A CORPUS-BASED STUDY OF THE FUNCTION AND STRUCTURE OF NUMERAL CLASSIFIER CONSTRUCTIONS IN S’GAW KAREN

by

Collin Olson

A Thesis Submitted to the
Faculty of the Graduate School at
Middle Tennessee State University
in Partial Fulfillment
of the Requirements for the Degree of
Master of Arts
in English

Murfreesboro, TN
May 2014

Thesis Committee:

Dr. Aleka Blackwell

Dr. Tom Strawman
ACKNOWLEDGEMENTS

I could not have completed this project without the patience and guidance of Dr. Aleka Blackwell. It was in Dr. Blackwell’s classes that I first decided to pursue a career in linguistics, and I was fortunate to serve as her research assistant while completing my MA. I am very thankful to have had an advisor who has stuck with me through what became a complicated and prolonged project.

I give many thanks to Dr. Tom Strawman, who has been a good friend and mentor for the past fourteen years.

The S’gaw Karen community in Smyrna, Tennessee has been instrumental in gathering the data for this project and helping me to learn about the S’gaw Karen language, culture, and history. Working with the S’gaw Karen community as an ESL teacher has been a rewarding experience both professionally and personally, and I hope that the S’gaw Karen people thrive in their new home.

Most importantly, I give thanks for the support from my family. I am especially grateful for the patience and encouragement of my wife, Selena, and our two daughters.
ABSTRACT

This thesis examines the numeral classifier system in S’gaw Karen, a language in the Tibeto-Burman subfamily. Numeral classifiers are an areal feature of Sino-Tibetan languages. This thesis contributes to our understanding of numeral classifier systems in this language family by examining both the function of numeral classifiers in S’gaw Karen and the structure of noun phrases when numeral classifiers are present.

The data for this project include a phonetically transcribed corpus of four stories created in consultation with three native speakers from Eastern Burma and targeted elicitations.

A detailed phonological analysis of the S’gaw Karen variety spoken by the consultants is included. In S’gaw Karen, numeral classifiers individuate and enumerate the noun, specify salient semantic features of the noun, and are used as anaphoric devices. The syntactic structure of noun phrases involving numerals and classifiers are examined in several different NP contexts, including those with modifiers and determiners.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>LIST OF TABLES</td>
<td>x</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>xi</td>
</tr>
<tr>
<td>LIST OF ABBREVIATIONS</td>
<td>xii</td>
</tr>
<tr>
<td>CHAPTER ONE: PURPOSE &amp; INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Purpose</td>
<td>1</td>
</tr>
<tr>
<td>1.2 An Introduction to the Karen Peoples &amp; Languages</td>
<td>3</td>
</tr>
<tr>
<td>1.2.1 Who Is Karen?</td>
<td>3</td>
</tr>
<tr>
<td>1.2.2 Population of Ethnic Karen &amp; Karenic Speakers</td>
<td>7</td>
</tr>
<tr>
<td>1.2.3 Geographic Location of the Karen People</td>
<td>9</td>
</tr>
<tr>
<td>1.2.4 Modern History of the Karen People</td>
<td>11</td>
</tr>
<tr>
<td>1.2.5 Current Status of the Karen People &amp; the Endangerment of the Karenic Languages</td>
<td>12</td>
</tr>
<tr>
<td>1.3 Introduction to the S’gaw Karen Language</td>
<td>14</td>
</tr>
<tr>
<td>1.3.1 Dialect Used in This Study</td>
<td>14</td>
</tr>
</tbody>
</table>
1.3.2 S’gaw Karen’s Language Family ....................................... 14

1.3.3 Literature Review of the Karenic Languages ....................... 16

1.3.3.1 Early Missionary Influence & Scholarship ......................... 16

1.3.3.2 Recent Research on Karenic Languages .......................... 18

1.3.3.3 Research on other Kareni Languages ............................. 21

CHAPTER TWO: METHODOLOGY & PHONOLOGY .......................... 23

2.1 Corpus Methodology .................................................................. 23

2.2 Phonological Description & Transcription Conventions ............... 25

2.2.1 S’gaw Karen Consonants ....................................................... 26

2.2.2 S’gaw Karen Vowels .............................................................. 29

2.2.3 S’gaw Karen Tones ................................................................. 31

2.2.4 S’gaw Karen Syllable Structure ............................................. 32

CHAPTER THREE: S’GAW KAREN NUMERAL CLASSIFIERS .......... 34

3.1 Literature Review on Numeral Classifiers ................................. 34

3.1.1 Overview of Numeral Classifiers .......................................... 34

3.1.2 Semantics of Numeral Classifier Systems .............................. 38

3.1.3 Typology of Numeral Classifiers ......................................... 41
3.1.4 Functions of Numeral Classifiers

3.1.4.1 Enumeration

3.1.4.2 Definiteness & Discourse Pragmatics

3.1.4.3 Classification

3.1.4.4 Anaphoric Use

3.1.5 Syntax of the Noun Phrase with Numeral Classifiers

3.2 The Numeral Classifiers System in S’gaw Karen

3.2.1 Introduction

3.2.2 Types of Numeral Classifiers in S’gaw Karen

3.2.3 Functions of Numeral Classifiers in S’gaw Karen

3.2.3.1 Enumeration

3.2.3.2 Anaphoric Use

3.2.3.3 Definiteness & Discourse Pragmatics

3.2.4 Structure of the Noun Phrase with Numeral Classifiers

3.2.4.1 The Classifier Constituent in the Noun Phrase

3.2.4.2 Demonstratives in the Noun Phrase

3.2.4.3 Adjectives in the Noun Phrase
3.2.5 The S’gaw Karen Numeral Classifier System within the Tibeto-Burman Family ................................................. 40

3.3 Discussion ........................................................................................................... 72

BIBLIOGRAPHY ........................................................................................................ 75

APPENDICES ............................................................................................................ 92

APPENDIX A: S’gaw Karen Corpus .......................................................................... 93
A1.1 “A Rabbit per Bush” ......................................................................................... 93
A1.2 “Bad Plans” .................................................................................................... 105
A1.3 “One Stupid Cat” ............................................................................................ 109
A1.4 “Ku Naw Lay & Naw Moo Aye” ....................................................................... 115

APPENDIX B: The S’gaw Karen Writing System .................................................. 129
A2.1 The S’gaw Karen Writing System .................................................................. 129
A2.2 S’gaw Karen Consonants .............................................................................. 132
A2.3 S’gaw Karen Consonant Clusters .................................................................. 133
A2.4 S’gaw Karen Vowels ...................................................................................... 134
A2.5 S’gaw Karen Tones ........................................................................................ 135
A2.6 /ə/ with No Tone Phoneme and the Mid-High Tone ......................................... 136
LIST OF TABLES

Table 1: The Consonantal Inventory of S’gaw Karen ........................................29
Table 2: The Vowels of S’gaw Karen ..............................................................30
Table 3: S’gaw Karen Consonant Clusters ......................................................33
Table 4: Written Representation of S’gaw Karen Consonants ......................133
Table 5: Written Representation of S’gaw Karen Consonant Clusters ..........134
Table 6: Written Representation of S’gaw Karen Vowels .........................135
Table 7: Written Representation of S’gaw Karen Tones ..............................136
LIST OF FIGURES

Figure 1: The Sino-Tibetan Family ..............................................................................4

Figure 2: The Karenic Language Subfamily ..................................................................5

Figure 3: The Karen Homeland ......................................................................................9

Figure 4: The Eastern States of Burma ..........................................................................10

Figure 5: Distribution of Numeral Classifiers ...............................................................37

Figure 6: The Structure of the S’gaw Karen Syllable for /kʰwè/ ‘saw’ .........................131
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLF</td>
<td>numeral classifier</td>
</tr>
<tr>
<td>POSS</td>
<td>possessive</td>
</tr>
<tr>
<td>1/2/3</td>
<td>first/second/third person</td>
</tr>
<tr>
<td>PL</td>
<td>plural</td>
</tr>
<tr>
<td>SG</td>
<td>singular</td>
</tr>
<tr>
<td>OBJ</td>
<td>object</td>
</tr>
<tr>
<td>DEM</td>
<td>demonstrative</td>
</tr>
<tr>
<td>IRR</td>
<td>irrealis</td>
</tr>
<tr>
<td>VOC</td>
<td>vocative</td>
</tr>
<tr>
<td>NEG</td>
<td>negative</td>
</tr>
<tr>
<td>INT</td>
<td>intensifier</td>
</tr>
<tr>
<td>Q</td>
<td>interrogative</td>
</tr>
<tr>
<td>FUT</td>
<td>future</td>
</tr>
<tr>
<td>EXCL</td>
<td>exclamative</td>
</tr>
<tr>
<td>COP</td>
<td>copula</td>
</tr>
<tr>
<td>NOM</td>
<td>nominative</td>
</tr>
<tr>
<td>TOP</td>
<td>topic pronoun form (cf. Jones 1961:18)</td>
</tr>
<tr>
<td>IMP</td>
<td>imperative</td>
</tr>
<tr>
<td>REFL</td>
<td>reflexive</td>
</tr>
</tbody>
</table>
CHAPTER ONE: PURPOSE & INTRODUCTION

1.1 PURPOSE

This research project has two objectives. First, I was interested in conducting field research. I was fortunate enough that a large group of S’gaw Karen refugees from Eastern Burma relocated close to my home, and I made connections to this community as an English teacher in a local high school. Working with this community is a great opportunity, for it offers the chance to work with a group of speakers who would otherwise be geographically inaccessible to me.

