daughter was expelled from school. Verbs made passive in these constructions are nominalized by an² ‘thing’. For example, in (96) kaap⁵ ‘bite’ (verb) is nominalized to an² kaap⁵ ‘bite’ (noun), and then followed by laj² khaam² ‘got crossed’, meaning that the subject ‘got crossed’ by a bite. In (97) an² nominalizes ha⁴ ɔk¹ ‘made to exit’ becoming an² ha⁴ ɔk¹ ‘expulsion’, hence in this phrase, the daughter ‘got crossed’ by expulsion. This construction does not exist in any other Tai or Tibeto-Burman language in the region.

4.6 Conditional

Conditional sentences in Shan-Ni can be made by with sang³ pi⁴, pi⁴ or (sang⁴) pɯ⁴. I did not find any distinctions in their usage, though sang³ pi⁴ seems to be the most common variant in spoken language. The conditional is followed by a verb, such as tang² ‘cook’ in (98) or han³ ‘see’ in (100). In both of these sentences the subject is omitted, but when this is not the case, as in (99) with nam⁵ ‘water’ and in (101) with lom⁴ ‘wind’, the subject precedes the conditional marker sang⁴ pɯ⁴. In (102) the adverb later also precedes the conditional pɯ⁴.
(98) sang³pi⁴ tang² khaw² nam⁵ pi⁴ hau⁴ nam³
if cook rice water PROH give much
‘If [you] cook rice, don’t put too much water.’

(99) nam⁵ sang³pi⁴ nam³ ᵇ⁶paap⁵paap⁵ si³ mo³kyin⁴ li⁴
water if many soft.RDP LNK NEG eat good
‘If [you add] too much water it will be soft and not tasty.’

(100) pi⁴ han³ aw⁴ pɔk⁵ maa⁴ kɤn⁴
if see take return come SHOULD
‘If [you] see [it], [you] should bring it back.’

(101) lom⁴ sang⁴pɯ⁴ hɛng⁴ sɤ² phaa²  sing⁴ waj⁵  ni⁴ni⁴  naa⁵ caang²
wind if strong shirt cloth tighten keep good.RDP TAG know

(102) kha¹nang³ pɯ⁴ thop¹ kɔ⁵ hɔng⁵ khaa² naa³
later if meet PRT call POL isn’t.it
‘If [you] find it back later, please call me’

4.7 Dependent clauses with an² ‘thing’, mɤ² ‘time’ and ti² ‘place’
Dependent clauses are usually initiated by an² [the thing that...], ti² [the place that] or mɤ² [the time that] and are closed by the definite marker naj⁵/ne². Sentences with an² ‘thing’ might be analyzed as relative clauses, while this would be less suitable for sentences with mɤ² ‘time’ and ti² ‘place’. Because the usage and structure of these sentences in Shan-Ni is identical, I will discuss them together. an² may or may not be preceded by a head noun specifying the ‘thing’ that is talked about. (103) includes the head noun kaa⁴ ‘car’, specifying that the thing the interlocuter had talked about was a car. In (104) lɔng² ‘work’
precedes \textit{an²}, specifying that the thing that was ordered was work. These sentences would still be correct when removing the headnoun.

(103) \textit{kaa⁴ an² mɛɯ⁴ waa² naj⁵ su⁵ kaa⁵ de⁴}
\textit{car thing 2SG say DEF buy GO Q}
\[H\] \[DC\]
‘Did you buy the car you were talking about?’

(104) \textit{ma¹su³ pi²nɔng⁵ lɔng² an² kaw⁴ cau⁴waj⁵ naa⁵}
\textit{2PL sisters work thing 1SG order KEEP PRT}
\[H\] \[DC\]
‘You sisters, the work I ordered you...’

Without a head noun, not much changes in the sentence, it is just that the type of thing is not specified by a noun preceding it, like above with ‘car’ and ‘work’. In (105) the relative clause is ‘the thing we wanted to do’ and in (106) ‘thing(s) that give strength’. In (105) the preceding sentences are about gold mining, so it is clear from the context that the ‘thing’ refered to is the activity gold mining. In (106) the interlocutor is advised to eat ‘things that give strength’, so it is clear this refers to food. Similar to the ellipsis discussed at the beginning of this chapter where already obvious subjects are omitted, obvious head nouns can also be omitted.

(105) \textit{an² khaj² het¹ naj⁵ tang⁴long³ het¹ si³ laj²}
\textit{thing want do DEF everyone do LKN GET}
\[DC\]
‘the thing [we] wanted to do, everybody can do.’

(106) \textit{an² te¹ yang⁴ heng⁴ naj⁵ kyin⁴}
\textit{thing FUT EXIST strength DEF eat}
\[DC\]
‘Eat things that give strength.’

This could be seen as posthead (102, 103) and internally headed relative clauses (104, 105), but that would suggest two different types of relative clauses. Since this is again about information structure in which a known and active argument is omitted, it should not be read that way. The head noun is there, either literally, or in the understanding of the speaker and
listener. If it is there, it precedes the relative clause. If it would be seen as and internally headed relative clause, the headnoun would be an². Since an² is included regardless of the presence of an external headnoun, that analysis does not work for Shan-Ni.

Temporal dependent clauses start with ms² ‘time’ and end in the definite marker naj⁵, similar to relative clauses with an² thing’. In (106) ms² meu⁴ thɤng³ naj⁵ refers to ‘the time you arrived’, and is followed a direct question about what happened in the recent past. (107) can be read in a more conditional way, with ms² me² ma¹ yang⁴ kaa⁵ naj⁵ ‘the time mother isn’t there anymore’ referring to any situation in which the mother is not present anymore.

(107) ms² meu⁴ thɤng³ naj⁵ pi² cai⁴ meu⁴ thɤng³ kaa⁵ de⁴
  time 2SG arrive DEF brother 2SG arrive GO QP
  [DC ]
  ‘When you arrived, had your brother arrived already?’

(108) ms² me² ma¹ yang⁴ kaa⁵ naj⁵ pɔ² naj⁵ me² kɔ⁵ nɤng² jaw⁵
  time mother NEG exist GO DEF father DEF mother CLF-human one FINISH
  [DC ]
  ‘When there is no mother anymore, the father is already a mother.’

Any noun, noun phrase or clause about physical places (including ‘at’ people) starts with ti², yaa¹, nyaa⁴ or yaa¹ ti². Similar to an² and ms², spatial referents can head relative clauses. In (109) it used in front of a classifier structure, ending with the demonstrative leu⁴ ‘at’ rather than naj⁵. In (110) naj⁵ is also not used, but it is clear that nyaa⁴ lau⁴ hɤn⁴ haw⁴ ‘in our household’ is a relative clause because otherwise the personal pronoun kaɯ⁴ ‘1SG’ would have to be a possessive pronoun instead.

(109) ti² mahr² pu¹ ta² ya² song⁵ ko⁵ leu⁴ luk⁵ sao³ sip¹ ko⁵ yang⁴ ne²
  place 3PL grandparents two CLF-people at child.F 7 CLF-person exist DEF
  [DC ]
  ‘The grandparents had seven daughters.’

(110) nyaa⁴ lau⁴ hɤn⁴ haw⁴ kau⁴ an² jau¹ ji³ pen⁴
  MEET inside house 1PL 1SG thing big LNK be
  [DC ]
  ‘In our household I’m the oldest.’

80
Chapter 5: Grammatical markers

Apart from the different types of sentences discussed in Chapter 4, there are also a number of words that have developed into grammatical markers, each with their own unique function. Some of these have developed from regular words, whereas others also exist as grammatical markers in other Tai languages but often are slightly different or have additional functions in Shan-Ni. Some of these markers may have quite different functions depending on their context, and in general this relates to two patterns. Firstly, they all have some relation to the meaning of the word they derived from, and secondly, they often mirror Tibeto-Burman grammatical markers in different ways. To illustrate this, I will point out the correspondences these markers have with Burmese. Nevertheless, as elsewhere, it has to be taken into account that Shan-Ni is and has for a long time been in contact with a large number of other neighbouring Tibeto-Burman languages, hence similarities may also be due to contact between Shan-Ni and these languages. Of course the fact that interviews went through Burmese facilitates this connection, but all these markers are common in daily oral communication and non-elicited texts as well. It is striking that each of these markers facilitates Shan-Ni/ Burmese bilinguals to communicate the same nuances through the different languages in their repertoire, though none of them are direct lexical borrowings, nor translations of the Burmese equivalents.

5.1 The nominalizer an² ‘thing’

Noun phrases are often headed by an² ‘thing’, referring to ‘something that is X’. It can precede words from a variety of word classes, modifying them into a noun (phrase). Followed by only a verb, an² can form a new noun. The nouns for ‘food’, ‘death’, ‘love’ and ‘dance’ (111) are formed that way. This way of forming new nouns is quite productive, for example with an²kaap⁵ ‘bite’ (112) derived from the verb ‘bite’ and an²haaj² ‘cry’ (113) derived from the verb ‘cry’.

(111) a.  an²  kyin⁴  b.  an²  taaj⁴  c.  an²  hak⁵  d.  an²  kaa²
     thing   eat       thing   die      thing   love    thing   dance
     ‘food’   ‘death’   ‘love’   ‘dance’

(112) kau⁴  maa³  an²  kaap⁵  laj²  khaam²  kaa⁵
     1SG  dog    thing bite get   CROSS  GO
     ‘I was bitten by a dog.’
an² can also be used with attributive verbs followed by the linker si³ and the verb pen⁴ 'be' to make superlative constructions. For example, to speak about the oldest (115), the smallest (116) or the most most beautiful (117).

(113) an² haaj² ka⁵ yin⁵ kan⁴ kwaa⁴ yaw⁵
thing cry PRT hear VPL go FINISHED
'you hear the cry [slowly] ceasing'

(114) nyaa⁴ lau⁴ hən⁴ haw⁴ kau⁴ an² jau³ ji³ pen⁴
MEET inside house 1PL 1SG thing big LNK be
'I'm the oldest in our house.'

(115) nang¹ nɛ² an² ɔn¹ ji³ pen⁴
this DEF thing small LNK be
'This is the smallest.'

(116) an² saan⁴ ni³ pen⁴
thing beautiful LNK be
'the most beautiful'

With demonstratives an² can refer to objects specifying the distance, as in 'that/this thing (over there)'. This construction may or may not include the definite particle naj⁵/ne⁴/naa⁵, which is obligatory when an² introduces a full noun phrase. In (118) an² and nang¹/ne² 'this' together form 'this thing', in (119) an² and pun⁴/naj⁵ together form 'that thing' and in (120) an² and pun⁴/than² refers to 'that thing over there'.

(117) an² nang¹/ne² kaa² aaj⁴te⁴ka⁵ mɔ³sɔn³
thing this car friend teacher
'This is my friend's teacher's car.'
Lit. This thing is my friend's teacher's car

(PV_EMP_16)
(118) အန် ပြင် နေ ဟု လေး လုပ်
an² pun⁴ nai⁵ hu⁵ ma¹ cau² le⁴
ing there this know NEG true QP
'You know that thing, right?'

(BSS_96)

(119) အန် ပြင် သို့ ဝေ မည် ကုန် ကုန်
an² pun⁴ than² aw⁴ maa⁴ kyaang¹kyaang¹
ing over.there take come fast.RDP
'Bring that thing over there quickly.'

(BSS_26)

Nouns can precede an² in this construction, specifying the type of 'thing' an² refers to. Hence in (121) an² naj⁵ 'this thing' is specified by hs⁵ 'meat' making it 'this meat (thing)'. In (122) the noun lɔng² comes before an² making it 'work thing', but here instead of naa⁵ following an² directly, there is a whole sentence in between an² and naa⁵.

(120) လာ အန် နေ ကုန် ကုန် ယုန် စီ နာ ရာ လုပ်
hs⁵ an² naj⁵ kau⁴ kwaa¹ su⁵ maa⁴ nan⁵ na¹hu⁴ ni⁴ yu¹ le⁴
good thing this 1SG go buy COME that how good STAY QP
'How good is this meat I bought?'

(BSS_96)

(121) မည် စီ ပြင် နေ လာ ကုန် ကုန် ဗျူ နာ
ma³su³ pi²nɔng⁵ lɔng² an² kau⁴ cau⁴ waj⁵ naa⁵
2PL sisters work thing 1SG order KEEP PRT
'You (sisters), the work I ordered you…'

(BSS_210)

When an² modifies a more complex noun phrase, the noun phrase needs to be closed with the particle naj⁵/ ne² to mark what is included in the noun phrase. Whereas above an² nominalizes verbs, in the examples below this is not the case. For example, an² khaj² het¹ naj⁵ is 'the thing [we] wanted to do' (122) and an² te¹ yang⁴ heng⁴ naj⁵ 'the thing that gives strength' (123).

(122) အန် ကူ ညှင် နေ
an² khaj² het¹ nai⁵
ting want do DEF
'the thing [we] wanted to do.'

(BSS_39)
5.2 The verbal plural marker *kan⁴*

The verbal plural marker *kan⁴* is used in Shan-Ni to refer to things that are done by multiple people, either reciprocally or simultaneously. The reciprocal function of *kan⁴* is common in other Tai languages, while the simultaneous function is not. In the examples below *kan⁴* is used reciprocally. In (124) *kan⁴* is used to show that this couple will continue quarrelling with each other their whole lives. Similarly, in (125) *kan⁴* is used in the second part to show that [we] will speak to each other.

(124) | CursorPosition | CursorPosition |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>pho³ ta¹ me⁴ sɔng³ kɔ⁵ caa⁵ caa⁵ thaa⁴ thaa⁴ u¹ kan⁴ tɔ¹ co²</td>
<td>husband.wife two CLF.human angry.RDP.swear.RDP STAY VPL until.life</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>'That couple will keep quarrelling their whole lives.'</td>
<td></td>
</tr>
</tbody>
</table>

(125) | CursorPosition | CursorPosition |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>lɤn⁴ laa² te¹ maa⁴ thop¹ nɔ³ te¹ thop¹ tɔ¹ kan⁴ yu¹</td>
<td>month face FUT come meet isn’t.it FUT meet talk VPL STAY</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>'Next month you will come meet [me] right? [We] will speak to each other.'</td>
<td></td>
</tr>
</tbody>
</table>

In most contexts, *kan⁴* is used to refer to situations in which actions are done by multiple persons at the same time. Though this has overlap with reciprocity, it is broader than that. Simultaneous actions only require the participation of multiple persons, but is does not require people to do a certain action "to each other". Both the speakers I worked with and the book Basic Spoken Shan-Ni consistently translate *kan⁴* as *kya* [ကော်], which in Burmese is used as a plural marker for verbs. Okell and Allott (2001:16) describe Burmese *kya* as a marker which "emphasizes (a) the mutuality of the action, or (b) the fact that several actors were engaged separately, severally or (c) (confusingly) that all actors acted together". The description "(a) the mutuality of the action" fits the reciprocal examples discussed above, but the examples below indeed fit the last two parts of the description better. In (126) two sisters are pounding rice the whole day, which is not reciprocal, but is a joint action. In (128) the parents are both worried about their child, not about each other, so also here it is simultaneous rather than reciprocal.
They are doing it [pounding rice] the whole day.

'We eat together.'

'[The parents] are also worried.'

'[You] can also drive along with the car.'

'The curry is also not good.'

5.3 The particle $kɔ^5$

The particle $kɔ^5$ ‘also’ is a homonym of the human classifier $kɔ^5$. They are easily confused, but their functions are quite different. Both words also occur as separate lexemes in Phake (Morey 2005a: 145). In most sentences $kɔ^5$ can be translated directly as ‘also’ after a verb (129) or noun (130). In (129) someone is offered a ride because the driver goes into the same direction, and (130) is a common phrase hosts say about their own food when they have guests.

Sometimes when $kɔ^5$ is used, it cannot be directly translated as ‘also’, nor does it really have a translation. Often $kɔ^5$ follows a clause, like in the examples below (131, 132, 133). Except for closing a clause it does not seem to have a specific function in this context. In Phake this
would mark a topicalized phrase, indicating that the core sentence is following (Morey 2005a: 248). This analysis does not entirely work for Shan-Ni, as in (133) it does follow a conditional clause but it is not necessarily topicalized.

(131) kha¹nang³ pu⁴ thop¹ ko⁵ hɔng⁵ khaa² nəa³
later if meet PRT call POL isn't it
'If you find it back later, please call me'

(132) an² man⁴ phaan³ nə² ko⁵ caa⁴ pa⁴ying⁴ man⁴ kha¹laj² nə² ma¹ laj²
thing 3SG poor DEF PRT for girl 3SG want DEF NEG GET

thop¹ nyaa⁴ pe⁵
encounter meet can
'Because of his poverty he cannot marry the girl he wants.'

(133) mə² hə² nai⁵ ko⁵ phun³ tok¹ sək⁵ sək⁵ si³ yik⁵ yik⁵ cək⁵ cək⁵
time monsoon DEF PRT rain fall trickle.RDP LNK moist.RDP trickle.RDP

moist.
trickle.
trickle.

'R in the monsoon season it rains a lot.'

Another function of ko⁵ is as an indefinite in combination with an interrogative, together forming an indefinite pronoun. In this construction the indefinite pronoun is at the beginning of the phrase, whereas when used for question it would usually come at the end. The examples below are all negated. For example, in (134) as ka¹sang⁴ ko⁵ refers to 'nothing', in (135) mə² leuw⁴ ko⁵ to 'never' and in (136) phew² ko⁵ to 'nobody'.

(134) ka¹sang⁴ ko⁵ ma¹ ko⁴
what PRT NEG afraid
'not afraid at all'

(135) mə² leuw⁴ ko⁵ pi¹ maa⁴
time which PRT PROH come
'never come.'
This function is not limited to interrogatives either; Shan-Ni can also use kɔ⁵ as an indefinite marker after a noun, for example below in (137) kaa² kɔ⁵ ‘any dance’. In (138) kɔ⁵ follows caa⁴ kyin⁴ caa⁴ yam⁵ ‘for eating, for chewing’, meaning that there was nothing to eat or to chew.

(137) ma¹ taang² kaa² kɔ⁵ kaa²
NEG way dance PRT dance
‘I also don’t know any dance.’

(138) caa⁴ kyin⁴ caa⁴ yam⁵ kɔ⁵ ma¹ yang⁴ kaa⁵ yaap¹ maa⁴
for eat for chew PRT NEG EXIST GO difficult COME
‘There was nothing to eat or to chew, it became difficult.’

5.4 The linker si³

The particle si³ is a linker used to connect two phrases. It can occur as si³, ni³, yi³ or i³. It has several functions, most of which have some overlap with Burmese lo [လိ့]. Depending on the context it could be translated as ‘that’, ‘and’ or ‘because’, or remain untranslated. The most basic usage of si³ is to link two verbs that are separate actions, rather than part of the same serial verb construction contributing to the same action. In (139) hɔng⁵ ‘call’ and taaj⁴ ‘shout’ are part of the same action, but simultaneously the children are also sitting, hence si³ is used to connect nang² ‘sit’ to hɔng⁵ taaj⁴ ‘(call) shout’. In (140) the boy is ‘faring a tree’ tm⁵maj⁵ khyi¹ and ‘climbing (going upstream) back’ khun² pok⁵ maa⁴ to where he came from. The two actions connected by yi³ show that he is doing both things simultaneously, though they still remain described as two different things. At the same time, khun² ‘climb’ and pok⁵ ‘return’ are not separated by another yi³, since they contribute to the same action.

(139) luk⁶on¹ nang² si³ hɔng⁵ taaj⁴ su¹ u¹
child sit LNK call shout towards STAY
‘The children are sitting and shouting at each other.’
When it links two verbs, \(si^3\) can follow both verbs to show multiple things happened in a certain time frame. In (141) \(wa^2 yi^3 \) ‘talk’ \(kap^1 yi^3 \) ‘communicate’ are both connected to \(te^4 pɔk^4 maa^4 kaa^5\) ‘was going back’ because the subject did both when he was going back. In (142) the astrologer explains the king that (a) he will get rid of the illness and (b) he will become better, which are both followed by \(si^3\), linking them both to \(waa^2 \) ‘say’.

\[
(140) \quad an^2 \ nyaa^4 \ ts^n^2 \ maj^5 \ khyi^1 \ yi^3 \ te^1 \ khun^2 \ pɔk^5 \ maa^4 \\
\text{thing} \ MEET \ tree \ ride, \ LNK \ FUT \ climb \ return \ COME
\]

(On the tree he faires back.)

(141)
\[
wa^2 yi^3 \ kap^1 yi^3 \ te^4 pɔk^4 maa^4 kaa^5 \ lstå^5 \ px^4 px^4 \\
talk \ LNK \ communicate \ LNK \ FUT \ return \ COME \ GO \ blood \ all \ over
\]

‘Talking and discussing [with his mother] he was going back, and the blood was all over.’

(142)
\[
yo^1 \ ga^1 \ te^1 \ haaj^3 \ si^3 \ te^1 \ li^4 \ maa^4 \ si^3 \ waa^2 \\
illness, BM FUT loose \ LNK \ FUT \ good \ come \ LNK \ say
\]

‘You will get rid of the illness and get better, he said.’

(143)
\[
khyet^1 to^4 \ nɔng^2 \ paaj^2 \ aw^4 \ yi^3 \ kwaa^1 \\
frog CLF.animal one run take LNK go
\]

‘The frog ran out.’

(144)
\[
tang^1 \ seik^1 \ maa^4 \ si^3 \ kwaa^1 \\
with goat, BM come LNK go
\]

‘With the goat he comes and goes.’
A very common expression with *si³* is *si³ma¹laj²*, the equivalent of ‘cannot’. It is used in combination with any verb, telling the interlocutor something cannot be done. For example in (147), the interlocutor is told (s)he cannot go to the pagoda festival today. This does not imply that the speaker is the one putting on the restriction, or that the interlocutor is not allowed to go. If the interlocutor would not be allowed to go, this would be expressed with *ma¹ hau⁴ kwaa¹*, as shown in (148). This construction also works without the negator *ma¹*, as shown in (149) with *khaw² si³ laj² yaw⁵* ‘can enter already’.

The linker *si³* is also part of the construction *si³ u¹* with the TAM marker *u¹* ‘stay’. This is discussed in more detail in Chapter 6.
5.5 The prohibitive $pi^1$

The prohibitive $pi^1$ ‘prohibits’ the action indicated by the verb it precedes. In (150) it precedes $taan^2$ $khwaam^4$ ‘talk’, in (151) $u^1$ ‘stay’, and in (152) $maa^4$ ‘come’. As it is about actions the speaker prohibits the listener to do, there is no TAM marking, and since it usually part of a conversation where the referent is clear, there will not be any subject expressed.

(150) $pi^1$ $taan^2$ $khwaam^4$ nam^3 nam^3
PROH talk language a.lot.RDP
‘Don’t talk a lot.’

(151) nyaa^4 thaj^2 $pi^1$ $u^1$ nyaa^4 nam^5 $nɔng^3$ long^3
MEET here PROH stay MEET water lake big
‘Don’t stay here, at Indawgi.’

(152) mɤ^4 lɛɯ^4 kɔ^5 $pi^1$ $maa^4$
time INDF PRT PROH come
‘Never come.’

(153) $pi^1$ $khaan^5$ aan^1 lik^5
PROH lazy count write
‘Don’t be lazy to study.’

(154) nang^1 nan^5 $pi^1$ het^1 nang^1 ne^2 het^1
that PROH do this do
‘Don’t do that, do this.’

5.6 Plurality

As already briefly mentioned in section 3.1.3, one of the grammatical functions of the no longer used personal pronoun $khaw^3$ ‘3pl’ is that of a plurality marker for humans. To indicate plurality, it is placed behind two important participants of the group it applies to, for example in (156) $pɔ^2$ $lɛŋ^5$ ‘biological father’ and $me^2$ $lɛŋ^5$ ‘biological mother’ are followed
by khaw³ to refer to the concept ‘biological parents’. The combination of pu¹ ‘grandfather’ and yaa² ‘grandmother’ refers to old people in general. Any prepositions or morphology can be marked on both elements of the pair, like the reduced a¹ (from ang¹ ‘for’) preceding both pu¹ ‘grandfather’ and yaa² ‘grandmother’ in (156). When it is specifically about the two people mentioned, for example grandfather and grandmother in (157), sɔng³ kɔ⁵ ‘two people’ is used instead of khaw³.

(155) pɔ² lɛng⁵ me² lɛng⁵ khaw³
father give.birth mother give.birth 3PL
‘biological parents’

(BSS_80)

(156) a¹ pu¹ a¹ yaa² khaw³
for grandfather for grandmother 3PL
‘for grandparents/elderly people’

(BSS_78)

(157) pu¹ ta¹ yaa² sɔng³ kɔ⁵ yang⁴ a¹
grandfather and grandmother two CLF.people EXIST PRT
‘There were a grandfather and a grandmother.’

(NN_ITL_3)

The same process with khaw³ as a plural marker also exists in Khamti, sometimes in combination with nai¹ ‘this’ as nai¹ khaw⁵ (Inglis 2014: 40). However, Khamti uses khaw³ as a general plural marker, whereas Shan-Ni restricts the usage of khaw³ to people. Plural objects or animals in Shan-Ni can be marked with the marker ne¹/le¹, which is probably derived from the definite marker naj⁵/nɛ². It specifies the difference between the indefinite concept which may or may not involve plurality and a concept that is definitely plural. Hence in example (158), tin⁴ man⁴ muu⁴ man⁴ is unmarked hence denotes the concept of ‘his limb(s)’ without specifying whether this is plural or not. If we want to specify whether the limbs are plural or not, we can say muu⁴ man⁴ tin⁴ man⁴ ne¹ for ‘his limbs’ or tin⁴ man⁴ paa² nzng² or e.g. ‘his foot (one side)’. Unmarked by anything this works the same; in (161) the dishes are referred to as ‘flat plate round plate’ without a definite marker, leaving it open what type of dishes there actually are and whether it is plural or not.

(158) tin⁴ man⁴ muu⁴ man⁴
foot 3SG hand 3SG
‘his limb(s)’

(BSS_126)
Although in some sentences ne¹ could be seen as a definite marker, there are sentences in which this does not seem to be the case. This is also the only function in which the definite marker can be pronounced with an /l/. Though it is common for naj⁵ to be reduced to ne² or maybe even ne¹, le¹ is only used as a plural marker. Still, it is not used for people. The examples below are almost identical in structure, though in (162) ngo⁴ ‘cows’ is marked by le¹ while in (163) kon⁴ ‘people’ remains unmarked.

5.7 Summary

The grammatical markers discussed in this chapter are all very common in Shan-Ni daily speech. Most of them do not exist or do not have the exact same function in other Tai languages, though they do have a Burmese equivalent. This is most evident with the verbal plural kan⁴ and the linker si³. kan⁴ exists as a reciprocal marker in other Tai languages, but in
Shan-Ni it is commonly used for simultaneous actions or actions done by multiple participants, similar to Burmese cha. si³ in other Tai languages is mostly used between verbs, but in Shan-Ni is also used as a linker, similar to the usage of the Burmese lo, often meaning ‘because’ or ‘that’.

The word for ‘thing’ an² in Shan-Ni is used as a nominalizer, identical to the Burmese prefix a-. Similar to Burmese it can also be part of superlative structures. The difference is that Shan-Ni also uses it before demonstratives, which does not happen in Burmese. The only other Tai language in which a cognate of an² has similar functions is Khamti, which Inglis (2014) describes in detail in his dissertation.

The markers ko⁵, pi¹ and ne¹ / le¹ also allow direct translation to and from Burmese. Though they may have additional functions, they allow for bilinguals to express common Burmese expressions through Shan-Ni, in ways that do not always exist in other Tai languages. The plural marker ne¹ / le¹ is unique to Shan-Ni and like the Burmese twe not obligatory, but used for emphasis. The difference is that Shan-Ni does not use this plural marker for people, but uses khaw³ instead. In Khamti, khaw⁵ is used as a plural marker for both people and objects.

<table>
<thead>
<tr>
<th>Shan-Ni</th>
<th>Function</th>
<th>Burmese equivalent</th>
<th>Other Tai languages</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>an²</td>
<td>Nominalizer</td>
<td>အ- [a-]</td>
<td>Only in Khamti</td>
<td>Not for demonstratives in Burmese</td>
</tr>
<tr>
<td>kan⁴</td>
<td>Verbal plural</td>
<td>ƭ [cha]</td>
<td>Only reciprocal</td>
<td>In Shan-Ni reciprocal and multiple participants/simultaneous, like Burmese</td>
</tr>
<tr>
<td>ko⁵</td>
<td>‘also’</td>
<td>လည် [leeh]</td>
<td>Common</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Clausefinal particle</td>
<td>-</td>
<td>Topicalization in Phake</td>
<td></td>
</tr>
<tr>
<td>si³</td>
<td>Linker</td>
<td>လ [lo]</td>
<td>Between verbs</td>
<td>Between verbs and as ‘because’ or ‘that’</td>
</tr>
<tr>
<td>pi¹</td>
<td>Prohibitive</td>
<td>မ.. [ma. ne]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ne¹ / le¹</td>
<td>Plural</td>
<td>ေတွ [twe]</td>
<td>Absent, khaw⁵ in Khamti</td>
<td>Both non-obligatory</td>
</tr>
</tbody>
</table>
Chapter 6: Tense, Aspect, Mood (TAM) markers

Shan-Ni has nine tense, aspect and mood markers. Individually or in combination with each other, they place events in time. The TAM markers have grammaticalized from words that still exist in their original function in the language. To a certain extent, they correspond with the TAM markers in Phake and Aiton described by Morey (2005a, 2008). Because of the extensive description Morey has given on the TAM markers, I will discuss the TAM markers in Shan-Ni mostly in comparison to his work.

The TAM markers derived from words (mostly verbs), having to different degrees retained transparency to the original words they derived from. For example waj⁵ ‘keep’ has retained the same form as both regular verb and TAM marker, while the TAM marker kaa⁵ ‘go’ is derived from the same origin as the verb kwaa¹ ‘go’. Some of these markers have a reduced vowel (/ə/ or /ɤ/) in comparison to the word they derived from. In those cases the script has often retained the vowel of the source word. Morey (2005a: 321) categorizes these markers

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>kaa⁵</td>
<td>kaa⁵</td>
<td>GO</td>
<td>Marker of past time</td>
<td>Similar</td>
</tr>
<tr>
<td>maa⁴</td>
<td>maa²</td>
<td>COME</td>
<td>Marker of past time, used when the scope of the process indicated by the verb is in some way towards the speaker</td>
<td>Similar</td>
</tr>
<tr>
<td>waj⁵</td>
<td>wai³</td>
<td>KEEP</td>
<td>Marker of past time, used when the process indicated by the verb had or has some continuation</td>
<td>Similar</td>
</tr>
<tr>
<td>yaw⁵</td>
<td>ʒau³</td>
<td>FINISHED</td>
<td>Indicates that the process of the verb is complete, or that a section of the text has been completed</td>
<td>Similar</td>
</tr>
<tr>
<td>te¹</td>
<td>tak¹/ta¹/ti¹</td>
<td>-</td>
<td>Marker of future action, or action which should be undertaken</td>
<td>Similar</td>
</tr>
<tr>
<td>u¹</td>
<td>uu¹</td>
<td>STAY</td>
<td>Emphatic, continuous/progressive in other sources</td>
<td>Similar to other sources, not to Morey</td>
</tr>
<tr>
<td>kɔm¹</td>
<td>-</td>
<td>SHOULD/YET</td>
<td>-</td>
<td>Something expected to occur in the future, but has not occurred yet</td>
</tr>
<tr>
<td>kaw¹</td>
<td>-</td>
<td>OLD</td>
<td>-</td>
<td>Something the speaker has already experienced at some point in the past</td>
</tr>
<tr>
<td>laj²</td>
<td>dai³/nai³</td>
<td>GET</td>
<td>In preverbal position, indicates that the process has been or will be accomplished</td>
<td>Same, possibly more complex</td>
</tr>
</tbody>
</table>

94
in three different groups: those that precede the verb, those that can either follow the verb
directly or come at the end of the core sentence, and those that can only come at the end of
the core sentence. An overview of the TAM markers in Shan-Ni and a comparison with their
description in Morey (2005a, 2008) is provided below.