My interest in field research and language documentation arose from my attempts as an English as a Second Language instructor to learn more about the S’gaw Karen language for my own instruction. The available research on this language in terms of its structure, history, and its speakers is scant and outdated. Information regarding the S’gaw language is limited because most research relies solely on Robert Jones’s (1961) study of S’gaw Karen. While Jones’s work is the most comprehensive study available, the manner in which it is written is not easily accessible, and the analysis of the data is incomplete. A new, comprehensive study of the S’gaw Karen
language is long overdue for this language of over a million speakers (Lewis et al. 2013), but that project is beyond the scope of a master’s thesis.

A more limited and focused project seemed appropriate. The development of a modest corpus of a representative sample of the language would serve me as an apprenticeship in conducting field research, and the corpus could then be used to examine a particular feature of the S’gaw Karen language. Thus, the second purpose of this project was to use the corpus that I developed to describe a grammatical feature of S’gaw Karen. I chose to examine the numeral classifier system in S’gaw Karen.

While some research is available concerning numeral classifiers in S’gaw Karen, I was interested in describing the system of numeral classifiers more fully. Jones’s (1961) treatment of numeral classifiers is an excellent start, but it remains limited because Jones failed to differentiate numeral classifiers and measure words. Ratanakul (2001b) examined the etymology of different classifiers in the Mae Chaem dialect of S’gaw Karen and provided some analysis of the structure of the noun phrase with numeral classifiers. However, Ratanakul’s (2001b) examination of the syntactic role of the numeral classifier is limited to only a few constructions and is more representative
of the dialect of S’gaw Karen spoken around Chiang Mai rather than the S’gaw Karen spoken in Burma (Rattanaporn 2012: 3).

Building upon Ratanakul’s (2001b) excellent compilation of numeral classifiers, this thesis examines the structure of the noun phrase with numeral classifiers in a variety of constructions, including those with adjectives and demonstratives. My descriptions of these various constructions rely on the data in my corpus which consists of a series of S’gaw Karen children’s stories as well as elicitations of S’gaw Karen consultants.

1.2 AN INTRODUCTION TO THE KAREN PEOPLES & LANGUAGES

1.2.1 WHO IS KAREN?

The term Karen refers to both an ethnic group consisting of a variety of subdivisions and also to a language subfamily within the larger Sino-Tibetan family (see Figure 1).
Frequently, the adjectival term “Karenic” is used to refer to the language group while the term “Karen” is used to refer to the ethnic group. While the subdivisions of Karen ethnicity and Karenic languages often overlap in terminology (i.e. often an individual who is ethnically S’gaw Karen, for example, speaks the S’gaw Karen language), this is not always the case. Figure 2 displays the divisions of the Karenic language subfamily, but these divisions are not necessarily representative of the ethnic relationships among Karen peoples.
Individuals may identify themselves as members of one of the Karen ethnic subgroups without having the ability to speak the subgroup language. Furthermore, marriage between different Karen ethnic groups as well as between other non-Karen ethnic groups is common, but individuals may choose to identify themselves in a specific way. For example, a consultant recounted an individual who identified as Pwo Karen ethnically. Her mother was ethnically Pwo Karen, and her father was ethnically S’gaw Karen. The individual was fluent in S’gaw Karen but knew only a little Pwo Karen. Nevertheless, she identified herself as ethnically Pwo Karen like her mother. Similarly,
individuals may self-identify as being ethnic Karen without being able to speak a Karenic language. As Charney (2009: 8) writes,

In many spheres of Burmese life, identities and identifications that [are] fluid, syncretic, multiple, or even undefined [are] common. Whether in terms of religion, ethnicity, or culture, it [is] not unusual for an individual or a group to change [his or her] self-identifications in different contexts.

The label of being Karen may be the result of self-identification by individuals or it may be a term applied from outside groups used to differentiate the Karen from themselves (i.e. not Burmese). As a result, the relationships between the languages within the Karenic subfamily are far easier to map than are the ethnic identities of the Karen people. Many scholars divide the Karen subfamily into three divisions, Northern, Central, and Southern, mirroring the geographical locations of the majority of their speakers (see Figure 2 above). The languages within these subdivisions are not mutually intelligible and as such are distinct languages and not dialects. While individuals may identity themselves as belonging ethnically to one of these subdivisions, there is also a larger sense of Karen ethnic identity. The Karen ethnic group is one of the largest ethnic minorities in Burma (the majority group being the Burman people), and self-
identification within this group is often in opposition to the dominant Burmese
government. As a result, it is important not to assume that the Karen ethnic group is
necessarily made up of the same individuals that speak Karenic languages.

1.2.2 POPULATION OF ETHNIC KAREN AND KARENIC SPEAKERS

Accurate statistics regarding the precise number of both ethnic Karen and
speakers of Karenic languages in Burma and Thailand are unavailable for two reasons.
First, the last Burmese census was taken in 1983 (Lewis et al. 2013). Since that time,
numerous Karen groups have been involved in a civil war in Burma and there has been
a mass refugee migration to Thailand. Secondly, as discussed in section 1.2.1, being
ethnically Karen and speaking a Karenic language are not mutually inclusive. As a
result, estimates vary widely regarding the population of Karen people and Karenic
speakers. Such data are difficult to collect in an era of constant warfare and shifting
populations.

The often interchangeable uses of Karen to refer to an ethnic group and to a
language subfamily make estimating the population of ethnic Karen people and Karenic
language speakers even more difficult. It is not always clear if scholars are referring to
populations of ethnic Karen alone, to speakers of Karenic languages only, or to a
combination of the two groups. Ethnologue, a database of known languages in the world, estimates around 3.7 million speakers of Karenic languages but relies heavily on the 1983 census (Lewis et al. 2013). David Bradley (1997: 46) argues for 3.9 million, and in the most recent study of Karenic, Ken Manson (2011: 1) suggests between 6 and 10 million ethnic Karen. However, Manson notes that not all ethnic Karen speak a Karenic language, especially those living toward Central Burma. David Solnit (1997: xiii) argues that speakers of Karenic languages are the largest ethnic minority in both Burma and Thailand. Solnit estimates the number of Karenic speakers to be between 3 to 4.5 million, but he does not differentiate between ethnic Karen and Karenic speakers.

Estimates on the number of Karenic languages differ (see Figure 2 above). Lewis and colleagues (2013) provide information on twenty-one Karenic languages (one being extinct) while Manson estimates the number of Karenic languages to be between twenty and thirty (2011: 1).

The Karenic languages with the most speakers are S’gaw Karen, Pwo Karen, Pa’o Karen, and Kayah (also widely known as Karenni). The most widely spoken of these is S’gaw Karen, spoken by an estimated 1.5 million speakers worldwide (Lewis et al. 2013).
1.2.3 GEOGRAPHIC LOCATION OF THE KAREN PEOPLE

The Karenic languages are spoken at the extremity of the Sino-Tibetan and Tibeto-Burman linguistic area (see Figure 3).