Five TAM markers in Shan-Ni look similar to the ones described by Morey for Phake and
Aiton. However, their functions and usages in Shan-Ni are not completely identical. Two
Shan-Ni TAM markers are not included in Morey's description for the Tai languages of Assam.
One of these is kɤn¹, which in the Tai languages of Assam does mean 'should', but is not used
as a TAM marker. In Shan-Ni it is frequently used as a TAM marker, especially in combination
with the negator ma¹, referring to things that have not happened yet. The other marker is
kaw¹, which in other Tai languages simply means 'old'. In Shan-Ni, kaw¹ marks whether
events have already been experienced in the past. The grammaticalization of verbs into TAM
markers is thus an active process, with markers existing in both varieties, having slightly
different functions and Shan-Ni introducing new TAM markers. This does not only provide
insights into the different functions Shan-Ni can express, but also into the historical processes
behind these developments.

Most Shan-Ni's TAM markers refer to the past when used alone, but have a different
function in combination with one or two other TAM markers. They do not all fit nicely into
the categories tense, aspect and mood, and have other features distinguishing them from one
another. For that reason, I follow Morey's glosses referring to the original verbs they derived
from, writing them in small caps to mark that they are grammaticalized. Hence, the lexical
gloss 'go' refers to the regular verb 'go' whereas 'GO' refers to the TAM marker which is a
grammaticalization of the verb 'go'. Following this system, I have given the Shan-Ni specific
TAM markers kaw¹ and kɤn¹ the glosses OLD and SHOULD/YET, referring to their original
meanings. For the future marker te¹ I gloss FUT, because Morey's gloss WILL does not match
the origin nor function of the marker in Shan-Ni.

In this overview, laj² 'GET' is only discussed in preverbal position in combination with the
future marker te¹. Although it is possible for laj² to be a TAM marker postverbally as well (as
it is in other Tai languages), the sound changes and additional functions TAM markers
acquire in Shan-Ni make that many particles could be interpreted as having derived from laj².
At this point, I cannot say anything conclusive about the post verbal developments of laj².
Touching upon that level of complexity goes beyond the scope of this thesis and should be
left for future research.

TAM marking in Shan-Ni is not obligatory, nor exclusive. A sentence can have zero, one,
two or more TAM markers. There is no combination I have found to be impossible, though I
will only discuss the combinations that are most common and are influencing the way the
Shan-Ni speak about time.
6.1 Phrases with TAM markers

Shan-Ni marks tense, aspect and mood either before the verb or at the end of the clause. The examples below illustrate three ways to express time, and will later be elaborated on in their own sections. In (a), \( \text{kaa}^1 \) 'GO' is used because father and mother left in a particular moment in the past. In (b), \( \text{yau}^5 \) 'FINISHED' is used, expressing that three days have come to an end. In (c) \( \text{k\(\text{"}^n\text{"} \)) \) 'YET' is used to express the parents have not returned yet.

\[
a) \quad \text{po}^2 \quad \text{me}^2 \quad \text{khaw}^3 \text{\(\text{"}^k\text{"} \))^1 \quad \text{taang}^4 \text{\(\text{"}^k\text{"} \))^4 \quad \text{kwaa}^1 \quad \text{kaa}^5 \\
\text{father} \text{mother} \text{PL} \text{ leave} \text{ path} \text{ go} \text{ go} \\
\text{’Father and mother went travelling...’}
\]

\[
b) \quad \text{saam}^3 \text{ \(\text{"}^w\text{"} \))^4 \quad \text{yang}^4 \quad \text{yau}^5 \\
\text{three} \text{ day} \text{ have} \text{ FINISHED} \\
\text{’...three days have passed already...’}
\]

\[
c) \quad \text{ma}^1 \quad \text{thong}^3 \quad \text{\(\text{"}^p\text{"} \))^4 \quad \text{maa}^4 \quad \text{k\(\text{"}^n\text{"} \))^1 \\
\text{NEG} \text{ arrive} \text{ return} \text{ come} \text{ YET} \\
\text{’...they haven’t return yet.’}
\]

(BSS_78)

The same set of TAM markers can be used to refer to the current state of the sun, using, \( \text{kaa}^1 \) for the time after sunrise, \( \text{yau}^5 \) FINISHED for daytime, when they sky has shown “its face” already and \( \text{k\(\text{"}^n\text{"} \)) \) for the sun not having risen yet.

\[
a) \quad \text{phaa}^5 \text{ \(\text{"}^l\text{"} \))^4 \quad \text{kaa}^1 \\
\text{sky} \text{ bright GO} \\
\text{’after sunrise’}
\]

\[
b) \quad \text{phaa}^5 \text{ \(\text{"}^l\text{"} \))^2 \quad \text{yaw}^5 \\
\text{sky} \text{ face} \text{ FINISHED} \\
\text{’daytime’}
\]

\[
c) \quad \text{phaa}^5 \quad \text{\(\text{"}^m\text{"} \))^1 \quad \text{\(\text{"}^l\text{"} \))^4 \quad \text{k\(\text{"}^n\text{"} \))^1 \\
\text{sky} \text{ NEG bright YET} \\
\text{’before sunrise’}
\]

(BSS_66, 67)
6.2 The terminative aspect marker *kaa⁵* ‘go’

The marker *kaa⁵* [ɕɔː] is derived from *kwaa¹* [ɕɔː] ‘go’. It is used post-verbally to describe an event that happened in the past, without specifying the moment it happened. It corresponds with the Tai Aiton equivalent described by Morey (2005a: 326) as “the simple past time, indicating that the action of the verb is ‘gone’”. In this way, it is a terminative aspect marker, in contrast with the inchoative aspect marker *maa⁴*. It is used as the most general past marker, with *maa⁴*, *yau⁵* and *kau¹* being used in more specified contexts. While the TAM marker *kaa¹* in Shan-Ni differs from the verb it derived from, *kwaa¹* ‘to go’, it does correspond with both the TAM marker and the regular verb in the Tai languages of Assam, both being *kaa¹*. In (164), one of the examples also given in the introduction, *kaa⁵* is used to say the parents went travelling, in (165) to say that the boy slept and in (166) that his returned. All of these actions were completed at a particular moment in the past 18.

(164) aymV arBF cE0f tGuf wMif; ugm um h po² mɛ² khaw³ɔk¹ taang⁴ kwaa¹ kaa⁵ father mother PL leave path go go ‘Father and mother went travelling…’

(BSS_77)

(165) mɤ² man⁴ lǝ⁵ɔn¹ lɔn⁴ kaa⁵ time 3SG child sleep go ‘When he was asleep…’

(KSN_FS_5)

(166) pa¹ying⁴ mɛɯ⁴ pɔk⁵ kaa⁵ wife 2SG return go ‘Your wife has returned.’

(EP_LP_28)

Similar to the function of *kaa¹* in Tai Aiton (Morey 2005a: 326), *kaa⁵* in Shan-Ni can mark multiple verbs in the same construction. In Shan-Ni this is a strategy often used in narratives to describe a sequence of actions, usually repetitive. In the examples below, the main character is running away from home because her mother-in-law put glass in the rice pot, so her daughter-in-law would cut herself. In (167), she is running down the stairs while blood is falling on the stairs. In (168), she arrived in the banana forest and is trying to clean her wounds while continuing to run.

---

18 The verb *ln⁴* ‘sleep’ also means ‘laying down’, hence *kaa⁵* is used because he already laid down.
The blood fell onto the ladder.'

'Around the banana trees she was cleaning and going.'

'He shot the mouse and the mouse died.'

'After going to Nammilaung, go to Maing Naung.'

19 Village names were changed to sound Burmese, but the original names are Shan-Ni. Therefore ‘Maing Naung’ is actually mɤng⁴nɔng³ ‘lake city’ and ‘Nammilaung’ nam⁵mi³lɔng² the place where bears come down to the water.
6.3 The inchoative aspect marker `maa⁴ `COME’

The inchoative aspect marker `maa⁴ `COME’ is grammaticalized from the verb `maa⁴ `come’. It marks the addition of an initial boundary to an action (Ljungqvist 2003: 22). It recognizes the initial phase as a separate state, distinguished from the other state(s) occurring or imagined in the background. It may be interpreted as a single point in time, or it can refer to an intermittent state varying in length. (Zanned 2011: 195). In other words; it marks the beginning of a verb, without specifying any details on the final boundary or the duration of the verb. When `maa⁴ is the only TAM marker, it is used for past events. In many cases, `maa⁴ and `kaa⁵ can be used in similar contexts, like in (172) and (173). With `kaa⁵ in (173), it is clear that the interlocutor is already done cooking, while in (172) with `maa⁴ it is only specified that the interlocutor started cooking. (S)he might be finished, but she might also still be cooking.

(172) мɛɯ⁴ tang² phak¹ sang⁴ maa⁴
2SG  cook  curry what  COME
‘What curry did you (start) cook(ing)?’

(BSS_116)

(173) tang² phak¹ sang⁴ kaa⁵
cook  curry what  GO
‘What curry did [you] cook?’

(BSS_96)

In (175) and (176) `maa⁴ initiates a narrative; something has happened and it is not known yet how it ended. In (175) the widow had a dream, and the sentences after that specify what happened in that dream. In (176) we learn that the frog escaped, started a narrative in which the boy and the dog search for the escaped frog.

(174) мɛ² maa²  phan³ han³ maa⁴
lady widow  dreamsee  COME
‘The widow had a dream.’

(MK_IH_4)

(175) maa³ ti²  lɔt⁵ maa⁴
dog  place  escape  COME
‘[The frog] escaped from the dog’s place.’

(KSN_FS_16)
In (176) and (177) the verbs ‘give’ and ‘jump’ do not allow for a long duration, so it is clear that these (past) actions have already been completed. In these sentences, \textit{maa} is used because it marks the initial phase of something different than that which was happening in the background.

(176) maak\textsuperscript{1}un\textsuperscript{3}huj\textsuperscript{1} rɛŋ\textsuperscript{2}hau\textsuperscript{4} maa\textsuperscript{4}  
\textit{coconut CLF.round one give COME}  
‘I gave them two bananas and a coconut to make sweets.’  
(KSN\_CAUS\_21)

(177) caa\textsuperscript{4} si\textsuperscript{3}  lɤ\textsuperscript{4}  hɤn\textsuperscript{4}  wɛn\textsuperscript{3}  tuk\textsuperscript{1}  maa\textsuperscript{4}  
\textit{for LNK on house jump fall COME}  
‘[He] jumped down from the house.’  
(BSS\_20)

With stative verbs, \textit{maa} refers to things that “became” something. It can refer to the state of a person, for example in (178) where someone became rich, or other physical things, such as rising water (179), or to situations, like in (180) where the lack of food made life difficult.

(178) man\textsuperscript{4} nai\textsuperscript{5} caa\textsuperscript{4} mɤ\textsuperscript{2} lɛɯ\textsuperscript{4}  mi\textsuperscript{4} maa\textsuperscript{4}  
\textit{3SG PRT for time which rich COME}  
‘When did he become rich?’  
(BSS\_37)

(179) nam\textsuperscript{5} ne\textsuperscript{1} yau\textsuperscript{1} maa\textsuperscript{4}  
\textit{water DEF big COME}  
‘The water became big.’  
(MK\_IH\_17)

(180) caa\textsuperscript{4} kyin\textsuperscript{4} caa\textsuperscript{4} yam\textsuperscript{5} ko\textsuperscript{5} ma’ yang\textsuperscript{4} kaa\textsuperscript{5} yaap\textsuperscript{1} maa\textsuperscript{4}  
\textit{for eat for chew PRT NEG EXIST GO difficult COME}  
‘There was nothing to eat or to chew, it became difficult.’  
(NN\_ITL\_11)

For Tai Aiton and Phake, Morey (2002: 328) describes \textit{maa\textsuperscript{2}} as “...marker of past time [...] in the specific context of the action referring towards the speaker or place where the speaker is.” This definition could work for (174), where a dream came towards the widow, or (177) where someone jumped down from a house. It does not work for sentences like (175), where the frog escaped (to an unknown place), or in other sentences where the direction does not seem to be specified by \textit{maa}. Rather, it marks the initial phase of an action, without specifying a final boundary. It is also not exclusively used for the past; in combination with
the future marker te’ it can also be used as ‘become’, as in (181) where the water will rise (become big) in seven days.

(181) cet¹ wan⁴ po⁴ waa² nam⁵ te¹ yau¹ maa⁴
    seven day enough say water FUT big COME
    ‘In seven days the water will rise.’

6.3.1 Combining [ maa⁴ ] and [ kaa⁵ ]

In the previous sections we have seen that the inchoative maa⁴ is used in sentences where the speaker refers to actions that were initiated in the past (regardless of their ending), while the terminative kaa⁵ is used to talk about actions that were finished in the past (regardless of their beginning). The combination of the two markers, maa⁴ kaa⁵, is used for events that were both initiated and finished in the past. In the examples below, the narrator describes the moment the subject had already arrived on the mountain (182), the water had risen already (183), and the subject had returned already (184). Within the Indawgyi History narrative, example (183) actually follows (179), discussed above for maa⁴. The sentences only differ in the usage of maa⁴ versus maa⁴ kaa⁵, describing a sequence saying ‘The water got big’ (it started, but we do not know whether it ended) ‘The water got big already’ (it started and finished, that’s where we are now). This repetition with a change from maa⁴ to maa⁴ kaa⁵ is common in storytelling.

(182) nyaa⁴ li⁵³ li⁴³ thɔn⁵ maa⁴ kaa⁵
    MEET on mountain arrive come GO
    ‘They had arrived on the mountain’

(183) nam⁵ ne¹ yeu¹ maa⁴ kaa⁵
    water DEF big COME GO
    ‘The water got big already.’

(184) pok⁵ maa⁴ kaa⁵ phɔ³lung⁵ man⁴ pok⁵ maa⁴ kaa⁵
    return COME GO husband 3SG return COME GO
    ‘When she came back her husband also had come back.’
6.4 The past marker waj⁵ ‘KEEP’

The past marker waj⁵ derived from the verb ‘keep’ and is used to mark the initiation of an action, with the intention of having a lasting result. Morey (2005a: 331) describes the function of the Tai Phake equivalent waj³ to be an action with the intention to have a lasting result, though there is no mention of the action being deliberately initiated by the subject. When speaking about waj⁵ in Shan-Ni, speakers tend to explain waj⁵ by translating it literally to the Burmese particle [htaa], derived from the verb [htaa] ‘put’. Okell & Allot (2001: 142) describe the function of htaa as “V permanently, V and leave, V so that there is some lasting result.” The most important factor for Shan-Ni however remains that it is something deliberately initiated. In (185) waj⁵ is used show that the speaker gave the book to the hearer, and in (186) to show that the item was bought last month. In both cases the actions where deliberate, and initiated a new situation. In (187) waj⁵ is used to refer the work the speakers ordered the sisters, emphasizing that she ordered them to work, and they are still obliged to complete it. In (188) waj⁵ follows leng⁵ ‘bare’, referring to a pig giving birth, hence also marking an action indicating a new state.

(185) phɯn³lik⁵ kau⁴ hau⁴ waj⁵ kaa⁵ meu⁴ pat¹ jaw⁵ le⁴ book 1SG give KEEP GO 2SG read FINISHED QP ‘Did you read the book I gave you?’ (PV_NS_3)

(186) pon¹ mɤ² mɤ² lɤn⁴ ne⁵ sɯ⁵ waj⁵ last time time month DEF buy KEEP ‘It was bought last month.’ (NY_NMLe_17)

(187) ma’su³ piⁿong⁵ lɤŋ² an² kau⁴ cau⁴ waj⁵ naa⁵ 2PL sisters work thing 1SG order KEEP PRT ‘You sisters, the work I ordered you...’ (BSS_210)

(188) mu³ to⁴ nɤŋ⁵ leng⁵ waj⁵ naï⁴ toj⁴ com⁴ kɤn¹ pig CLF.animal one give.birth KEEP DEF look follow SHOULD ‘A pig just gave birth come take a look.’ (BSS_12)

Okell & Allot (2001) use V here to refer to ‘verb’, and can be replaced with a verb of choice.
Another usage of \textit{waj} is as an imperative, telling someone to do something. For example in (189) where the speaker asks the interlocutor to clean the plates after using them, or in (190) where the speaker urges the interlocutor to pick the fruit carefully.

(189) oun⁴ taa⁵ sii⁴ oun⁴ sii⁴ oun⁴ taa⁵
kyin⁴ yau⁵ laang⁵ pok⁵ waj⁵
eat FINISHED clean return KEEP
‘After eating clean [the plates] again.’

(190) sii⁴ sii⁴ taa⁵ sii⁴ sii⁴ taa⁵
mak¹ maj⁵ sɔng⁴ nɤng² tim⁴ tim⁴ kyip¹ waj⁵
fruit yellow CLF.fruit one slow slow pick KEEP
‘Pick the fruit carefully.’

6.5 The perfective marker \textit{yaw} ‘FINISHED’

The verb \textit{yau} ‘finish’ is grammaticalized from the verb ‘finish’ to a past marker describing actions that have been completed. In most contexts, it can be translated literally to English as ‘already’. They describe actions with some duration, whereas \textit{kaa} and \textit{maa} are used for actions with a shorter time span. Similar to the English word ‘already’, there is an emphasis on the action being completed, perhaps against the expectation of the interlocutor. In the examples below, \textit{yaw} is used to describe a situation where three days have passed already, someone has eaten ‘everything’ emphasizing duration and completeness of the action.

(191) saam³ wan⁴ yang⁴ yau⁵
three day have FINISHED
‘…three days have passed already…’

(192) man⁴ tang⁴ long³ kyin⁴ si³ kwaa¹ yau⁵
3SG everything eat LNK go FINISHED
‘He ate everything.’

(193) su⁵ yau⁵ le²
buy FINISHED Q.POL
‘Did you buy it already?’
According to Morey (2005a: 337) yau⁴ in Phake is a perfective aspect marker, because it can also be used to speak about the future. In Shan-Ni this is also possible, though in those cases yaw⁵ has to be combined with the future marker te¹. It can be used to speak about things that will be completed in the future, for example an age that will be reached at the moment referred to (194). It is also used in combination with te⁷ to speak about things that are happening already, for example when people are coming back already (195).

(194) ကဝ⁴ နှင့် နာမတွေ နာမတွေ ကြား ကြား သို့ ယှောင်⁵ ယှာင် စီ⁵ စီ သို့ ယှာင် ယှာင်
  kaw⁴  nai⁵  saam³  sip¹  te³  thon²  yaw⁵
1SG  DEF  three  ten  FUT  reach  FINISHED
'I will be thirty years old then.'

(BSS_32)

(195) သို့ ပြေ⁵ ယှာင်⁵
  te¹  pok⁵  yaw⁵
FUT  return  FINISHED
'Coming back already.'

(BSS_33)

6.6 The past marker kaw¹ ‘OLD’

The past marker kaw¹ is used to refer to actions that have already happened, that already have been experienced. It is derived from the verb ၂ိုး kaw¹ ‘be old’, hence referring to something that is part of the past. It is similar to the usage of the Burmese particle [ဖျဖြ] hpuu which also denotes something that has already happened or the speaker has already experienced. This TAM marker is not included in Morey’s (2005a, 2008) descriptions. In Inglis (2014) I only found kaw⁵ as the verb ‘be old’. It is thus possible that this is a development limited to Shan-Ni, which considering the presence of a particle with a similar function in Burmese, might be result of Burmese influence. The Burmese particle hpuu itself does not have a connection with the concept of being old, it is more likely to be a grammaticalization of the verb hpuu, which means ‘to visit’. It is used in questions and sentences where English would use ‘ever’ or ‘never’ (in combination with negation), as in examples (196) and (197). It can also be used to declare how many times something has already happened or has already been experienced, as in (198).

(196) ကိုး ကိုး နှင့် ကြား ကြား ကြား ကြား ကြား ကြား ကြား ကြား
  keu⁴  caa⁵  man⁴  ma¹  han³ kaw¹
1SG  for  3SG  NEG  see  OLD
'I have never seen her before.'

(CM_YT_EMP)
105

6.8 The progressive marker \( u^1 \) ‘STAY’

The progressive marker \( u^1 \) derives from the verb ‘stay’. As a regular verb, it is used for both someone’s living place as well as to talk about someone’s state of being, for examples in the phrase \( u^1 ni^4 le^4 ‘how are you?’ \). As a TAM marker it is a progressive marker, used to describe actions that are going on at the moment referred to, usually in combination with the linker \( si^2 \). Diller (1992: 24) describes it as a progressive-continuative marker, marking progressive aspect. Morey (2005a: 334) only found \( u^1 \) with emphatic meaning often used in combination with adjectives. I did also not really find \( u^1 \) by itself, only in combination with the linker \( si^3 \). However, it is included in the Basic Spoken Shan-Ni book, and a possible explanation might be that because TAM markers are usually phonologically reduced, the pronunciation of \( u^1 \) could be very subtle. Therefore, I do include the examples from the book here, but I would like to emphasize that its usage without \( si^3 \) needs further research. Written Shan-Ni allows variation between \[ u \] and \[ yu \], though \( yu \) never comes up in spontaneous speech or elicitation.

105

Often \( u^1 \) is preceded by the linker \( si^3 \), meaning that something is happening or someone is doing something; for example, children who are playing or a person who is eating a lot. When preceded by an adjective, it refers to a temporary state, for example being hungry. In example (34), \( si^2 u^1 \) is combined with \( yaw^5 \) to emphasize ‘we’ are hungry already.
6.8 The marker \textit{kɤn}¹ ‘should’ or ‘yet’

\textit{kɤn}¹ is used to describe situations that should happen or did not occur yet. It can either express a request of the speaker for the interlocutor, or in combination with the negator, to speak about things the speaker did not do yet or things that did not happen yet. It is not mentioned by Morey (2005a) or Inglis (2014) as having any meaning beyond ‘should’, but is used very frequently in Shan-Ni, especially in combination with the negator \textit{ma}¹. When not negated, \textit{kɤn}¹ can be used to mark a request. This usage of \textit{kɤn}¹ is closer to the original meaning of the lexeme ‘should’. In (204) \textit{kɤn}¹ is used to invite someone to take a look at newborn piglets, and in (205) it is used for a request in a conditional phrase.

\begin{itemize}
  \item (204) \textit{mu}³ \textit{to}⁴ \textit{nɤng}² \textit{leng}⁵ \textit{wai}⁵ \textit{nai}⁵ \textit{toj}⁴ \textit{com}⁴ \textit{kɤn}¹
  
  \begin{tabular}{llllllllllll}
    \textit{pig} & \text{CLF.animal} & \text{one} & \text{give.birth} & \text{KEEP} & \text{DEF} & \text{look} & \text{follow} & \text{SHOULD} \\
  \end{tabular}

  ‘A pig just gave birth, come take a look.’

  \item (205) \textit{sang}³ \textit{pi}⁴ \textit{het}¹ \textit{maa}⁴ \textit{mok}³ \textit{kɤn}¹ \\

  \begin{tabular}{llllllllllll}
    \textit{thing} & \text{if} & \text{do} & \text{come} & \text{inform} & \text{SHOULD} \\
  \end{tabular}

  ‘If you do it let me know.’
\end{itemize}
The usage of \( kɤn^1 \) in combination with the negator is identical to the Burmese particle \( [\text{ေသး}] \) \( tee \). Since this usage is unique to Shan-Ni and not found in other Tai languages, it is very likely to be Burmese influence. Similar to \( kaw^1 \) it appears to be a Shan-Ni word which gained the function of a common Burmese particle, probably as a result of widespread Burmese/Shan-Ni bilingualism among the Shan-Ni.

\[(206)\] \( \text{kwa}^1 \text{thɛu}^4 \text{ ma}^1 \text{ hu}^5 \text{ kɤn}^1 \)  
\[ \text{go where NEG know YET} \]
\[ \text{‘I’m not sure where I’m going yet.’} \]

\[(207)\] \( \text{ma}^1 \text{ lai}^2 \text{ yap}^5 \text{ kɤn}^1 \)  
\[ \text{NEG GET sow YET} \]
\[ \text{‘[I] didn’t get it sowed yet.’} \]

\[(208)\] \( \text{ma}^1 \text{ thɤng}^3 \text{pɔk}^4 \text{ ma}^4 \text{ kɤn}^1 \)  
\[ \text{NEG arrive return COME YET} \]
\[ \text{‘...they didn’t return yet.’} \]

6.9 The future marker \( te^1 \)

The future marker \( te^1 \) is the only TAM particle that occurs preverbally. The most basic usage of \( te^1 \) is to refer to events that will happen in the future. In \( (209) \) it is used for an illness that will be cured, and in \( (210) \) to talk about what is going to happen next months. For all events happening in the future, \( te^1 \) is an obligatory marker.

\[(209)\] \( \text{yɔ}^1 \text{ga}^1 \text{ te}^1 \text{ haaj}^3 \)  
\[ \text{illness.BM FUT loose} \]
\[ \text{‘The illness will be cured.’} \]

\[(210)\] \( \text{lɔm}^4 \text{ laa}^2 \text{ te}^1 \text{ maa}^4 \text{ thop}^1 \text{ nɔ}^3 \text{ te}^1 \text{ thop}^1 \text{ tɔ}^1 \text{ kan}^4 \text{ yu}^1 \)  
\[ \text{month face FUT come meet TAG FUT meet talk SIM STAY} \]
\[ \text{‘Next month you will come meet [me] right? [We] will speak to each other.’} \]

\( te^1 \) is also used for things that are about to start. In \( (211) \) it is used to announce the speaker is going to tell a story, and in \( (212) \) it is used to say that the father goes into the forest. When
combined with *kwaa* 'go' it might be that someone is already on their way, for example in (212) where the father is already walking into the forest.

(211) a’cang¹ te¹ khaj²pyaa⁴ pom¹ ne² now FUT tell show story DEF ‘Now I’m going to tell a story.’

(212) pɔ² man⁴ te¹ kwaa¹ lɛɯ⁴ te¹ tɤn² father 3SG FUT go inside forest ‘The father goes into the forest.’

To express uncertainty, things that the speaker assumes to be true but is not entirely sure of, or things that the speaker proposes to the interlocutor, *te¹* is used in combination with *laj²* 'GET'. In (213), the main character of the story is wondering whether the blood he sees is his wife’s blood. In (214), the speaker tells her fellow villagers that they could run before the flood. (215) is a typical market situation, where the speaker is asking the how much (s)he should give, asking the vendor to tell reveal the price.

(213) lɤt⁵ pa¹ying⁴ kew⁴ te¹ laj² pjen⁴ sang⁴ i³ blood wife 1SG FUT GET be what LNK ‘Maybe that’s my wife’s blood’

(214) cet¹ wan⁴ pɔ⁴ waa² nam⁵ te¹ jau¹ maa⁴ haw⁴ te¹ laj² paaj² seven day enough say water FUT big COME 1PL FUT GET run ‘In seven days the water will rise, we could run.’

(215) haw⁴ kha’lau⁴ te¹ laj² hau² 1PL how.much FUT GET give ‘How much should I give?’

Diller (1992) and Morey (2005a) include the function of expression of volition or intention to the marker *te¹* for the languages of Assam. For an irrealis this would make sense; the future spoken about is of course the one expected or desired to happen. However, in Shan-Ni *te¹* distinguishes itself from *kɤn¹*: whereas *kɤn¹* expresses volition in relation to the
interlocutor’s actions, te¹ simply refers to the expected future, which can include one’s own actions. In order to explicitly express volition, Shan-Ni can use the auxiliary verb khaj² ‘want’, shortened in speech in front of a verb to kə².

(216) ဗူး သား စားနွှဲမှု
ma¹ khaj² non⁴
NEG want sleep
‘I don’t want to sleep.’

(217) သား စားနွှဲမှုကို မ၁ သား
khaj² kyin⁴ sang⁴
want eat what
‘What do you want to eat?’

In combination with other TAM markers, it may be used to talk about the past as well as the future. When te¹ is combined with kaa⁵ or maa⁴ kaa⁵ it refers to something someone was going to do. In that context it can be interpreted as being about volition. In combination with yaw⁵ or maa⁴, it refers to things that will be completed in the future. This distinction shows that while these are all past markers, in combination with te¹ they may refer to very different points in time. In (218) te¹ pɔk⁴ maa⁴ kaa⁵ is used to say that someone was about to go back, and in (219) te¹ sɔt¹ phaj⁴ kaa⁵ they were already about to light the fire. These events both refer to things that were about to happen in the past. In contrast, in (220) te¹ thon² yaw⁵ is about the age the speaker will reach in future, and (221) te¹ yaur⁴ maa⁴ is about something that will become big.

In combination with other TAM markers, it may be used to talk about the past as well as the future. When te¹ is combined with kaa⁵ or maa⁴ kaa⁵ it refers to something someone was going to do. In that context it can be interpreted as being about volition. In combination with yaw⁵ or maa⁴, it refers to things that will be completed in the future. This distinction shows that while these are all past markers, in combination with te¹ they may refer to very different points in time. In (218) te¹ pɔk⁴ maa⁴ kaa⁵ is used to say that someone was about to go back, and in (219) te¹ sɔt¹ phaj⁴ kaa⁵ they were already about to light the fire. These events both refer to things that were about to happen in the past. In contrast, in (220) te¹ thon² yaw⁵ is about the age the speaker will reach in future, and (221) te¹ yaur⁴ maa⁴ is about something that will become big.

(218) ဝက် ရောင် ဝတ် ကျွန်ုပ် ပန် ပယ် ပယ် ပယ် ပယ် ပယ်
wa² ji³ kap¹ ji³ te¹ pɔk⁴ maa⁴ kaa⁵ lə⁵ pə³ pə³ pə³ pə³ pə³
talk LNK communicate LNK FUT return COME GO blood all over
‘After discussing, he was going back and there was blood all over.’

(219) သား စားနွှဲမှုကို မ၁ သား
te¹ sɔt¹ phaj⁴ kaa⁵
FUT light fire GO
‘When they were going to light the fire....’

(220) ကျောင်း မောင် စားနွှဲမှု သီး သီး သီး သီး သီး
kaw⁴ nai⁵ saam³ sip¹ te¹ thon² yaw⁵
1SG DEF three ten FUT reach FINISHED
‘I will be thirty years old then.’
In seven days the water will rise.