![Figure 3: The Karen Homeland (Keyes 1979)](image)

Within this area, the Karenic languages form the Eastern most extreme of Sino-Tibetan languages spoken in South Asia, bordered by the Tai-Kadai languages spoken in Thailand (Solnit 1997: xiii). The Karen languages are primarily spoken alongside the
Eastern border of Burma with Thailand in states which were initially established as semi-autonomous states when the British dissolved their Burmese colony in 1948 (Charney 2009: 66-67, Silverstein 1997: 15) (see Figure 4). These states include Kayah state, Kayin state, Tanintharyi state, and portions of the Mon and Shan states. Additionally, Karen languages are spoken natively, though most likely bilingually, farther west in the Irrawaddy Delta region as well as farther east across the Thai border (Lewis et al. 2013; Manson 2001: 1; Solnit 1997: xiii).

Figure 4: The Eastern States of Burma (Dedering 2010)
1.2.4 MODERN HISTORY OF THE KAREN PEOPLE

The role of the Karen people in Burma changed dramatically with British colonialism. Because the Karen are one of the largest ethnic minorities in Burma and because Christianity was spreading through the regions inhabited by the Karen (see §1.3.3.1), the British relied heavily on Karen soldiers during their occupation of Burma (Charney 2009: 53). Since Burma was ruled by an ethnic Burman monarch, the British were able to rely on the already present ethnic tensions between the Burmans and the Karen to employ the Karen as a native army to control the rebellious Burman majority (Charney 2009: 55). However, when the British withdrew to India in the face of an impending Japanese invasion from Thailand during World War II, one of the first acts of the liberated Burmans was to disarm and attack the Karen soldiers (Charney 2009: 55).

After the Japanese were later forced out of Burma by the British in 1945, the British began to dismantle their colonial system. Efforts were made by the British to establish semi-autonomous states for the various major ethnic groups in Burma. For example, the Kayin or Karen state (see Figure 4 above) was created to serve as the state of the ethnic Karen people (Charney 2009: 66-67). As such, a large number of Karenic
speakers live in that state. However, the delegations representing the various ethnic
groups receiving states, such as the Karen, the Chin, the Karenni, etc., were dissatisfied
that these states were not given more authority to secede from Burma and refused to
sign the Panglong Agreement of 1947 that established the Kayin state (Charney 2009:
74). War between ethnic Karen military organizations (such as the Karen National
Union and Democratic Karen Buddhist Union) and the Burmese controlled government
broke out soon after in 1949. This war continues to this day. Because of the
overwhelming numbers of ethnic Burmans, the economic control of the junta, and
infighting between the different Karen groups (which resulted in the Democratic Karen
Buddhist Union allying itself with the Burmese government), the relatively tiny military
wing of Karen National Union has been reduced to guerilla warfare and retreat for
several decades.

1.2.5 CURRENT STATUS OF THE KAREN PEOPLE & THE ENDANGERMENT
OF THE KARENIC LANGUAGES

The Burmese military have adopted a scorched-earth policy of warfare against the
ethnic Karen forces. This tactic includes the destruction of Karen villages and cities, the
kidnapping and forced conscription of children into the Burmese military, and laying
mines in Karen areas so that retreating forces and displaced families cannot return. As a
result, hundreds of thousands of Karen people have been displaced as refugees by the war (IDMC 2013). Many thousands of these refugees have fled to camps in Thailand. In the largest of these refugee camps, Mae La, the estimated population of approximately 43,000 as of January 2014 (TBC 2014) is predominately ethnic Karen (IDMC 2013). As these refugees are granted third country resettlement in countries such as the United States, Sweden, Norway, Canada, and Australia, the geographical location of ethnic Karen and Karenic language speakers expands greatly. As a result, these languages are more likely to experience rapid and diverse change. The number of native speakers is also likely to decline as children born in resettlement countries learn the language taught in that country's schools rather than the language of their parents. For example, L1 literacy instruction is limited in resettlement countries and refugee camps, with Lewis and colleagues (2013) citing L1 literacy rates of between 10% and 30% among Karen people living in Thailand; however, it is not clear if these statistics refer only to ethnic Karen born in Thailand or include Burmese-born Karen living in Thai refugee camps. If the civil war in Burma continues, education and literacy remain absent in rural areas, and migration to countries speaking predominately Indo-European languages increases, it is likely that much of the diversity of the Karenic languages will disappear forever.
1.3 INTRODUCTION TO THE S’GAW KAREN LANGUAGE

1.3.1 DIALECT USED IN THIS STUDY

Much of the more recent work on Karenic languages has focused on the Karenic languages spoken in Thailand, especially around Chiang Mai. As my consultants all come from within Burma and speak S’gaw Karen, I am focusing primarily on the S’gaw Karen dialect spoken in Burma. The dialect of S’gaw Karen spoken in Burma is different with regards to both vocabulary and phonology than the dialect spoken in Thailand (Rattanaporn 2012: 3). Furthermore, Jones (1961) is the only study on Burmese S’gaw Karen, and he relied on a single speaker in his research. One can assume that there are a variety of dialects of S’gaw Karen spoken in Burma, but until further research can be conducted, I am forced to make generalizations about how representative the dialect spoken by my consultants is. More information on the backgrounds of my consultants can be found in §2.1.

1.3.2 S’GAW KAREN’S LANGUAGE FAMILY

The Karenic languages are a member of the Sino-Tibetan language family (see Figure 1 above) and the Tibeto-Burman subfamily more narrowly (cf. Solnit 1997: xiii; Graham & LaPolla 2003: 18). Much of the early research on the Karenic languages
centered on whether or not the Karenic languages were part of the Tibeto-Burman family at all. Many of the common typological features of the Tibeto-Burman subfamily are absent in the Karenic languages, leaving some early scholars to label the Karenic languages as a separate, anomalous group, related perhaps to the Tai-Kadai languages (Luce 1959: 3). Most notably, the Karenic languages are typologically subject-verb-object ordered (SVO) while the other Tibeto-Burman languages are subject-object-verb (SOV) (Dryer 2003: 43). However, the Karenic languages are now accepted as being part of the Tibeto-Burman language family (cf. Thurgood & LaPolla 2003: 19; Solnit 1997: xiii), and the syntactic differences between the Karenic languages and the other Tibeto-Burman languages are attributed to the influence of bordering languages within the Mon-Khmer and Tai-Kadai families, which are both SVO typologically (Thurgood & LaPolla 2003: 19). Within the Karenic subfamily, S’gaw Karen is part of the Southern branch (see Figure 2). S’gaw Karen is closely related to Pwo Karen, but the two languages are not mutually intelligible.¹

¹ Native speaking Pwo and S’gaw informants indicated that they could not understand each other’s language. However, both languages are part of the southern branch of the Karenic subgroup, and my own examination of Pwo Karen indicates that there is a great deal of grammatical and lexical similarity with S’gaw Karen.
1.3.3 LITERATURE REVIEW OF THE KARENIC LANGUAGES

1.3.3.1 EARLY MISSIONARY INFLUENCE & SCHOLARSHIP

Europeans first made contact with Karen people in 1828, shortly after the British began annexing Burmese land in 1824 (Marshall 1922: 296; Charney 2009: 5).

According to legend, this occurred when Adoniram Judson, the founder of the American Baptist Mission in Burma, purchased a Karen man named Ko Tha Byu who had been sold into slavery as punishment for crimes he committed. As The Reverend Marshall (1992: 296) writes,

> If one were planning to start a movement to transform the life and religion of a race, one would not be expected to choose a savage bandit—a cutthroat who had taken part in the murder of at least thirty persons—to promote his enterprise.

The Reverend Judson was successful in converting Ko Tha Byu, who then helped spread Christianity among the Karen people. As Baptist Missionaries continued to convert the Karen to Christianity, their contact led to the first anthropological studies being written about the Karen, including *On the Karens and Their Language* by Edmund Cross (1854) and Reverend Harry Marshall’s (1922) *The Karen People of Burma: A Study in Anthropology and Ethnology*. In fact, the writing systems used for
the S’gaw Karen and Pwo Karen languages (both Southern Karenic languages) were both adapted from Burmese scripts by missionaries: S’gaw Karen by Reverend Jonathan Wade in 1832 and Pwo Karen by Dr. Francis Mason sometime before 1846 (Jones 1961: v; Mason 1846). The role that missionaries and churches play in Karenic linguistic studies and Karen culture and diaspora remains involved today, with many scholars continuing to rely on missionaries in Burma and Thailand to gather data (cf. Luce 1959: 5).