6.10 Summary

Table 31 summarizes the functions of Shan-Ni TAM markers and how they relate to verbs. Shan-Ni has at least five past markers, and more when including the possible combinations and the negation of kɤn¹ for things that have not happened yet. All of these markers are of Tai etymology, though their functions as TAM markers are likely to have developed under influence of Burmese and other neighbouring languages. Shan-Ni TAM markers illustrate two types of developments: changes in form, and changes in meaning. Changes in form are phonological changes that TAM markers have gone through, resulting in a difference between

<table>
<thead>
<tr>
<th>TAM markers and their functions in relation to verbs (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>maa⁴</td>
</tr>
<tr>
<td>kaa⁵</td>
</tr>
<tr>
<td>maa⁴ kaa⁵</td>
</tr>
<tr>
<td>waj⁵</td>
</tr>
<tr>
<td>jaw⁵</td>
</tr>
<tr>
<td>kaw¹</td>
</tr>
<tr>
<td>u¹</td>
</tr>
<tr>
<td>si³ u¹</td>
</tr>
<tr>
<td>kɤn¹</td>
</tr>
<tr>
<td>ma¹...kɤn¹</td>
</tr>
<tr>
<td>te¹</td>
</tr>
<tr>
<td>te¹...yaw⁵</td>
</tr>
<tr>
<td>te¹...maa⁴</td>
</tr>
<tr>
<td>te¹... (maa⁴ kaa⁵)</td>
</tr>
</tbody>
</table>
the word it derived from and the grammaticalized TAM marker. In some cases, they can be observed from a difference between Shan-Ni and other Tai languages, or because the Shan-Ni writing system still retains the old form. The clearest example of this is kaa⁵'go' which has developed from kwaa¹'go'. In pronunciation however, both kaa⁵ and maa⁴ often get a /ə/ or /ɤ/. The progressive u¹ can be written as both u¹ and yu¹ though yu¹ is rarely used in spoken Shan-Ni.

The other development is the acquisition of new functions. The usages of kɤn¹ and kaw¹ are unique to Shan-Ni. While the other TAM markers have developed from verbs, especially kaw¹ 'OLD' is an odd one out. They are however not odd, when considering Burmese, which has particles with the same function that are used very frequently. The past marker waj⁵ 'KEEP' does exist in other Tai languages, although the function in Shan-Ni appears to have additional features not present in other Tai languages. In other Tai languages this marker refers to actions which are intended to have a long-lasting result, in Shan-Ni they are actions that have been initiated deliberately by the speaker. This is identical to the function of the Burmese marker [တား ] htaa, a particle derived from the verb 'put'.

Other TAM markers share small differences with the TAM markers of the Tai languages of Assam, and do not show any clear signs of influence from other languages. Combinations of TAM markers allow Shan-Ni speakers to express themselves about time in even more ways. The past marker yaw⁵ and (maa⁴) kaa⁵ differ slightly from each other when used individually, though they could be interpreted in the same way depending on the context. In combination with the future marker te¹, the usage diverges: te¹...yaw⁵ is used for something that will be completed, while te¹...(maa⁴) kaa⁵ means something was going to happen.
Chapter 7: 
Shan-Ni and other languages: Genetic affiliations and contact

The different features of Shan-Ni discussed in the preceding chapters, provide clues about its position within the Tai language family. This chapter is dedicated to a more extensive discussion on the relation between Shan-Ni and the other Tai languages of the region, going deeper into the processes described in earlier chapters. It also looks into the extent to which diverging features in Shan-Ni correspond with Burmese, indicating where these could be a result of contact. Burmese is the language with which Shan-Ni speakers have had most contact in the last few centuries. The other languages spoken in the region are mostly also Tibeto-Burman languages, and share features with Burmese, hence similarities may be the result of older contact between Shan-Ni and these languages as well. This chapter will highlight a few features that stand out, and for which there is sufficient data available for the languages in the comparison. The phonology, including consonant clusters and tones, the interrogatives, demonstratives and TAM markers are relatively well described for the other Tai languages. The formation of disyllabic nouns and its reduction to sesquisyllabic structures (introduced in Chapter 3) will be included shortly as well, as a special feature within Shan-Ni. As discussed in section 1.5, there are historical accounts suggesting different instances of migration of the Tai groups in Northern Myanmar. This chapter will conclude by looking at the extent to which the connections between Shan-Ni and different Tai groups fit with the possible migration patterns suggested in Chapter 1. Based on this I will suggest a preliminary genetic affiliation for Shan-Ni, which is meant to set the foundation for more extensive genetic research in the future.

7.1 Phonology

Both Edmondson (2008) and Chantanaroj (2007) compare different varieties of Southwestern Tai languages based on the phonology. Edmondson’s conclusion is that the phonology shows influence of migration from the east into the Shweli river valley, and Chantanaroj’s analysis shows a great number of correspondences between the Tai varieties of the Shweli river valley and the different Shan-Ni locations included in her study. Edmondson recorded a number of sound changes in Shan-Ni and Tai languages in the region which are enlisted in Table 32, alongside their pronunciation in Shan-Ni, and the places in

<table>
<thead>
<tr>
<th>Proto Tai phoneme</th>
<th>Shan-Ni pronunciation</th>
<th>Locations with same pronunciation (Edmondson 2008)</th>
</tr>
</thead>
<tbody>
<tr>
<td>*kh</td>
<td>/x/</td>
<td>Dehong (China), Bhamo, Puta-O</td>
</tr>
<tr>
<td>*f</td>
<td>/pʰ/</td>
<td>Kachin state</td>
</tr>
<tr>
<td>*k(h)w</td>
<td>/kw/</td>
<td>Dehong and Kachin</td>
</tr>
<tr>
<td>*k</td>
<td>/tɕ/</td>
<td>Outside Shan state</td>
</tr>
</tbody>
</table>
which the pronunciation of these consonants corresponds with Shan-Ni. Overall, this shows that Shan-Ni shares most phonological features with other locations within Kachin state and the Dehong region in China. According to Edmondson, the shift from *k → /tɕ/ is very common in the region, but this shift is not included in any of the language descriptions of Tai languages in the region (e.g. Inglis 2014, Morey 2005, Young 1985). The consonant clusters are the most interesting part of Edmondson’s comparison, and will be discussed in more detail in the next section.

7.1.1 Consonant clusters

<table>
<thead>
<tr>
<th>Table 33 Consonant clusters in Shan-Ni and their cognates in other Tai languages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Shan-Ni</strong></td>
</tr>
<tr>
<td>kwaa⁴/kaa⁵</td>
</tr>
<tr>
<td>khwam⁴</td>
</tr>
<tr>
<td>khwaaj⁴</td>
</tr>
<tr>
<td>kyin⁴</td>
</tr>
<tr>
<td>kyep¹</td>
</tr>
<tr>
<td>khyem³</td>
</tr>
<tr>
<td>khyek¹</td>
</tr>
<tr>
<td>khyet¹</td>
</tr>
<tr>
<td>pyet¹</td>
</tr>
<tr>
<td>pyet¹</td>
</tr>
<tr>
<td>pyin⁴/pin⁴</td>
</tr>
<tr>
<td>maak¹phyit¹</td>
</tr>
<tr>
<td>pyet⁵</td>
</tr>
<tr>
<td>myit⁶</td>
</tr>
<tr>
<td>myet¹</td>
</tr>
<tr>
<td>myin⁵</td>
</tr>
</tbody>
</table>

Both consonant clusters /kw/ and /tc/ <ky> included in Edmondson (2008) are common in Shan-Ni. The cluster /kw/ is retained from proto-Tai, but has reduced to a /k/ in many Tai varieties. The Tai varieties in Kachin state (including Shan-Ni) and the Dehong region in China have retained the old form. In Tai Mau (Young 1985) many words allow variation in

21 Sources: Morey 2005a (Phake), Young 1985 (Tai Mau), SEAlang Database (Shan Gyi), Enfield 2007 (Lao).
k~kw and kh~khw. The /k/ itself did change in many words to a /tɕ/, represented as <ky>. Edmondson does not discuss the other consonant clusters or palatalizations common in Shan-Ni. Edmondson mentions palatalization as a possible next stage in loss of consonant clusters, but he did not find this in his own data (Edmondson 2008: 201).

In Shan-Ni palatalization occurs between the consonants /k, p, ph, m/ and vowels /e, i, ɛ/. Table 33 shows the consonant clusters in Shan-Ni and their cognates in other Tai languages. Palatalization of the consonants /k, kh, p, ph, m/ is part of many Shan-Ni words with Tai origin. It is most common in combination with the velar stops /k, kh/, which correspond with Edmondson’s findings for other Shan varieties. There are also many palatalizations with the labial stops /p, ph/, and less with the labial nasal /m/. I would need more data to confirm this, but I did not find any instances of any of these consonants followed by /-ɛt, -et, -it/ without palatization. In combination with other final consonants and the vowels /e, i, ɛ/, palatization of the initial consonants /k, kh, p, ph, m/ in combination with is possible but not obligatory.

Shan-Ni shares the cluster /kw/ with Shan Gyi and Lao. In Tai Mau /k/ and /kw/ are allophones, and may be both used within the same words. In Shan-Ni the word kwaa¹ is the verb ‘go’ and kaa⁵ is a TAM marker grammaticalized from the verb go. Hence both these variants are present, though they have different meanings. The interesting thing is that in Phake this same TAM marker exists in this form, but it is identical to the ungrammaticalized verb ‘go’ in Phake. This supports the narrative that there might have been a pre-existing Tai group in the Mogaung area, and that migrants came in together with the ruler of Mong Mao around 1200. In that scenario the reduction to kaa⁵ and its grammaticalization to a TAM marker could be a development in the sedentary Tai group, which merged with the incoming Tai language of migrants that had retained the cluster /kw/. The variant with /kw/ could then be used for the verb ‘go’ but not for the TAM marker, because it was not present in the variety that had retained /kw/. The problem with that theory is that it would not explain why Phake does have kā¹ both as the verb ‘go’ and as the TAM marker.

7.1.2 Tone

Figure 9 shows all the Tai tone boxes of the region. Note that the numbers indicate how the tones are grouped together within individual languages, not how they compare across languages. The colours give an indication of how the tones are grouped together across different languages. They might have developed different contours and pitch levels, but the splits are indicators of more salient changes, and therefore provide more reliable information about the historical relations between these different languages. I have tried to take the most recent, reliable tone boxes available, but this does not take away from the fact that regional and individual nuances do effect the tone boxes. My Indawgyi tone box does not correspond with the tone box recorded by Wyn Owen for Shan-Ni as spoken in Myitkyina, which is a regional difference speakers are also aware of. For Khamti there are many different tone boxes, though most of them are from very old sources and the region unspecified. I have taken
Figure 9 Tone boxes of different Tai languages of the Southwestern Tai language family

<table>
<thead>
<tr>
<th>Shan-Ni (Indawgyi)</th>
<th>Khamti (Chindwin, adapted from Dockum 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A      B       C     DS     DL</td>
<td></td>
</tr>
<tr>
<td>1      3       1      2      1</td>
<td></td>
</tr>
<tr>
<td>2      4       1      2      1</td>
<td></td>
</tr>
<tr>
<td>3      2       5      1      2</td>
<td></td>
</tr>
<tr>
<td>4      2       5      1      2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tai Phake (adapted from Morey 2005a)</th>
<th>Shan-Ni (Myitkyina, adapted from Owen 2013)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A      B       C     DS     DL</td>
<td></td>
</tr>
<tr>
<td>1      6       1      3      1</td>
<td></td>
</tr>
<tr>
<td>2      2       1      3      1</td>
<td></td>
</tr>
<tr>
<td>3      5       1      2      1</td>
<td></td>
</tr>
<tr>
<td>4      4       6      1      2</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tai Aiton (adapted from Diller 1992)</th>
<th>Khamyang (adapted from Morey 2005b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A      B       C     DS     DL</td>
<td></td>
</tr>
<tr>
<td>1      4       6      5      6</td>
<td></td>
</tr>
<tr>
<td>2      3       1      5      6</td>
<td></td>
</tr>
<tr>
<td>3      2       1      5      6</td>
<td></td>
</tr>
<tr>
<td>4      1       1      5      6</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Tai Nua (adapted from Chantanaroj 2007)</th>
<th>Tai Mao (adapted from Chantanaroj 2007)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A      B       C     DS     DL</td>
<td></td>
</tr>
<tr>
<td>1      1       4      5      4      1</td>
<td></td>
</tr>
<tr>
<td>2      2       4      5      4      1</td>
<td></td>
</tr>
<tr>
<td>3      3       2      6      2      2</td>
<td></td>
</tr>
<tr>
<td>4      3       2      6      2      2</td>
<td></td>
</tr>
</tbody>
</table>
the tone box by Dockum (2014), which is a recent one of the Chindwin region, where Shan-Ni is spoken as well. For Phake and Khamyang I use the tone boxes by Morey (2005a, 2005b), and for Aiton I use the work of Diller (1992). Chantanaroj’s (2007) thesis is on Tai Nua, but she also includes an overview of the tone boxes of different languages including Tai Mao.

The Indawgyi Shan-Ni tone box corresponds most with Khamti (Chindwin) and Tai Phake. The only difference with Khamti is that it distinguishes B4-C123 (blue) from C4-DS4-DL4 (orange), which in Khamti is the same tone. In Phake B4-C123 is split up further between C123 (blue) and B4 (red). In Myitkyina Shan-Ni and Khamyang B4 is also different, but grouped together with A23 (red). This A23-B4 group also exists in Tai Mao and Tai Nua, though in these languages it also includes DS4 and for Tai Nua DL4. Tai Mao and Tai Nua furthermore group DL123 with A1, which is not the case in any of the other languages. Aiton is also interesting, because it groups A234 together like Indawgyi Shan-Ni, Khamti and Tai Phake do, but follows the exact same pattern of Tai Nua in row 4 (B4-C4-DS4-DL4). The different correspondences between these tonal splits are summarized in Table 34.

| Table 34 Tone splits in Shan-Ni and other Tai languages (based on Figure 9) |
|-----------------|---|---|---|---|---|---|
|                | A23 | B4 | C4 | DS4 | DL4 | DL123 |
| Khamti (Chindwin) | x   |   | x  |   | x  |   |
| Shan-Ni (Indawgyi) | x   | x | x  | x  |   | x  |
| Phake           |     |   | x  | x  | x  | x  |
| Aiton           |     | x | x  |   |   |   |
| Shan-Ni (Myitkyina) |     |   | x  | x  | x  | x  |
| Khamyang        |     | x | x  |   |   |   |
| Tai Mao         |     | x | x  |   |   | x  |
| Tai Nua         |     | x | x  |   |   | x  |

The developments shown in Table 34 do not show a gradual move from one tonal pattern to the other. Rather, my assumption is that it shows the different ways migrants (related to the Tai Nua) and the sedentary population (related to the Tai Khamti) have adapted their tones to each other at different times, places and probably in different proportions of population groups. Myitkyina Shan-Ni tonal splits are identical to Khamyang, which would suggest that they might be closely related. Although little is known about the Khamyang, no sources suggest they might originate near Myitkyina.

---

22 She cites Robinson (1994) and Gedney (1973), but she does not specify per language which source she uses. Since I compare Shan-Ni with Tai Mao elsewhere I wanted to include it, but if new work on Tai Mao is published this part should be revised.