The first attempts at describing the Karenic languages were also taken on by missionaries. Wade’s (1861) Karen Vernacular Grammar was the result of several decades of research and data collection. Wade’s (1861) text focuses primarily on the S’gaw Karen language, and the majority of the text is written in S’gaw Karen rather than in English. Much of his data and glossings were later used by Mason (1846: v) in A Synopsis of a Grammar of the Karen Language, Embracing Both Dialects, Sgau and Pgho, or Sho. Both works were intended to teach other missionaries how to speak S’gaw and Pwo Karen so that they could interact with Karenic speakers outside of the cities. In some instances, the scholars’ familiarity with Indo-European languages influenced their analyses and descriptions of the Karenic languages. Additionally, at this
time, it was not understood that S’gaw and Pwo were distinct languages or ethnic
groups (Charney 2009: 8). Mason (1846: iii) acknowledged that his work was an
incomplete start, writing that his work is of “temporal utility only, and is expected to be
superseded in a few years.” Reverend David Gilmore (1898: 3) attempted this update
with *A Grammar of the Sgaw Karen*. In the introduction, he writes that his work is
intended as an update to Mason’s earlier work, which was difficult to find and “very
perplexing to a neophyte.” Additionally, Gilmore (1898) writes that the American
Baptist Karen Mission in Burma requested that he complete the grammar because there
was no way for new missionaries to Burma to learn S’gaw Karen.

**1.3.3.2 RECENT RESEARCH ON S’GAW KAREN**

There was very little published concerning the Karenic languages until the mid-
20th century. The increase in attention placed on the Karenic languages, in particular
S’gaw Karen, was focused primarily on what the Karenic languages could contribute to
the study of the characteristics and typology of the Tibeto-Burman subfamily. Often,
however, the atypical characteristics of the Karenic languages relative to other Tibeto-
Burman languages led many scholars to conclude that the Karenic languages were part
of a separate family (Luce 1959: 3), a remote branch of the Sino-Tibetan family but not
the Tibeto-Burman subfamily (Benedict 1972: 6, 128; Lehman 1979: 220; Thurgood 1985: 384), or that Karen was an isolate (Jones 1975 qtd. in Solnit 2003: 624). As discussed previously, Karenic is now an accepted branch of Tibeto-Burman.

Much of the discussion on Karen during this rise in the interest in Tibeto-Burman typology relied primarily on Robert B. Jones’s (1961) *Karen Linguistic Studies: Description, Comparison, and Texts*, adapted from research completed while he wrote his dissertation (Jones 1961: 3) and updated after a year of fieldwork in Burma (Obituary 2007: 1). Indeed, much contemporary research cites Jones’s (1961) book as the sole resource on S’gaw Karen. Jones’s (2007: 1) obituary states that “this monograph is among the most thoroughgoing studies of any Tibeto-Burman language and is still the most authoritative single-volume study of the Karen family.” In particular, Jones (1961) examines numerous Karenic languages and dialects in order to reconstruct proto-Karen, and these reconstructions were foundational for a number of other scholars’ work (cf. Robbins 1969, Benedict 1971, Manson 2001). Within TB studies, these proto-Karen reconstructions have been helpful in reconciling the unique Karenic languages within TB typology (cf. Manson 2001).
It should also be noted that all of the Karen studies discussed so far, including Wade (1861), Mason (1846), Gilmore (1898), and Jones (1961), have focused on dialects spoken around Moulmein. This is likely because Moulmein was the location of the British presence during colonial Burma as its first capital between 1826 and 1852, and Moulmein continued to maintain a large Anglo-Burmese population after the capital moved to Rangoon (Charney 2009: 13, Silverstein 1977: 21). The dialects represented in these texts may not be representative of other S’gaw Karen dialects spoken outside of the city. Mason (1846: v) even mentions that he could not yet publish a complete grammar “until he is better acquainted with the language as spoken out of the Provinces.”

While there has not been a general study of S’gaw Karen updating Jones’s (1961) work, there has been some scholarship focusing on individual aspects of the language. Hilary Chappell (1992) examined benefactive and causative constructions in S’gaw Karen, while Kato (1993) and Carol Lord and Louisa Benson Craig (2004) studied verb serialization and concatenation respectively. Suriya Ratanakul has been the most prolific scholar of S’gaw Karen, publishing on syntax (1980), verbal constructions (1981), adverbs (2001a), and numeral classifiers (2001b). Unfortunately, some of her
work is extremely difficult to find or is available only in Thai, such as her S’gaw Karen dictionary (1986).

1.3.3.3 RESEARCH ON OTHER KARENIC LANGUAGES

Since Christian missionaries completed the early scholarship on Karenic languages and focused on the Christianized S’gaw, there was little appreciation that the different Karen ethnic groups were quite varied and spoke different languages until recently. For example, Charney (2009: 8) writes,

Despite the substantial cultural, linguistic, and religious differences among Karen groups, a Karen identification based on the practices of only one group, that of the Christianized Sgaw Karen, was applied to the Karen in general, in part because this community had more records than others and because they had been the main subject of missionary records.

While there have been numerous scholarly contributions by Thai authors, there is little scholarship on other Karenic languages beside S’gaw available in English. There was some interest in Pwo Karen in the early twentieth century. A Manual of the Pwo-Karen Dialect by C.H. Duffin (1913) and a Pwo/English bilingual dictionary by W.C.B Purser and Tun Aung (1922) were both published by the American Baptist Mission Press.
There has been no updated, comprehensive study of Pwo since then, but there has been a number of narrower, more focused contributions to the study of Pwo (cf. Kato 1999; Dawkins & Phillips 2009). Research on Bwe is likewise limited, with Henderson’s (1990) *Bwe Karen Dictionary* the most significant. The Kayah languages are the best studied with regards to modern scholarship. David Solnit’s (1997) work on Eastern Kayah-Li is an excellent example of the updated, comprehensive, and modern analysis needed for S’gaw Karen.
CHAPTER TWO: METHODOLOGY & PHONOLOGY

2.1 CORPUS METHODOLOGY

The corpus for this project consists of four traditional children’s stories translated from S’gaw Karen into English. The stories were published by Drum Publications and are available for download at http://www.drumpublications.org. Three consultants helped in the translations. Two consultants provided a word-by-word translation. One was eighteen years old at the time of translation and had lived in the United States for four years. She was originally born in Ta Law Thaw refugee camp in Thailand after her mother fled Burma while pregnant. Her family moved to Mae La camp shortly after, and she lived there her whole life before moving to the United States. The second consultant was also eighteen years old at the time of translating and had lived in the United States for five years. He was also born in Ta Law Thaw in Thailand along the border with Burma, which is now an abandoned refugee camp. After a year there, he moved to a small village in Thailand named Soe Klo, which served as a temporary refugee camp for many families, and then moved to Mae La refugee camp four years later. He lived there permanently until moving to the United States. Both consultants were proficient English speakers. When necessary during translating, both

After the two consultants had completed the word-by-word translations, I compared and corrected the translations with both consultants together along with an additional consultant who was nineteen years old at the time. She was born in K’mah Klut in Burma and lived for ten years in No’Po refugee camp in Thailand before moving to the United States. She had lived in the United States for four years at the time of translation. The three consultants and I examined the two translations provided by the first two consultants together in order to decide upon the most accurate translation. We discussed the proper translation for lexical items since the two consultants sometimes gave similar, but slightly different translations. The consultants discussed the word in question and decided on the most accurate translation. We also examined grammatical elements that were impossible to translate. For example, one consultant consistently translated numeral classifiers as either “+s” or “plural.”

After the translations had been completed and compiled, I analyzed each sentence and added interlinear morpheme glossing (see Appendix One). These annotated sentences were used as the main resource to analyze the structure of noun
phrases with numeral classifiers. Information about constructions not present in the translations was elicited from the three consultants who helped build the corpus and one additional consultant. She was thirty-nine years old at the time of the elicitations and was from Shwe Gon in Burma, which is located in Kayin State. She lived for two years in Mae La refugee camp in Thailand and six years in the United States. She currently works as an English translator and has a native knowledge of S’gaw Karen and Burmese and is fluent in English.

2.2 PHONOLOGICAL DESCRIPTION & TRANSCRIPTION CONVENTIONS

The following description of the phonological system for S’gaw Karen is based on work with the native consultant from K’mah Klut. Major contributions to S’gaw Karen phonology include Jones (1961), Namkung (1996), which is largely a summary of Jones’s work, and Baa (2001). My research conducted with the consultant has shown Baa’s treatment of S’gaw Karen to be more reflective of my consultant’s speech than Jones’s on a number of significant features of S’gaw Karen phonology; however, this description differs from Baa’s in ways that will be noted below.
2.2.1 S’GAW KAREN CONSONANTS

The following analysis is the result of working with a native consultant to establish minimal pairs for the consonants of S’gaw Karen. This analysis resulted in an inventory of nine stops (p, pʰ, b, t, tʰ, d, k, kʰ, and ?), four nasals (m, n, ŋ, and ŋ), eight fricatives (θ, s, sʰ, š, x, Y, h, and ɦ), two affricates (tʃ and dʒ), three approximants (w, j, j), and one lateral (l) (see Table 1).

Minimal pairs for the consonants are provided below:\1:

/t/ – /θ/  /tí/  ‘uncle’  /θí/  ‘die’
/dʒ/ – /tʃ/  /dʒáʔ/  ‘ask’  /tʃáʔ/  ‘sew’
/b/ – /d/  /bá/  ‘worship’  /dá/  ‘youngest child’
/n/ – /ŋ/  /ná/  ‘witch’  /ŋá/  ‘internal’
/ŋ/ – /m/  /ŋá/  ‘borrow’  /má/  ‘wife’
/h/ – /ɦ/  /há/  ‘to cook by steaming’/háʔ/  ‘attention particle’
/x/ – /ɣ/  /xē/  ‘to run’  /ɣē/  ‘cane’

\1 In the literature on Karen languages, it would seem the convention is to transcribe the aspirated phonemes as a full /h/ rather than the superscript /ʰ/ (cf. Solnit 1997, Jones 1961, Namkung 1996, Ratanakul 2001b). I have chosen to follow Baa’s (2001) more accurate transcription using superscript aspirations so that it is clear that these are aspirations rather than coarticulations with a voiceless glottal fricative.
This analysis of the consonantal inventory of S’gaw Karen differs from that proposed by Jones (1961) regarding affricates and glottal fricatives. Jones (1961) argued for the existence of the palatal stop /c/ and aspirated palatal stop /cʰ/. However, the native speaker used as a consultant in this study produced the postalveolar affricates /tʃ/ and /dʒ/. This analysis is in line with Baa’s (2001) inventory which also lists /tʃ/ and /dʒ/ rather than /c/ and /cʰ/. Additionally, Jones (1961) did not distinguish between the voiced and voiceless glottal fricatives as is done here. Interestingly, though, he reported the voiceless glottal fricative in Bassein S’gaw and the voiced glottal fricative in Moulmein S’gaw (Namkung 1996: 170-73). A possible reason for this oversight is that /ɦ/ occurs infrequently, primarily to describe animal sounds or in grammatical particles.
Jones (1961) also includes /z/ in his description of S’gaw Karen whereas I do not. Other scholars’ transcriptions also include /z/ (cf. Ratanakul 2001b, Rattanaporn 2012). My consultants did not produce /z/, but they indicated that variations between /z/ and /s/ or /sʰ/ are acceptable. I suspect that some S’gaw dialects employ /z/ in the same instances where other dialects use /s/ or /sʰ/, but a more detailed examination of this issue is outside the scope of the project.

The proposed inventory differs from Baa’s (2001) with regard to the dental fricative and postalveolar fricative. Baa (2001) argued that Jones (1961) confused the dental fricative /θ/ for a fortis dental stop /t/. However, the consultant used in this project pronounced /θ/ rather than /t/. Baa (2001) also argues that the postalveolar fricative /ʃ/ is fronted in S’gaw and argues for /s/ instead. On this point Baa (2001) might be correct, but I have chosen to follow Jones (1961) because the frequency of this phoneme is very small and does not occur once in my corpus or in Baa’s (2001) corpus. It is possible that the sound is a borrowing for foreign words, such as Jesus /jèʃú?/, which would account for its infrequency.
Table 1: The Consonantal Inventory of S’gaw Karen

<table>
<thead>
<tr>
<th></th>
<th>Bilabial</th>
<th>Dental</th>
<th>Alveolar</th>
<th>Post – Alveolar</th>
<th>Palatal</th>
<th>Velar</th>
<th>Glottal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stop</td>
<td>p</td>
<td>b</td>
<td>t</td>
<td>d</td>
<td></td>
<td>k</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>pʰ</td>
<td></td>
<td>tʰ</td>
<td></td>
<td></td>
<td>kʰ</td>
<td></td>
</tr>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>η</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>n̥</td>
</tr>
<tr>
<td>Fricative</td>
<td>θ</td>
<td>s</td>
<td>š</td>
<td></td>
<td></td>
<td>x</td>
<td>y</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>sʰ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tʃ</td>
</tr>
<tr>
<td>Approximant</td>
<td>w</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>j</td>
</tr>
<tr>
<td>Lateral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>l</td>
</tr>
</tbody>
</table>

2.2.2 S’GAW KAREN VOWELS

An examination of the vowel system of S’gaw Karen with a native consultant has indicated that the language has four front vowels (i, e, e, and a), one central vowel (ə), and four back vowels (u, u, o, and o). Only the high back vowels show a tense/lax distinction (see Table 2). There are no diphthongs in S’gaw Karen.

Below are the minimal pairs of the vowels of S’gaw Karen:

/i/ – /e/ /iː/ ‘narrow’ /ɪː/ ‘to bite’
/a/ – /e/ /má/ ‘wife’ /mé/ ‘teeth’
/o/ – /o/ /oː/ ‘to have’ /ɔ/ ‘to eat’
The proposed vowel inventory differs from Jones’s (1961) only in that he argues that S’gaw Karen has a high central unrounded vowel (which he transcribed as /y/ though /i/ may be more accurate using current IPA standards) rather than /ɯ/. However, the consultant in this study produced /ɯ/ and not /i/. Baa (2001) likewise includes /ɯ/ in his inventory and not /i/, but /i/ is used in transcriptions by other scholars (cf. Ratanakul 2001b, Thomas & Brunelle 2010, and Chappell 2009).

Table 2: The Vowels of S’gaw Karen

<table>
<thead>
<tr>
<th></th>
<th>Front</th>
<th>Central</th>
<th>Back</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>i</td>
<td></td>
<td>ɯ u</td>
</tr>
<tr>
<td>Close-Mid</td>
<td>e</td>
<td>ē</td>
<td>o</td>
</tr>
<tr>
<td>Open-Mid</td>
<td>ɛ</td>
<td></td>
<td>ɔ</td>
</tr>
<tr>
<td>Open</td>
<td>a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.2.3 S’GAW KAREN TONES

Tones are phonemic in S’gaw Karen and are represented orthographically. However, the number of tones is controversial (cf. Chappell 2009, 15; Finkeldey 2011, Rattanaporn 2012, 93-94). In this study, six distinct tone phonemes were established with the native consultant. These include one mid-high tone (´), a mid-high tone checked by a glottal stop (ʔ´), a mid tone (˘), a mid-low tone (´), a mid-low tone checked by a glottal stop (ʔ˘), and a low tone (˘).2

Minimal pairs for the six tone phonemes are:

/´ʔ/ – /˘/ /màʔ ‘son-in-law’ /mà ‘to do’

/˘/ – /´ʔ/ /mû ‘hate’ /mûʔ ‘ceremony’

/´/ – /˘/ /kà ‘open’ /kà ‘grill’

Jones (1961: 9) argues for three tones, each with an allophone. These allophones are checked by a voiceless glottal stop. Thus, in his analysis there is a high, mid, and low tone, and each can be shortened by the glottal stop. Namkung (1996) follows Jones’s (1961) analysis, and Solnit (1997: xvii) uses Jones’s transcription of tones, but changes Jones’s high tone with a glottal stop to a high falling final-creaky tone. Baa’s

2 Rather than tone bars, accents will be used to transcribe tones according to the IPA standard for the convenience of the reader.
(2001) analysis differs greatly from Jones’s (1961). Baa (2001) argues instead for six distinct tones, which include four level tones and two contour tones that are checked by a glottal stop. The level tones are mid-high, mid, mid-low, and low. One contour tone is high falling, and the other contour tone is low falling. A detailed analysis of the differences between Jones’s (1961) and Baa’s (2001) examination of S’gaw tones can be found on pages 30 to 32 of Baa’s work. While the consultant for this project indicated six distinct phonemic tones, I have also relied on Baa’s (2001) analysis of these tones and followed his transcription methodology closely.