23 Morey (2005a: 20) is unsure about the history of the Khamyang, writing that he only found one pamphlet without sources mentioned saying the Khamyang had lived in Sagaing near the Indian border for 500 years before they moved to Assam in 1780 (Morey 2005a:20). However, they also use the initial /l/ like speakers of Myitkyina Shan-Ni do, in words that in Sagaing and India always are pronounced with an /n/. This connection should be further investigated.
Tai Nua and Tai Mao are distinct from all others by grouping DL123 with A1, while all the other languages group DL123 with B123 and DS123 together. Geographically, the other groups are all northwest of where Tai Nua and Tai Mao are spoken, and they share a history under the rulers of Mogaung, while the Tai Nua and Tai Mao do not. However, around 1200 Mogaung was under Mong Mao in the Shweli region, where Tai Mao is spoken today. The parabaik manuscript included in the thesis of Hla Maw Maw (2017: 85) describes a royal order to the mayor of Maing Mao (near Myitkyina) to ensure peace, accommodate and provide for the increased population and divide the township into four new regions. Edmondson (2008: 203) refers to a large scale migration around that same time from the Tai Nua area into the Shweli area, where the Tai Mao live. It is very likely that the same wave of migration from the east came to both the Shweli valley as well as the Myitkyina region. The significantly stronger similarity between Tai Nua and Tai Mao may point to a stronger influence of Tai Nua on Tai Mao than on the other languages. There are three possible reasons for this: (a) the proportion of migrants might have been bigger in the Shweli valley; (b) the sedentary population in the Shweli valley may have spoken a language which was already more similar to Tai Nua prior to the arrival of the migrants; (c) the migrants may have first gone to the Shweli valley and later gone up to Mogaung/Maing Mao. The last theory supports that they came along with the new king Sam Long Hpa who was a brother of the king of Mong Mao, and that the migrants had a royal endorsement, as mentioned in the parabaik manuscript.

Aiton, Phake, Indawgyi Shan-Ni and Khamti group A234 are together, while Myitkyina Shan-Ni, Khamyang, Tai Mao and Tai Nua have a split between A23 and A4. This might show a stronger influence of Tai Nua/Tai Mao migrants in Myitkyina in comparison to the other areas. Myitkyina Shan-Ni and Khamyang also group A23 together with B4, like Tai Mao and Tai Nue do as well. In Indawgyi Shan-Ni and Khamti B4 is grouped together with C123. All groups except for the Khamti distinguish C4-DS4-DL4 from C123. This is a development all groups have gone through, except the Khamti who remained more isolated in the north.

The Aiton split DS4-DL4 off from C4 in the same way Tai Nua does. The Aiton presumably left Mogaung earlier than the Phake, though it is not exactly sure when. Morey (2005a) assumes that the Tai Aiton have been in Assam from at least 1500 onwards. This would explain their retention of the Tai Nua DS4-DL4 tones. The Phake left Mogaung in 1775 for Assam, joining the Aiton who had come earlier. The A23 and B4 distinction Phake and Aiton have in common, might be a result of their renewed contact.

Both tonal developments and historical events need more reseach, but the fact is clear that migration of large population groups has affected tonal developments in Tai languages of northern Myanmar and the surrounding borderlands.

24 Discussed earlier in section 1.5.
7.2 The shape of words

7.2.1 Class terms and sesquisyllables

The formation of sesquisyllables is an active process in Shan-Ni (section 3.1.1), and can be found to a limited extent in other Tai languages, but nowhere as frequent as in Shan-Ni. It is connected to the process of forming compounds by adding class terms, followed by a reduction of the first syllable. However, it also used to expand the number of demonstratives (discussed separately in section 7.3.3) and distinguish between personal and possessive pronouns.

In Lao (Enfield 2007: 147) class terms apply to objects, and are not phonologically reduced to sesquisyllables. They are also very much still transparent and separable compounds, of which each syllable also has a meaning when used individually. The process to get there however, is identical to the formation of class terms in Shan-Ni. The class terms are the first part of the compound, indicating the category (e.g. ‘fish’) and the second part can be either a name of a specific thing within that category (e.g. ‘mackerel’) or a descriptive term, forming in a new meaning together with the class term (e.g. ‘gold’ to form ‘goldfish’).

The application of class terms and sesquisyllables to personal pronouns is especially interesting, because in Khamti and Phake all personal and possessive pronouns are the same, with their position specifying whether they are personal (preceding the noun) or possessive (following the noun). In Shan-Ni this is also the case, though the second and third person plural\(^{25}\) are prefixed by \(mə³\), probably derived from the pronouns for second and third person singular, \(mɛɯ⁴\)'2\(SG\)' and \(man⁴\)'3\(SG\)'.

Phake and Aiton have what Morey (2002: 229) refers to as “noun class markers”. In his description most of these are already intransparent, and do not have a full-form equivalent like some nouns in transition in Shan-Ni have. Due to this lack of transparency, he also does not link their semantics to full lexemes, though he does describe their functions. The forms and their descriptions correspond with Shan-Ni. The two noun class markers that are not phonologically reduced in Phake he does consider to be be full nouns, but he does not discuss a possible transition.

The partial presence of this process in Phake and Aiton, shows that the phonological reduction probably already took place in a time when the Phake and Shan-Ni were still living alongside each other, but that Shan-Ni expanded the use of class terms significantly, creating more disyllabic nouns than any other Tai language. The description of the formation of class terms in Lao facilitates the understanding of this process in Shan-Ni, which works exactly as Enfield describes, though it is more extensive.

The way these compounds are reduced phonologically, meets all the requirements listed by Butler (2014: 11) for sesquisyllables. The prosodic prominence is word final, the non-final syllables are reduced phonologically through a reduced segmental inventory; reduced

\(^{25}\) \(ma³\)'2\(PL\)' and \(ma³haw³\)'3\(PL\)' while the possessive pronouns are \(su³\)'2\(PL.POS\)' and \(khaw³\)'3\(PL.POS\)', corresponding to Phake and Khamti personal and possessive pronouns.
syllable shape and the non-final syllables do not constitute well-formed prosodic words on their own. The development of this as a very dominant feature in the language should be seen in relation to other languages in the area that also have sesquisyllables. These are all Tibeto-Burman and Austroasiatic languages, and in order to understand how this process has come into Shan-Ni more information and further research is required into the exact correspondence between these features and the contact history.

7.2.1 Interrogatives

Tai interrogatives usually correspond with the indefinite pronouns, with their function depending on their place within the syntax. All Tai languages in Table 35 have cognates of Shan-Ni leu⁴ ‘which’, pheu⁴ ‘who’ and mɤ² leu⁴ ‘when’/‘which time’, but the rest of the interrogatives vary.

The innovation theu⁴ ‘where’ is described by Young (1984: 27) as a contraction of ti² ‘place’ and leu⁴ ‘which’. This innovation is present in Tai Mau, Shan-Ni and Phake, but not in Khamti, Shan Gyi and Lao. In Shan Gyi ti; lau ‘which place’ is the uncontracted form of what Young suggests to be the source of theu⁴ ‘where’ its cognates, while Khamti and Lao interrogatives have been through different developments. In Khamti, amai⁵-nai⁵ ‘where’ is a compound of an³ ‘thing’, mai³ ‘here’ and nai⁵ ‘Q’. The Khamti word for ‘here’ mai³ does not exist in the other Tai languages in the table.

Table 35 Interrogatives in Tai languages\(^{26}\)

<table>
<thead>
<tr>
<th>English</th>
<th>Shan-Ni</th>
<th>Phake(^{27}) (+Aiton)</th>
<th>Khamti</th>
<th>Tai Mau</th>
<th>Shan Gyi</th>
<th>Lao(^{28})</th>
</tr>
</thead>
<tbody>
<tr>
<td>polar Q</td>
<td>le⁴/le²</td>
<td>-/nɔ⁶/naü⁵</td>
<td>hâa nɔ</td>
<td>bɔo³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>what</td>
<td>sang⁴</td>
<td>ka³saŋ⁶/ksang⁵-nai⁵</td>
<td>sänja sʰaɯ</td>
<td>nang³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>who</td>
<td>pheu⁴</td>
<td>naÜ⁶/an³lae³-nai⁵</td>
<td>lai</td>
<td>lau</td>
<td>N-daj³</td>
<td></td>
</tr>
<tr>
<td>where</td>
<td>theu⁴</td>
<td>thau⁶/phæu⁵-nai⁵</td>
<td>ph’ai ph’ai³</td>
<td>phaj³</td>
<td></td>
<td></td>
</tr>
<tr>
<td>how</td>
<td>nang¹læu⁴</td>
<td>nũŋ¹ hũ⁶</td>
<td>heu⁵læu⁵-sii⁵</td>
<td>hũ</td>
<td>hũ</td>
<td>-</td>
</tr>
<tr>
<td>when</td>
<td>mr² leu⁴</td>
<td>mɔ⁵naũ⁶/meeu⁵læu⁵-nai⁵</td>
<td>mə lai</td>
<td>mə: lɛɯ</td>
<td>mùa¹daj³</td>
<td></td>
</tr>
<tr>
<td>how much</td>
<td>kha¹læu⁴</td>
<td>kha¹læu⁴/ken⁵lae³-nai⁵</td>
<td>khai</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>how many</td>
<td>kha¹læu⁴</td>
<td>kha¹læu⁴/ken⁵lae³-nai⁵</td>
<td>-</td>
<td>laaj (+)</td>
<td>cak² (+CLF)</td>
<td></td>
</tr>
</tbody>
</table>

The innovation theu⁴ ‘where’ is described by Young (1984: 27) as a contraction of ti² ‘place’ and leu⁴ ‘which’. This innovation is present in Tai Mau, Shan-Ni and Phake, but not in Khamti, Shan Gyi and Lao. In Shan Gyi ti; lau ‘which place’ is the uncontracted form of what Young suggests to be the source of theu⁴ ‘where’ its cognates, while Khamti and Lao interrogatives have been through different developments. In Khamti, amai⁵-nai⁵ ‘where’ is a compound of an³ ‘thing’, mai³ ‘here’ and nai⁵ ‘Q’. The Khamti word for ‘here’ mai³ does not exist in the other Tai languages in the table.

---

\(^{26}\) Sources: Morey 2005a; (Phake), Inglis 2014 (Khamti), Young 1985 (Tai Mau), Glick & Sao Tern Moeng 1991 (Shan Gyi), Enfield 2007 (Lao).

\(^{27}\) Aiton is not included in the table. In Morey (2005a: 222) all Aiton interrogatives are cognates of the Phake interrogatives, the only differences are minor, regular phonological distinctions between Phake and Aiton.

\(^{28}\) This table does not do justice to the complexity of interrogatives in Shan Gyi and Lao. They are used here in comparison to Shan-Ni, but for a comprehensive overview see Enfield (2007) and Glick & Sao Tern Moeng (1991).
Most Shan-Ni interrogatives correspond with Phake, with the exception of ‘how many’ and the polar question particle le⁴/le². Cognates of the Phake word ki⁴ also exist in Khamti and Aiton. In Shan-Ni kha¹leu⁴ is used for both ‘how much’ and ‘how many’, though ‘how many’ requires a classifier or other word indicating the entity asked about. Lao and Shan Gyi also require a word indicating the type of entity for their own varieties of ‘how many’, though I could not find whether they had a separate interrogative for ‘how much’. For ‘how’, Shan-Ni, Phake and Aiton each use a variation of nang¹hɯ⁴, while in the other languages only use the second syllable of that compound (e.g. hɯ in Shan Gyi) feeling out nang¹ ‘that’.

For polar questions, all Tai languages have a different strategy. In all languages mentioned except for Shan-Ni and Khamti it is also possible to ask questions without a questions particle; however, how common this is and in which situations this is acceptable differs and is not comparable. Despite this possibility, all languages do have a polar question particle. Shan Gyi, Phake and Aiton use a variety of nɔ for polar questions, which may have derived from Burmese tag question ေနာ် nɔ ‘isn’t it’. In Shan-Ni this is only used as a tag question, while polar questions obligatorily end in le⁴/le². The Lao particle bòò³ is the most neutral of several polar question particles (Enfield 2007: 43). In Khamti the polar question particle is nai⁵, which is also obligatory marked behind every other question word, expect for kii⁵ ‘how many’ and heu⁵leu⁵sii⁵ ‘how’. Looking at the interrogatives, the relation between Shan-Ni, Phake and Aiton is clear. Innovations from Tai Mau are also found or developed further in these three languages, but not in Khamti. Shan Gyi and Lao are clearly more distant.

### 7.3.3 Demonstratives

Demonstratives in Shan-Ni have gone through two processes. Like the interrogatives, they show a few contractions, and like the nouns, they have become disyllabic. In some cases, both processes implied the addition of the same lexeme ti² ‘place’, for example in ti² thai², in which thai² itself is a contraction between ti² ‘place’ and naj⁵ ‘this’. The combination of these different processes allows Shan-Ni to be more specific in the way space is discussed. The contractions are described by Young (1985) for Tai Mao, and are also present in Phake, though none of these languages use disyllabic demonstratives.

<table>
<thead>
<tr>
<th>Shan-Ni</th>
<th>gloss</th>
<th>Tai Mau</th>
<th>gloss</th>
<th>Phake</th>
<th>gloss</th>
<th>Khamti</th>
<th>gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>ti² thai²</td>
<td>here</td>
<td>thài</td>
<td>here</td>
<td>than³</td>
<td>there</td>
<td>an²-nai¹</td>
<td>here.SG</td>
</tr>
<tr>
<td>ti²than²</td>
<td>there (near)</td>
<td>than</td>
<td>there</td>
<td>that³</td>
<td>here</td>
<td>a-meau⁴-nai¹</td>
<td>there.SG</td>
</tr>
<tr>
<td>pun⁴than²</td>
<td>there (far)</td>
<td>nài</td>
<td>this</td>
<td>nan⁴</td>
<td>that</td>
<td>an²-pun²-nai¹</td>
<td>distant.SG</td>
</tr>
<tr>
<td>nang¹ne²</td>
<td>this</td>
<td>nân</td>
<td>that</td>
<td>nai⁴</td>
<td>this</td>
<td>an²-nai¹-khau⁵</td>
<td>here.PL</td>
</tr>
<tr>
<td>pun⁴nan⁵</td>
<td>that</td>
<td>lan⁴</td>
<td>that</td>
<td>a-meau⁴-nai¹-khau⁵</td>
<td>there.PL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>an²nai¹</td>
<td>this thing</td>
<td>lai⁴</td>
<td>this</td>
<td>an²-pun³-nai¹-khau⁵</td>
<td>distant.PL</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

120
Phake uses the words *thai³* ‘here’ and *than³* ‘there’ similar to Shan-Ni, but does not combine them with *ti²* or *pun⁴*. In Morey (2005a) *pa¹lan³* ‘that side’ and *ti² nan⁵* ‘that place’ do occur, though monosyllabic determiners are more common in his Phake data and combinations with *than³* ‘there’ or *thai³* do not occur. Similarly, Phake uses the monosyllabic words *nai⁴* and *lai⁴* for ‘this’ and *nan⁴* and *lan⁴* for ‘that’, while these are disyllabic in Shan-Ni. Morey does not mention word-initial /n/ and /l/ to be allophones in Phake, while they certainly are in most words in Indawgyi Shan-Ni. The acceptance of both variants in Phake for the demonstratives is interesting, and they might be revealing a connection at an earlier stage where /n/-/l/ allophony was less extensive then it is now in Indawgyi Shan-Ni.

Khamti has developed a totally different system and includes *an³* ‘thing’ in all demonstratives, and also marks number by adding *khau⁵* ‘they’ for plural objects. Clearly, Shan-Ni has extended the Tai Mau system with the formation of disyllabic terms. However, the distinction Shan-Ni makes between ‘here’, ‘there (near)’ and ‘there (distant)’ is shared with Khamti, though it uses different terms to express that. However, Shan-Ni does not mark plurality with demonstratives like Khamti does.

**7.3 TAM markers**

TAM markation in Tai languages is rare. Table 37 shows an overview of the different TAM markers across Tai languages. The extensive TAM markation present in Shan-Ni is the result of a longterm development, which to different extents can be found in other Tai languages. In chapter 6, TAM markers were discussed in relation to Phake and Aiton, as described by Morey (2005a, 2008). Some of these TAM markers are also present in Khamti, Tai Mau and Shan Gyi. As these are grammaticalized words also exist in their ungrammaticalized form in Shan-Ni and other Tai languages, we can also expect these developments to at different stages in different languages.