2.2.4 S’GAW KAREN SYLLABLE STRUCTURE

The structure of the S’gaw Karen syllable is an initial consonant or consonant cluster, a vowel, and a tone3.

\[ C_1(C_2)V_1T \]

Toneless syllables only occur with the vowel /a/; however, /a/ can also be used with a tone phoneme. A medial consonant of /l/, /r/, /j/, /Ɣ/ or /w/ is optional for some initial consonants (see Table 3). Consonants and consonant clusters only appear in syllable initial positions.

---

3 This analysis of S’gaw Karen syllable structure differs from Baa’s (2001) in ways that are not necessary to detail here. Interested readers are referred to pages 27-28 of his work.
Table 3: S’gaw Karen Consonant Clusters

<table>
<thead>
<tr>
<th></th>
<th>/w/</th>
<th>pʰw</th>
<th>tw</th>
<th>tʰw</th>
<th>dw</th>
<th>nw</th>
<th>kw</th>
<th>kʰw</th>
<th>sw</th>
<th>sʰw</th>
<th>θw</th>
<th>jw</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>nw</td>
<td>lw</td>
<td>bw</td>
<td>pʰw</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/l/</td>
<td></td>
<td>pl</td>
<td>pʰl</td>
<td>bl</td>
<td>ml</td>
<td>kl</td>
<td>kʰl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Ɂ/</td>
<td></td>
<td>pɁ</td>
<td>pʰɁ</td>
<td>bɁ</td>
<td>ɁɁ</td>
<td>ɁɁ</td>
<td>kɁ</td>
<td>kʰɁ</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/j/</td>
<td></td>
<td>pʰj</td>
<td>mj</td>
<td>kj</td>
<td>kʰj</td>
<td>pj</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>/Y/</td>
<td></td>
<td>pY</td>
<td>pʰY</td>
<td>bY</td>
<td>mY</td>
<td>sY</td>
<td>sʰY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This analysis differs slightly from Jones (1961) and Namkung (1997) who argue that /ʔ/ is the only consonant that appears in a syllable final position because of the two tones (three according to Jones 1961) that are checked by a glottal stop. However, considering that the only instances of /ʔ/ appearing as syllable final occur with these two checked tones, I agree with Baa’s (2001) analysis that the glottal stop is part of the tone phoneme rather than a separate, syllable final consonant.
CHAPTER THREE: S’GAW KAREN NUMERAL CLASSIFIERS

3.1 LITERATURE REVIEW OF NUMERAL CLASSIFIERS

3.1.1 OVERVIEW OF NUMERAL CLASSIFIERS

The primary function of numeral classifiers is enumeration. Languages that have numeral classifiers do not have the same distinction between count and mass nouns that English does. In English, a count noun by itself is singular (e.g. *dog), is generally pluralized by affixing -s (e.g. dogs), is directly enumerated (e.g. two dogs), and can take an indefinite article (e.g. a dog). In contrast, mass nouns (e.g. rice) cannot be pluralized (e.g. *rices), require a measure word for enumeration (e.g. two cups of rice), and cannot take an indefinite article but require a quantifier (e.g. some rice, *a rice). In some languages, nouns behave more similarly to English mass nouns and require a numeral classifier for enumeration of the noun in a way similar to how English mass nouns require a measure word. Lucy (2004: 8) writes that in numeral classifier languages, nouns are “semantically unspecified as to a quantificational unit” and so...
numeral classifiers provide a unit so that the noun may be counted. Examples 1 – 4 below illustrate numeral classifier usage in a variety of languages.¹

(1) Japanese (Matsumoto 1993: 673)
enpitsu ni-hon
pencil two-CLF
‘two pencils’

(2) Burmese (Vittrant 2002: 132)
/khwe niʔ KəN/
dog two CLF:ANIMAL
‘two dogs’

(3) Javanese (Jones 1970: 3)
kertas sepuluh lembay
paper ten CLF:SHEETS
‘ten papers’

(4) Yucuna (Aikhenvald 2003: 106)
pajluhu-na yahui
one-CLF:ANIMAL dog
‘one dog’

Numeral classifiers are found in many of the languages of Southeast Asia, and much scholarship has been devoted to numeral classifiers in the languages of East and Southeast Asia, including Burmese (Vittrant 2002), Thai (Carpenter 1986), Mandarin (Zhang 2007, Zhang 2012, Fang & Connelly 2008), Korean (Sohn 2001), Hmong

¹ All examples drawn from the literature are shown with the same transcription as provided in the source text.
(Bisang 1993), Japanese (Downing 1996), a number of Austroasiatic languages (Adams 1989), and Vietnamese (Löbel 2000), among many others. Outside of languages of Southeast Asia, numeral classifier systems are found in the languages of the Pacific Northwest: Tlinglit (Sherzer 1976); of Central America: Tzeltal (Berlin 1968), Yucatek (Suárez 1983), Teribe (Quesada 2000); of South America: the Arawak languages (Aikhenvald 1996); of Eastern Africa: Boko (Jones 1998), Kana (Ikoro 1994); and of Eastern Europe and Western Asia: Hungarian (Beckwith 1992), Persian (Mahootian 1997), Turkish (Lewis 1967).\(^2\) See Figure 5 for a geographic distribution of numeral classifiers. Some languages have a small set of numeral classifiers, while many languages have large systems. For example, the Papuan languages Iwam and Abau have five and twelve numeral classifiers respectively (Aikhenvald 2003: 123), while Burmese is thought to have around 200 classifiers (Aikhenvald 2003: 103), and the Mae Chaem dialect of S’gaw Karen has approximately 90 numeral classifiers (Ratanakul 2001b).

\(^2\) Comprehensive lists of known languages with numeral classifiers can be found in Gil (2013) and Aikhenvald (2003).
Scholarly interest in numeral classifier systems flourished in the 1970s. The earliest studies focused on semantic descriptions of numeral classifiers within individual languages, such as Burling's (1965), Pe's (1965), and Becker’s (1975) work on Burmese classifiers. Cross-linguistic comparative studies focusing primarily on the syntax of numeral classifier constructions were conducted by Jones (1970) and Goral (1978) on numeral classifiers in Southeast Asia and Greenberg (1972, 1975) on languages worldwide. Typological studies on the semantics and functions of numeral classifier systems cross-linguistically established the semantic parameters for numeral classifier choice among speakers (Denny 1976, 1986; Adams & Conklin 1973), the role of
numeral classifiers in individuating the noun for enumeration (Croft 1994), the use of
numeral classifiers in anaphoric constructions (Allan 1977), and the role of the numeral
classifier in marking definiteness (Bisang 1999). More recently, a number of
morphosyntactic studies have examined numeral classifiers within the larger typology of
noun categorization, including Craig (1992), Grinevald (2001), and Aikhenvald (2003).

The following sections address the semantics, typology, and functions of
numeral classifier systems in more detail. The last section describes the structure of
noun phrases when numeral classifiers are present.

3.1.2 SEMANTICS OF NUMERAL CLASSIFIER SYSTEMS

In languages with numeral classifier systems, the numeral classifiers are
obligatory in the presence of a numeral, as shown in 5 and 6.

(5) Hmong (Bisang 1999: 122)
    ib    rab    riam
    one    CLF    knife
    ‘one knife’

(6) Thai (Allan 1977: 286)
    mǎˑ sìˑ tua
    dog    four    CLF
    ‘four dogs’
The numeral classifier typically specifies some semantic feature of the noun. In Hmong, /rab/ is used with nouns denoting implements and weapons (Jaisser 1987: 174), and in Thai, /tua/ is used with nouns denoting animals, tables, trousers, and shirts (Deepadung 1997: 54).

Much of the scholarly interest in numeral classifier systems has focused on trying to determine the semantics of the classifier system within a particular language.

According to Becker’s (1975: 115) analysis of the Burmese classifier system, all Burmese classifiers are organized on varying levels of distance away from Buddha.³ Goral (1978: 31-35) wrote that “semantically, [classifiers] can be associated with multi-dimensional prototypes, ranging from totally opaque, arbitrary [classifier] - Noun associations to clearly defined systems, as well as every possible mixture of these two extremes,” and he concluded that “semantic chaos […] reigns in the domain of the [classifier … and …] there is no point in searching for an overall pattern.”

³ According to Becker (1975: 166), the center of the Burmese numeral classifier paradigm is /shu/ which is used with Buddhas, the Law, and relics; /'pa/ which is used with deities, saints, and monks; /'u/ which is used with teachers and scholars; /yauʔ/ which is used with ordinary humans; and /kaunʔ/ which is used with animals, ghosts, dead bodies, and depraved people.
Uncovering an overall, logical semantic organization of classifiers and the nouns they pair with is made more difficult by the pragmatic and dialectal use of classifiers.

For example, classifiers are used in the earliest inscriptions found of the Burmese language from the 12th century. As Buddhism flourished in Burma, poetry and writing spread throughout the country and the frequency of classifier use increased (Pe 1965: 168-169). Similarly, in Cambodian, classifiers are only used in written speech and are a sign of formal style (Goral 1978: 26). Because classifier use is so influenced by education, status, and discourse, classifier usage may vary greatly from one person to the next. As Erbaugh (1986: 413) writes, “Level of formality, discourse type, especially narrative and poetry; length of speech turn, presence of the referent, familiarity of the referent, and age of the hearer are all critical, synergistic, and variable determinants of special classifier use.” Attempting to find the semantic basis of a classifier system, in which each classifier and noun pair is logical and consistent, might be an impossible task; however, typological examinations of the morphosyntactic role of numeral classifiers have proven fruitful.
3.1.3 TYPOLOGY OF NUMERAL CLASSIFIERS

There are four types of numeral classifiers: sortal classifiers, mensural classifiers, repeaters, and general classifiers. Many scholars consider sortal classifiers to be “true” numeral classifiers (Grinevald 2001: 64, Bisang 1990: 124). According to Lyons (1977: 115), a sortal classifier “individuates whatever it refers to in terms of the kind of entity that it is.” Examples 7 and 8 illustrate sortal classifiers.

(7) Burmese (Becker 1975: 117)

\begin{verbatim}
lu `le yau?
\end{verbatim}

person four CLF:ORDINARY.HUMAN
‘four people’

(8) Tashkent Uzbek (Beckwith 1998: 131)

\begin{verbatim}
bir bâs karâm
\end{verbatim}

one CLF:HEAD.SHAPED cabbage
‘one cabbage’

Unlike a sortal classifier, a mensural classifier “individuates in terms of quantity” (Lyons 1977: 115). Two features of mensural classifiers differentiate them from measure words. First, mensural classifiers generally follow the same syntactic construction as sortal classifiers, which is a useful distinction in the few languages in which numeral classifier syntax differs from that of measure words (Grinevald 2001: 64). Secondly, mensural classifiers can only be used with a small number of nouns
whereas measure words can be used with almost any noun. In example 9 below, the mensural classifier *han* in Korean “is used exclusively for measuring rice wine in terms of an institutionalized measuring cup” (Aikhenvald 2003: 115). Likewise, /tɛʔ/ in example 10 can only be used with a small number of nouns.

(9) Korean (Lee 1997 from Aikhenvald 2003: 115)

\[
\begin{array}{ccc}
\text{makke}i & \text{han} & \text{mal} \\
\text{rice},w\text{ine} & \text{one} & \text{MENS},\text{NUM},\text{CL}:\text{RICE},\text{WINE} \\
\end{array}
\]

‘one measure of makkeli (rice wine)’

(10) Burmese (Vitrant 2002: 133)

\[
\begin{array}{ccc}
\text{/ \text{Ch}i\text{N} } & \text{ni?} & \text{tɛʔ/} \\
\text{ginger} & \text{two-} & \text{CLF}:\text{SHOOT/GROWTH} \\
\end{array}
\]

‘two small pieces of ginger’

An additional important feature of mensural classifiers is that they do not exist in any language independent of sortal classifiers. Thus, all known human languages have measure words, but there is no language with both measure words and mensural classifiers but no sortal classifiers (Craig 1992: 280).

Repeaters refer to numeral classifiers that are used with either abstract nouns or nouns that are so unique that they belong to a very limited class. According to Goral (1978: 3), abstract entities are not countable and thus cannot be paired with a sortal
classifier. If a language has a low degree of numeral classifier grammaticalization\(^4\), then a speaker simply does not use a classifier when pluralizing abstract nouns. Similarly, because numeral classifiers classify nouns according to salient semantic features of the noun, a noun may be so unique that it does not fit into an existing class in the language. For example, the number of house-shaped objects is likely small; the number of country-shaped objects is extremely small. In such instances, speakers may simply not use a classifier. However, in languages in which numeral classifiers are becoming more grammaticalized, the use of a classifier is obligatory even if a classifier does not exist.

A special kind of classifier, called a repeater, is used in these circumstances. The term repeater is used when the noun is repeated in the classifier position so that the classifier slot is filled, as shown in 11 and 12.

(11) Burmese (Pe 1965: 293)

\[
\begin{array}{ccc}
\text{qéin} & \text{ta} & \text{qéin} \\
\text{house} & \text{one} & \text{CLF:HOUSE} \\
\text{‘one house’}
\end{array}
\]

\(^4\) Grammaticalization is the process by which content words (lexical items of an open class) become function words (grammatical markers of a closed class). The degree of grammaticalization of a numeral classifier system refers to whether a numeral classifier is used primarily in a grammatical way.
Repeaters are found in many numeral classifier languages, including Vietnamese, Chinese, Thai, and Indonesian (Goral 1978: 34-35). Repeaters are an important feature of numeral classifier languages because they indicate the degree of grammaticalization in the classifier system (Aikhenvald 2003: 361).

In a similar fashion to repeaters, many numeral classifier languages have a general classifier that may be used in any situation, see 13 below. General classifiers can be used in situations in which a speaker is not sure how to classify an object because it is new (for example, a new form of technology), but often general classifiers are indicative of social standing or education. For example, there are a number of numeral classifiers that were only used in Burmese theater slang and thus would not be used by common people (Pe 1965: 170). In Burmese, /khù/ is used as a general classifier. Burling (1965: 261) writes that "Burmese speakers readily advise the learner of the language to use khù when in doubt, and I have heard children use khù in situations in which an adult would probably use a more precise classifier.”
repeaters, the extent of use of general classifiers indicates the degree of grammaticalization of the numeral classifier system.

(13) Mandarin (Li & Thompson 1981: 104)

\[
\begin{array}{ccc}
\text{sān} & \text{ge} & \text{rén} \\
\text{three} & \text{CLF:GENERIC} & \text{person} \\
\text{‘three people’}
\end{array}
\]

3.1.4 FUNCTIONS OF NUMERAL CLASSIFIERS

3.1.4.1 ENUMERATION

The primary function of numeral classifiers is enumeration. Croft (1994: 162) writes that “enumeration […] involves two cognitive processes. One must identify multiple units of the same kind in order to count them. The two cognitive processes involved are individualization of units, and identification of them as being of the same kind.” In languages with a mass/count noun distinction, count nouns by themselves are already individuated (e.g. *dog) and are enumerated with plural markers (e.g. *dogs).

Mass nouns are not already individuated and require measure words for enumeration (e.g. *two rices but two cups of rice). In languages without a mass/count noun distinction, a noun by itself is neither singular nor plural. Bisang (1999: 114) provides the example of Chinese xìn ‘letter,’ which “can mean ‘letter, a letter, letters, the letter,
etc.’ according to the given context.” *Xìn* by itself is neither singular nor plural, and it is neither definite nor indefinite.

Languages without a mass/count noun distinction require numeral classifiers in order to individuate the noun referent so that it may be enumerated.\(^5\) Denny (1986: 298) summarizes the general theory of numeral classifiers put forth by Greenberg (1977) that "the noun refers to the collectivity of individuals, whereas the quantifying role of the classifier is to refer to a unit, a single individual, from this collectivity.” Thus, in languages without a mass/count noun distinction, nouns refer to sets of things rather than individual entities. In example 14 below, */rôm/ is by itself neither singular nor plural. The use of a numeral classifier individuates the noun allowing it to be singular in example 15 and plural in example 16. As one would expect, the absence of a numeral classifier in 17 creates an ungrammatical construction.

(14)  Thai (Hundius-Kolver 1983: 172)

| /rôm    | níí    |
| umbrella | DEM    |
| “this/these umbrella(s)” |

\(^5\) Early work on numeral classifiers argued for a typological connection between numeral classifiers and an absence of plural marking on nouns (Greenberg 1972). However, while this is generally the case, more recent research has found languages that employ both numeral classifiers and plural marking (Aikehnvald 2003: 100-101).
3.1.4.2 DEFINITENESS & DISCOURSE PRAGMATICS

Early scholars of numeral classifier languages argued that numeral classifiers are used as markers of definiteness (Goral 1978: 22). Numeral classifiers are often used to distinguish between new and old information, which is likewise one function of definite and indefinite articles in Indo-European languages, as shown in 18. In this example, the indefinite article is used when ‘the man’ is first introduced, and the definite article is used after because ‘the man’ is no longer new information.