Most sentences in Shan Gyi and Tai Mau do not have TAM markation, hence tense, aspect and mood can then be filled in by the interpretation of the listener. They do use the words for ‘already’ and ‘will’ in similar ways Shan-Ni uses them. I also found the words for ‘keep’ and ‘stay’ in dictionaries as ‘verbal particles’. In some texts they do occur in ways that could be interpreted as having similar functions the TAM markers in Shan-Ni, though these could also be interpreted as parts of serial verb constructions. I did not find evidence convincing enough to call them TAM markers, not in the last place because these are rare. Of course, their function within a serial verb construction could be the first step to further grammaticalization into a TAM marker, but they clearly did not reach that state yet.

The markers *maa⁴* and *kaa⁵* in Shan-Ni as inchoative and terminative aspect markers, mostly used for events in the past. In Phake and Khamti these are also present as past markers. Inglis (2014) glosses *kaa⁵* refers to distant time and *maa⁴* to the perfective. Unfortunately, he does not elaborate on it and does not include enough examples to see how

---

29 Narratives in Young (1985) and Glick & Sao Tern Moeng (1991)
the usage of these markers in Khamti corresponds with the usage of TAM markers in Shan-Ni.

The shape of these TAM markers is also interesting, since these in Shan-Ni these differ from the verbs they derived from. The TAM markers themselves share the same shape as they have in Khamti and Phake, but in these languages they also exist as normal verbs for ‘go’ and ‘come’. These verbs in Shan-Ni have different shapes, corresponding with Tai Mao.

### Table 37 The presence of Shan-Ni TAM markers in other languages

<table>
<thead>
<tr>
<th>Shan-Ni marker gloss</th>
<th>Shan-Ni gloss</th>
<th>Phake marker</th>
<th>Phake gloss</th>
<th>Khamti marker</th>
<th>Khamti gloss</th>
<th>Tai Mau marker</th>
<th>Tai Mau gloss</th>
<th>Shan Gyi marker</th>
<th>Shan Gyi gloss</th>
</tr>
</thead>
<tbody>
<tr>
<td>marker gloss</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kaw¹</td>
<td>OLD</td>
<td>yau⁴</td>
<td>yau⁴</td>
<td>yåu already</td>
<td>jåw⁵</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kɤn¹</td>
<td>SHOULD'/YET</td>
<td>uu¹</td>
<td>uu¹</td>
<td>-</td>
<td>ju²</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>u¹</td>
<td>STAY</td>
<td>û</td>
<td>STAY</td>
<td>uu¹</td>
<td>-</td>
<td></td>
<td></td>
<td>ju²</td>
<td>waj⁵</td>
</tr>
<tr>
<td>waj⁵</td>
<td>KEEP</td>
<td>wai¹</td>
<td>KEEP</td>
<td>-</td>
<td>wâi</td>
<td>keep</td>
<td></td>
<td>waj⁵</td>
<td>keep</td>
</tr>
<tr>
<td>te¹</td>
<td>FUT</td>
<td>ta¹</td>
<td>WILL</td>
<td>tii⁵</td>
<td>tî</td>
<td>will</td>
<td></td>
<td>te⁴</td>
<td>will</td>
</tr>
<tr>
<td>kaa⁵</td>
<td>GO</td>
<td>kâ¹</td>
<td>GO</td>
<td>kaa⁵</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>maa⁴</td>
<td>COME</td>
<td>mâ²</td>
<td>COME</td>
<td>maa⁴</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>kaw¹</td>
<td>OLD</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>kɤn¹</td>
<td>SHOULD'/YET</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

### 7.3.1 kaw¹ ‘OLD’ and kɤn¹ ‘SHOULD’/’YET’

The grammaticalisation of two TAM markers, kaw¹ ‘OLD’ and kɤn¹ ‘SHOULD’/’YET’ seems to be unique to Shan-Ni. They occur only marginally as lexical, ungrammaticalized items in other Tai languages. In Morey (2005a), both are mentioned as lexical items in lists cited from Banchob (1987), but they do not occur in any of Morey’s own examples. They are also not found in Khamti (Inglis 2014). They do both have Burmese equivalents that are frequently used in colloquial speech. Although these do not correspond entirely with Shan-Ni and have grammacalized from different words than their Shan-Ni equivalents, the fact that Shan-Ni speakers feel the need to have a TAM marker in the same situations may indicate that this a result of widespread Shan-Ni/Burmese bilingualism. Of these two markers, kɤn¹ ‘SHOULD’/’YET’ is used the most frequently. It is used in two ways, with and without negator. Without negator it is used for request from the speaker to the interlocutor. With negator it is used for events that have not happened yet or things that the subject did not do yet. The usage with negator is the most frequent and corresponds with the Burmese particle [ ေသး ] thee.

The examples below are shown together with their Burmese translations in the Basic Spoken
Shanni book. In (222) and (223) it is about things that the speaker did not do yet, and (224) refers to the time before sunrise, which is described as ‘the sky is not bright yet’. The usage of \textit{ma¹.kɤn¹} ‘not yet’ corresponds the Burmese \textit{ma...thee buu} ‘not yet’ in all instances I found, both in spoken and written examples given by bilinguals. The function of \textit{kɤn¹} without negation in Shan-Ni does not correspond with a specific Burmese particle, so it is not entirely the same, but the \textit{ma¹.kɤn¹} construction does allow Shan-Ni speakers a express a very frequently used particle in Burmese.

\begin{tabular}{ll}
\textbf{222} & a. \textbf{[SHAN-NI]} & b. \textbf{[BURMESE]} \\
 & \textit{\begin{tabular}{l}
ma¹ lai² yap⁵ kɤn¹  \\
\text{NEG GET SOW YET} \\
\text{‘[I] didn’t get it sowed yet’} \\
\end{tabular}} & \textit{\begin{tabular}{l}
ma \text{khyup thee buu}  \\
\text{NEG SOW YET NEG} \\
\text{‘[I] didn’t get it sowed yet.’} \\
\end{tabular}} \\
\text{BSS_33} & \\
\textbf{223} & a. \textbf{[SHAN-NI]} & b. \textbf{[BURMESE]} \\
 & \textit{\begin{tabular}{l}
ma¹ laj² sɯ⁵ kɤn¹  \\
\text{NEG get buy YET} \\
\text{‘[I] didn’t buy it yet.’} \\
\end{tabular}} & \textit{\begin{tabular}{l}
ma \text{we ya thee buu}  \\
\text{NEG buy get YET NEG} \\
\text{‘[I] didn’t buy it yet.’} \\
\end{tabular}} \\
\text{BSS_74} & \\
\textbf{224} & a. \textbf{[SHAN-NI]} & b. \textbf{[BURMESE]} \\
 & \textit{\begin{tabular}{l}
phaa⁵ ma¹ lɛng⁴ kɤn¹  \\
\text{sky NEG light YET} \\
\text{‘before sunrise’} \\
\end{tabular}} & \textit{\begin{tabular}{l}
moe ma \text{leng thee buu}  \\
\text{sky NEG bright YET NEG} \\
\text{‘before sunrise’} \\
\end{tabular}} \\
\text{BSS_66} & \\
\end{tabular}

For Phake \textit{küŋ¹} ‘should’ (cognate of \textit{kɤn¹} ‘SHOULD’ in Shan-Ni) is listed as an auxiliary verb in Morey (2005a: 234) citing Banchob (1987). His work does not feature examples of the usage of \textit{küŋ¹} in Phake. For ‘not yet’, there is one example by Morey (2005a: 350) with the word \textit{paj¹}. As far as I know this word does not exist in Shan-Ni. Later Morey (2005a: 259) does mention “\textit{pai1/pi1/pa1} ‘don’t’, ‘not yet’”. If this is indeed the same lexeme, it only exists as the prohibitive \textit{pi¹} in Shan-Ni.

The TAM marker \textit{kaw¹} ‘OLD’ is used in sentences that in English would include ‘ever’, ‘X times’ or ‘never’, but searching on these terms through the work of Morey (2005a) and Inglis (2014) did not yield any results, hence I do not know how Khamti, Phake and Aiton express these kind of constructions. \textit{kaw¹} ‘OLD’ does correspond with the Burmese marker \textit{hpuu} \text{[ɪ_k̂]}, itself a grammaticalization of the verb ‘visit’. Below are the Shan-Ni sentences from the
section on kaw¹ in Chapter 6 and their Burmese equivalents. In (225) ma¹.kaw¹ is used as ‘never’, in (226) kaw¹ is used for ‘ever’ and in (227) sɔng³ ha¹...kaw¹ is used for ‘two times’.

(225) a. [SHAN-NI]

kɛɯ⁴ caa⁵ man⁴ ma¹ han³ kaw¹  
1SG for 3SG NEG see OLD
'I have never seen her before.'

b. [BURMESE]

chendo tu go ma mjen hpuu buu  
1SG.M 3SG OBJ NEG see EVER NEG
'I have never seen her before.'

(226) a. [SHAN-NI]

nyaa⁴ yaan¹kong¹ thɤng³ kaw¹ de⁴  
MEET yangon arrive OLD Q
‘Have you ever been to Yangon?’

b. [BURMESE]

yangon jauk hpuu laa  
yangon arrive EVER Q
‘Have you ever been to Yangon?’

(227) a. [SHAN-NI]

ɤ⁴  sɔng³ ha¹...kaw¹  
EMP two CLF times arrive OLD
‘Yes, I have been there twice.’

b. [BURMESE]

ho hni khaa jauk hpuu de  
yes two time arrive EVER REAL
‘Yes, I have been there twice.’
The lack of these frequently used TAM marker in Phake and other Tai langauges means innovation must have taken place after the migration of the Phake in 1775. From that moment onwards the Burmese administration has been very influential in the Shan-Ni areas, hence this is very likely a result of bilingualism.

7.4 Summary

Shan-Ni shares features with the Tai languages to the northwest (e.g. Khamti, Phake and Aiton), as well as with the Tai languages to the southeast (e.g. Tai Nua and Tai Mao). While the presence of TAM markers is clearly shared with the northwestern languages, the interrogatives are similar to Tai Mao, as well as the retention of the cluster word initial clusters /kw/ and /khw/. Class terms are present in other languages, though never as common as they are in Shan-Ni. The addition of class terms is responsible for a large chunk of Shan-Ni’s disyllabic vocabulary. But the major application of the formation of disyllabic terms can be found in grammatical features. Shan-Ni favours the inclusion of all innovations, both those shared with the northwestern languages, and those shared with the southeast.
This is well represented in the demonstratives, which include the contractions present in Tai Mao and Phake, but through which disyllabic terms can also distinguish between different degrees of proximity like Tai Khamti does. This tendency to create disyllabic terms also allows Shan-Ni to distinguish between personal and possessive pronouns for the second and third person plural, which are the same in Khamti and Tai Phake. Similarly, the TAM markers kaa⁵ ‘go’ and mao⁴ ‘come’ are identical to the verbs they derived from in Khamti and Tai Phake, but have different shapes in Shan-Ni. These features together show how Shan-Ni has a strong tendency to make distinctions in the language, that go beyond the incorporation of features also present in other different Tai languages.

Though data on some of the languages mentioned above is lacking, Phake, Khamti and Tai Mao show clear relationships to Shan-Ni in different ways. A similar set of innovations from Tai Mao are present in both Shan-Ni and Tai Phake. However, features absent in Tai Mao but present in Tai Khamti are as well. This supports the theory that a large group of migrants from the Tai Nue/Tai Mao region integrated into the sedentary Tai group living in the area of Moguang, who are related to the Khamti. The Tai Phake, who probably left Moguang around 1775, share a lot of features with Shan-Ni, although the formation of disyllabic words is not that frequently presented in Tai Phake. The relationship between Shan-Ni, Khamti, Phake and Tai Mao is demonstrated in Figure 10.
Conclusions

Shan-Ni as spoken at the Indawagyi lake shares features with the Tai languages spoken in India, the Chinese-Burmese border, and Khamti further in the north of Myanmar. It also has some very unique features, which in other Tai languages are either absent or only marginally present. Though all the languages spoken in India and northern Myanmar have distinct features that may be read as Tibeto-Burman influence, Shan-Ni seems to have a few more recent developments that are more explicitly Burmese influence. In addition, it has taken some processes that exist in other Tai languages and built new processes on top of that. These developments point to several moments in history where groups merged or split off. Since their movements have been all either towards or from Mogaung, where the Shan-Ni live, knowledge of Shan-Ni is crucial in understanding the developments within the Southwestern Tai language family.

The phonology of Shan-Ni is similar to other Tai languages, with aspirated and non-aspirated voiceless stops, no voiced stops, a variety of nasals, fricatives and approximants, and a large vowel inventory. It has retained the initial consonant cluster /kw/, and palatalized the consonants /k, kh, p, ph, m/ when followed by the vowels /ɛ, i, e/, in which did not happen in other Tai varieties in the region is likely to be a relatively new innovation.

Shan-Ni is actively making nouns disyllabic through the addition of class terms. These are words that provide categories to a noun, for example ‘round’ or ‘time’ or ‘fruit’. These then form compounds, often with the old form, and then gradually reduce the phonology of the class term, making it a sesquisyllabic word of which the meaning of the first syllable is no longer transparent. A similar process is applied to determiners, but here different class terms added to the same word allows Shan-Ni to make additional distinction when talking about place. Several word classes, including determiners, question words and adjectives, incorporate shapes of words that have been formed through processes active in Tai Mao, but have developed further or fossilized in Shan-Ni.

The sentence structure in Shan-Ni is affected by ellipsis, the omission of arguments, and the presence of grammatical markers, including TAM markers. Ellipsis happens when the referent is active and clear, hence does not need to be mentioned anymore. New referents are often introduced with a classifier construction, and inactive referents are marked by the definite marker naj⁵. Possession can be expressed either through predicates with the existential verb yang⁴ or by putting the personal pronoun referring to the possessor behind the possessed. Similarly, the same words that are indefinite pronouns at the beginning of a phrase, are question words when placed at the end. Grammatical markers include a linker, a verbal plural marker, a prohibitive and others that each in their own way modify the sentence structure.

The TAM markers mostly correspond with the Tai Phake and Khamti, though the functions slightly differ and Shan-Ni has two additional TAM markers that are not found in any other Tai language, (ma¹...) kən⁴ ‘should/not yet’ and kəw¹ ‘has been experienced in the past. These do correspond to the function of the particle tee and hpuu in Burmese, hence it looks like a
recent innovation.

When looking at the representation of these features in other Tai languages, it seems very likely that the Shan-Ni are descendents of two Tai groups that merged around 1200 around Mogaung. The sedentary population probably spoke a language similar to Khamti, while a large group of migrants coming in spoke a language similar to Tai Mao and Tai Nua. The language most similar to Shan-Ni is Tai Phake, who left Mogaung in 1775 to Assam. The differences between Shan-Ni and Phake are thus developments that have occurred in the last 250 years. These include the extention of some pre-existing phenomena, but also some things are very clearly Burmese influence, which makes sense because from that time onwards the Burmese became dominant in previously Shan-Ni strongholds.

Some innovations in Shan-Ni still allow variation in certain items, making language change processes more transparent. Other innovations are regional, hence these too ask for a deeper study into language change. What this study has pointed out, is that Shan-Ni is a crucial link to understand the connections between Tai languages in northern Myanmar, northeast India, and southwestern China. This is not a linear relation, but rather one of various periods of migration causing populations to meet and separate. Since they were also at the center of a larger kingdom, which included different Tibeto-Burman peoples, a description of Shan-Ni allows us to study how these languages have been influenced by Shan-Ni and vice-versa. Tibeto-Burman studies that include comparison with “Shan” up to now only had Khamti and Shan Gyi to look at, which are both quite distinct from Shan-Ni grammatically.

This thesis only scratched the surface of Shan-Ni grammar and its possible contributions to understanding language change in the region. Future research can go both broader and deeper into what has been introduced here. Broader with the inclusion of Shan-Ni dialects spoken in other regions, deeper with in-depth research on individual features discussed here.
References


## Appendix

### Wordlist

<table>
<thead>
<tr>
<th>∞ - k</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ဘာမည်ကျော်ကြား</td>
<td>ka¹<code>kyi⁴</code>eng²</td>
<td>armpit</td>
</tr>
<tr>
<td>ဘ</td>
<td>kaa²</td>
<td>dance</td>
</tr>
<tr>
<td>ဘေ</td>
<td>kaa⁴</td>
<td>car</td>
</tr>
<tr>
<td>ဘု</td>
<td>kaa⁵</td>
<td>GO (TAM)</td>
</tr>
<tr>
<td>ဘဒာ</td>
<td>kaan⁵</td>
<td>break off</td>
</tr>
<tr>
<td>ဘု</td>
<td>kaap⁵</td>
<td>bite</td>
</tr>
<tr>
<td>ဘာ</td>
<td>kaj¹</td>
<td>chicken</td>
</tr>
<tr>
<td>ဘာဝ</td>
<td>kak¹</td>
<td>stutter</td>
</tr>
<tr>
<td>ဘာက</td>
<td>kam¹</td>
<td>brown</td>
</tr>
<tr>
<td>ဘာဦး</td>
<td>kam²`cum⁴</td>
<td>ball pen</td>
</tr>
<tr>
<td>ဘာဗု</td>
<td>kau⁴ /kεu⁴</td>
<td>1SG</td>
</tr>
<tr>
<td>ဘာကြီး</td>
<td>kan²</td>
<td>starve</td>
</tr>
<tr>
<td>ဘာကြီး</td>
<td>kan⁴</td>
<td>verbal plural (VPL)</td>
</tr>
<tr>
<td>ဘာချီ</td>
<td>kang²</td>
<td>cover</td>
</tr>
<tr>
<td>ဘာပုံ</td>
<td>kap¹</td>
<td>communicate</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>kat¹</td>
<td>cold</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>kaw¹</td>
<td>old</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>ko⁵</td>
<td>CLF for humans</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>ko⁵</td>
<td>also</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>ko¹</td>
<td>classifier for cups</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>kong²`kaang⁴</td>
<td>chin</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>kong²`kha⁴</td>
<td>throat</td>
</tr>
<tr>
<td>ဘာစော</td>
<td>kɔp⁵`sang⁴</td>
<td>why</td>
</tr>
<tr>
<td>ဘာဗု</td>
<td>kɔn¹</td>
<td>SHOULD / YET</td>
</tr>
<tr>
<td>ဘာဗု</td>
<td>ko⁴</td>
<td>afraid</td>
</tr>
</tbody>
</table>
koj²  banana
kon²  buttocks
kon⁴  person
ku²  classifier for pairs
kuj¹  cotton
kwaa¹  go
kwak¹  cup
kyaang¹  fast
kyep¹  pick up
kyin⁴  eat

kha¹laaj²  want
kha¹lau⁴  how much, how many
kha¹nang³  later
khaa²  politeness marker
khaa³  search
khaa³  upper leg
khaam²  cross (verb)
khaan⁵  lazy
khaj¹  egg
khaj²  want
khat¹  broken
khaw¹  enter
khaw²  rice
khaw³  3PL.POS, their
khɔ² mɯ⁴ wrist  lit: useful hand
khɔ²tin⁴  ankle  lit: useful foot
khɔ²tin⁴  slipper
khɔ⁴ neck
khɔj⁴ male genitals
khew³ green
ki³ female genitals
khɯn² attend
khun² climb
khuj² banana
khun³ho³ hair
khun³mu⁴ hair on the hand
khun³ping¹taa⁴ eyelashes
khun³taa⁴ eyebrows
khun³tin⁴ hair on the foot
khwaaj⁴ buffalo
khwam⁴ language
khyem³ needle
khyet¹ frog
khyew² teeth
khyen³ mu⁴ forearm
khyeng³ stiff
khyep¹ classifier for classifier for flat, hard objects
khyi¹ ride, drive
khyi²hu³ ear wax
khyi²khyew² teeth impurities
khyi²muk⁵ snot
khyi²taa² eye boogers

ng - ng

ngun⁴ money
ngo⁴ cow
ngu⁴ snake

∞ - c

caa⁴ for
caa⁵ angry
cau⁴ order
cau⁴ heart
cau⁴ maj² worried
cau⁴ on² happy
cau⁴ yau¹ angry
com⁴ follow
cong³ little
cet¹ seven
cum¹pu⁴li⁴ navel  lit: sink count navel

缰 - s

saaj² intestines
saaj²:nu⁴ small intestine  Lit: intestine small
saaj²king² appendix  Lit: intestine branch
saaj²long³ large intestine  Lit: intestine big
saaj²tap¹ inner organs  Lit: intestines liver
saam³ three
saan³ inside
saan⁴ beautiful
sang⁴ what, something
sap¹ strike
sap¹pa¹tu⁴ window
sap¹pa¹tu⁴ door
soj¹ glass pot
sɔk¹ mɯ⁴ elbow lit: elbow hand
sɔn³ study
sɔŋ³ two
sɔt¹ alight, burn
sɔt⁵ drink
səŋ³ gem
səŋ³khew³ jade
sə² shirt
si¹ four
si¹ colour
si¹khew³ green
si¹ hɛng³ yellow
si¹nɛng⁴ red
si³ LINKER, because, that, and
sip¹ ten
siw³ pimple
su⁵ buy
sut⁵ push
son²tin⁴ heel lit: heel foot
sop¹ mouth
su³ 2PL .POS, your
suk⁵ rinse

ny - ny
Nyaa⁴ meet, at

- t
taa¹koong¹ Yangon (city)
taa⁴tin⁴ bowl of the ankle lit: eye foot
taaj⁴ shout
taa⁴ die
taan² talk
taang⁴ way, manner
taap¹kon² hips lit: flat side buttocks
tau² classifier for bottles
tang² cook
tang⁴ with
tang⁴long³ everything
tap¹ liver
tap¹phɤk¹ spleen Liver
tɔ¹ talk
tɔ¹ until	tɔj⁴ look
te¹ FUTURE (TAM), will
te⁴ko⁵ friend
te¹ build
ten¹ hit
ten² forest
ten²maj⁵ tree
ti² place
ti²thai² here
ti²than² there
tin⁴ foot
to⁴ classifier for animals
to⁴ body
tok¹ fall
traa⁴ Buddhist law
tuk¹ fall
∞ - th

- thaa⁴ swear
- theu⁴ where
- thɔng² classifier for trees, plants
- thon² arrive
- thɔp¹ encounter, meet
- thun³ classifier for bags
- thung³ bag

n - n

- na¹hu⁴ how
- naa²cau⁴ chest (female) lit: face feeling
- naa²khyɛng² shin
- naa²ok¹ chest (male) lit: face chest
- naa²phaak¹ forehead
- naa⁵ TAG, isn't it?
- nam³ many
- nam⁵ water
- nam⁵leng²waan³ tea (sweet)
- nam⁵naaj⁴ spit lit: water saliva
- nam⁵naaj⁴khum³ saliva lit: water saliva ?
- nam⁵taa⁴ tears lit: water eye
- nang¹hu⁴ how
- nang¹ne² this
- nang² sit
- nang³ back
- nɔn⁴ sleep
- nɔng³ lake
ne¹ plural

c - p

pa³laa³ almost
pa'ying⁴ girl
paa² side
paaj³ run
paaj⁴ mun⁴ top of the hand    lit: top hand
paaj⁴ tin⁴ top of the foot    lit: top foot
paak¹ mouth
paan⁵ plate (flat)
pak¹ hundred
pa² father
poj⁴ fair
pak⁵ return
pom² round
pon²kaaw² calf (of leg)    lit: joint calf (of leg)
pot¹ lungs
pen⁴ COPULA, happen, be
pe⁵ can
pi¹ PROHIBITIVE, don't
pi³nong⁵ sisters
pi⁶ year
pin⁴ COPULA, happen, be
ping¹sop¹ lips
pu⁴ if
pu⁴nang⁵ that
pu⁴than² over there
pom¹ story
pon⁵  last, previous
pu¹  grandfather
pu¹ta¹ya²  grandparents
put¹  open
pyaa⁴  show
pyen⁴  COPULA, happen, be
pyet¹  duck
cyet¹  eight

æ - ph

pha¹lo³  onion
pha¹maaw¹  bachelor
phaa¹kyem²  cheek  lit: palm/sole hand
phaa¹muu⁴  hand palm  lit: sole/palm cheek
phaa¹tin⁴  sole  lit: palm/sole foot
phaa²  cloth
phaa² ho³  turban
phaa² sɤ¹  bedsheets
phaa² taj⁴  traditional clothes
phaa³  stone plates
phaa³kaan⁵  Hpakant (city), stone plates break off
phaa⁵  sky
phaa⁵lɛng⁴  daylight, sunshine
phaan³  poor
phaj⁴  fire
phak¹  dish next to rice
phan³  dream
pheur⁴  who
phyn⁴  bee
classifier for flat, flexible objects
husband
couple (husband and wife)
temple
pagoda
rain
to quarrel

NEG
3PL, they
mango
3PL
mad
come
horse
widow
fruit
chillies
orange
heart
kidneys
coconut
village
new
classifier for spoons
3SG (POS), he/she
pot
mɔj³sop¹ moustache lit: (short facial) hair mouth
mɔk¹ inform
mɔk¹ classifier for flowers
mɔk¹ya² flower
mɔng⁴ hour
me⁴ wife
men⁴ fly
mε² mother
mε² lady
me⁴ wife
mɛɯ⁴ 2SG, you
mɛ² time
mɛ² maj² hot season
mɛ² maj² now
mɛ² phok⁵ tomorrow
mɛ²mɛ²ngaa⁴ yesterday
mi⁴ rich
mu² lew⁴ when
mu² thɛɯ⁴ when
mu⁴ hand
mɯn⁴ 3SG (POS), he/she, his/hers
mu² pig
myet¹ fishing hook
myin⁵ turmeric
myit⁵ knife, slice (verb)

yaa¹ at, place
yaa² grandmother
| မအိုး | yaap¹ | difficult |
| မအိုး | yaaw⁴ | long |
| မအိုး | yam⁴ | wet |
| မအိုး | yam⁵ | chew |
| မအိုး | yang⁴ | EXIST, have |
| မအိုး | yau⁵ | finish, FINISHED (TAM) |
| မွေးအိုး | yew² | urine |
| မိုမိမိ | yuuk⁵ | wet |
| မိမိ | yu¹ | stay |
| မိမိ | yau¹ | big |
| မိမိ၀င် | yu⁴ | shoot |
| မိမိတွင် | yu³ga¹ | illness |

### ၃-၁

| မဝိဝိ | la¹n¹ | child |
| မဝိဝိ | la¹ling⁴ | monkey |
| မိုမိမိ | laa² | face |
| မိမိ၀င် | laang⁵ | wash |
| မိမိ၀င် | laap⁵ | free |
| မိမိ၀င်၀င် | laj² | GET |
| မိမိဒိုင် | lau⁴ | in |
| မဝိဝိးမိုမိမိ | lang³mu⁴ | back of the hand | lit: back hand |
| မဝိဝိးမိုမိမိ | lang³tin⁴ | back of the foot | lit: back foot |
| မိုမိမိ၀င် | loj⁴ | slow |
| မိမိ၀င်၀င် | loj⁴ | mountain |
| မိမိ၀င်၀င် | long² | work |
| မဝိဝိးမိုမိမိ | lot¹ lom⁴ | trachea | lit: pipe wind |
| မဝိဝိးမိုမိမိ | lot¹yɤ² | esophagus | lit: pipe food |
| မဝိဝိးမိုမိမိ | lot⁵ | escape |
lep⁵tin⁴ toenail  
leŋ⁴ which  
leŋ⁴ lit: nail foot  
leŋ⁵ give birth  
leŋ⁵mɯ⁴ lit: nail hand  
lɤ⁵ meat  
lɤn⁴ month  
lɤt⁵ blood  
lɯ⁴ good  
lɪk⁵ write  
lɪn⁴ earth  
lɪng⁵ lit: finger hand  
lɪw⁵mɯ⁴ finger  
lɪw⁵tin⁴ lit: finger foot  
long³ big  
lu³ particle implying plea  
lu³ mouse  
luk⁵ daughter, son  
luk⁵ get up, stand up  
luk⁵caai⁴ son  
luk⁵⁰n¹ child  
luk⁵sao³ daughter  

waa² speak, talk  
waam⁵ catch  
waan¹ classifier for round, open objects  
waan³ sweet
waan⁴ day
waj⁵ keep, KEEP (TAM)
wan¹ plate (round)
wan¹ma¹laj⁵ today

haa² five
haaj² cry
haaj³ loose
hak⁵ love
hau⁴ give
han³ see
haw⁴ 1PL (POS), we, our
hɔng⁵ call
heng³ thousand
het¹ do, make lit: shape waist
heng² dry
heng⁴ strength
heng⁵ɛng² waist
hɔ⁴ sweat
hɔ⁵ boat
hɔ⁶men⁴ airplane lit: head knee
hɔ⁴ house
ho³ head lit: spacious.place rice
ho³khaw¹ knee
ho³maa¹ shoulder includes ear
hong⁴khaw² stomach
hu³ ear
hu³nang⁴ nose lit: hole buttocks
hu⁴ classifier for holes, drops, dots
<table>
<thead>
<tr>
<th>Syllable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>hu⁴</td>
<td>hole</td>
</tr>
<tr>
<td>hu⁴kon²</td>
<td>anus</td>
</tr>
<tr>
<td>huj¹</td>
<td>classifier for round, solid objects</td>
</tr>
<tr>
<td>huj¹taa⁴</td>
<td>eye</td>
</tr>
<tr>
<td>hung⁴kha¹laaj³</td>
<td>stairs, ladder</td>
</tr>
<tr>
<td>hung⁵hen</td>
<td>school</td>
</tr>
</tbody>
</table>

### a (vowels)

<table>
<thead>
<tr>
<th>Syllable</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>a¹saak⁵</td>
<td>age</td>
</tr>
<tr>
<td>a¹sang⁴</td>
<td>now</td>
</tr>
<tr>
<td>aaj⁴</td>
<td>very</td>
</tr>
<tr>
<td>aaj⁴te⁴kɔ⁵</td>
<td>best friend</td>
</tr>
<tr>
<td>aan¹</td>
<td>count</td>
</tr>
<tr>
<td>an²</td>
<td>thing</td>
</tr>
<tr>
<td>an²hak⁵</td>
<td>love</td>
</tr>
<tr>
<td>an²kyin⁴</td>
<td>food</td>
</tr>
<tr>
<td>an²taaj⁴</td>
<td>death</td>
</tr>
<tr>
<td>ang¹</td>
<td>for</td>
</tr>
<tr>
<td>aw⁴</td>
<td>take</td>
</tr>
<tr>
<td>ɔk¹</td>
<td>leave</td>
</tr>
<tr>
<td>ɔk¹ɛk¹</td>
<td>brain</td>
</tr>
<tr>
<td>ε²</td>
<td>mushy</td>
</tr>
<tr>
<td>u¹</td>
<td>stay, STAY (TAM)</td>
</tr>
<tr>
<td>un¹</td>
<td>warm</td>
</tr>
<tr>
<td>up⁵</td>
<td>classifier for books</td>
</tr>
</tbody>
</table>