(18) I saw a strange man on the street today.
    The man was wearing an animal suit.

While there are some languages that use numeral classifiers as markers of definiteness (cf. Aikhenvald 2003: 117), it may be more accurate to say that this function of numeral
classifiers in classifying old and new information is simply similar to definiteness.

Numerals classifiers may be used in non-obligatory constructions to introduce new referents in a discourse, and thus they are often translated into English using indefinite articles; however, while the indefinite article may be required in English depending on the discourse context, one should not necessarily infer that the numeral classifier is marking indefiniteness, as shown in example 19 where the phrase can be translated as either definite or indefinite.

(19) Hmong (Bisang 1999: 152)

peb tug dev
three CLF dog
a. ‘three dogs (indefinite)’
b. ‘the three dogs (definite)’

Numeral classifiers are often used when a noun referent is first introduced in discourse even if numeral classifier use is not obligatory with the numeral one. Croft (1994: 161) argues that because numeral classifiers specify semantic features of the noun referent, numeral classifiers function in referent tracking, and their function in these situations “is to identify an entity that will be referred to later on in the discourse.” Numeral classifiers can also be used after a referent has already been
introduced in order to highlight that noun referent’s importance in the discourse. As Aikhenvald (2003: 322) writes,

[Numeral classifier use] frequently depends on the definiteness and pragmatic properties of the referent noun: whether it has just been introduced into the discourse; whether it is topically continuous; or whether it is pragmatically salient—i.e. either is in the focus or is important in the discourse.

Aikhenvald (2003: 325) asserts that speakers may use a numeral classifier for a noun referent that has already been introduced when there are multiple referents in the discourse and the speaker wants to “indicate higher salience of one referent relative to the other referents in the local contexts.” Thus, numeral classifiers are used in order to focus attention on important referents in the discourse independent of whether or not the noun referents are definite or indefinite.

3.1.4.3 CLASSIFICATION

Languages with systems of numeral classifiers tend to employ a rich numeral classifier system with different numeral classifiers for different semantic contexts. As a result, a speaker’s choice of numeral classifier highlights one such aspect of a noun’s referent. As Pe (1965: 170-171) writes,
A person may be classed as a human being or as an animal according to his behavior; and a horse or an elephant as merely an animal or as a mount; and a sword as just a straight long thing or as a weapon. The choice of the classifiers is prompted by the occasion.

In this way, numeral classifiers do more than merely individuate, enumerate, and mark the noun for definiteness or importance. They reflect the interaction between the speaker and the object, or as Becker (1975: 111) argues, "classifiers are part of a system for organizing experience."

Speakers of numeral classifier languages can select a numeral classifier for a noun based on the discourse context. For example, in Becker’s (1975: 113) often cited example, eight different classifiers are used with the noun ‘river’ in Burmese depending on the context:

- `myiʔ tə yaʔ` ‘river one place’ (e.g. destination for a picnic)
- `myiʔ tə tan` ‘river one line’ (e.g. on a map)
- `myiʔ tə hmwa` ‘river one section’ (e.g. a fishing area)
- `myiʔ tə́ sin` ‘river one distant arc’ (e.g. a path to the sea)
- `myiʔ tə thewe` ‘river one connection’ (e.g. tying two villages)
- `myiʔ tə́ pa` ‘river one sacred object’ (e.g. in mythology)
- `myiʔ tə khú` ‘river one conceptual unit’ (e.g. rivers in general)
- `myiʔ tə myiʔ` ‘river one river’ (the unmarked case)
In Japanese, referring to an individual with an animal classifier or a classifier used for inanimate objects is a form of insult (Allan 1977: 296). Allan (1977: 297) cites examples of classifier usage in Thai in which the classifier selected when referring to ropes is also used when referring to elephants but not snakes, and the classifier used when referring to animals or bodies is used when referring to snakes, but not elephants, see 20 a-d.

(20) a. Thai (Allan 1977: 297)
   nu· sì· tua
   snake 4 CLF:body
   ‘four snakes’

   b. cháˑŋ sì· chŷag
   elephant 4 CLF:rope
   ‘four elephants’

   c. * nu· sì· chŷag
      snake 4 CLF:rope

   d. * cháˑŋ sì· tua
      elephant 4 CLF:body

3.1.4.4 ANAPHORIC USE

In anaphoric constructions, a morpheme serves as a referent to a previously mentioned person or object. In English, pronouns are used in anaphoric constructions
referring to count nouns, as shown in example 21 a,b; measure words are used in anaphoric constructions referring to mass nouns, as shown in example 22.

(21) a. Have you seen the movie?

    b. Yes, I watched it yesterday.

(22) I drink a lot of tea. I already drank two cups today.

    Numeral classifiers are used in anaphoric constructions in languages that do not have mass/count noun distinctions. Because numeral classifiers specify salient semantic features of the noun referent, they may be used in referent tracking, as shown in examples 23 and 24.

(23) Vietnamese (Bisang 1999: 148)

    `dây là sách Lan mua một quyền tôi mua hai quyền
    this be book Lan buy one CLF I buy two CLF
    `này
    this
    ‘These are books. Lan bought one (of them), I bought these two (of them)

(24) Burmese (Vittrant 2002: 137)

    /θəyeʔθi we yiN `ŋa `luN yu-Khé Pa/
    mango buy if Ø - five CLF:3D take-VM(mvt) POL
    ‘If you buy mangos, bring me back five [round objects]’
3.1.5 SYNTAX OF THE NOUN PHRASE WITH NUMERAL CLASSIFIERS

Allan (1977: 288) postulates a universal principle that "a classifier concatenates with a quantifier, locative, demonstrate, or predicate to form a nexus that cannot be interrupted by the noun which it modifies." The Kegboig languages spoken in Nigeria are the only languages attested thus far in which the numeral classifier forms a constituent with the head noun rather than the numeral (Aikhenvald 2003: 110-111; cf. Okori 1994). Based on analyses of available numeral classifier languages, the classifier generally forms a constituent with the numeral. Only rarely does a classifier form a constituent with another element, such as a determiner or a modifier (Aikhenvald 2003: 101). Greenberg (1972: 31) proposes four possible syntactic noun phrase constructions when the numeral and classifier form a constituent.

Type 1: [NUM - CLF] - N
Type 2: N - [NUM - CLF]
Type 3: [CLF - NUM] - N
Type 4: N - [CLF - NUM]

Type 1 constructions are found in languages such as Chinese, Vietnamese, and Hmong. Thai and Burmese are representative of Type 2 constructions, which are the most
frequently found orders. Type 3 and Type 4 constructions are relatively rare (Aikhenvald 2003: 105).

3.2 THE NUMERAL CLASSIFIER SYSTEM IN S’GAW KAREN

3.2.1 INTRODUCTION

S’gaw Karen is part of the Karenic language group in the Tibeto-Burman language subfamily, which belongs to the larger Sino-Tibetan language family. Numeral classifiers are a characteristic feature of the Sino-Tibetan family and are found in all Tibeto-Burman languages (Bradley 2012: 180). A number of studies have examined the classifier systems of member languages, including Mandarin (Zhang 2007, Fang & Connelly 2008, Tai 1992), Cantonese (Pacioni 1997), Burmese (Vittrant 2002, Burling 1965, Pe 1965), Newari (Bhaskarareo & Joshi 1985), Yi (Bradley 2001), Eastern Kayah-Li (Solnit 1997), and Tibetan (DeLancey 1998), among others. Lists of numeral classifiers and the nouns they pair with in S’gaw Karen are found in Gilmore (1898), Jones (1961, 1970), and Ratanakul (2001b). Jones (1961) also analyzes the structure of the noun phrase when numeral classifiers are present in a variety of contexts.

When a numeral is present in the noun phrase, numeral classifiers are obligatory in S’gaw Karen. This section will describe the types of numeral classifiers that are
found in S’gaw Karen, the function of numeral classifiers, and the structure of the noun phrase involving numeral classifiers.

3.2.2 TYPES OF NUMERAL CLASSIFIERS IN S’GAW KAREN

S’gaw Karen has the following types of numeral classifiers: sortal classifiers, mensural classifiers, and repeaters. Sortal classifiers are used in enumeration and individuate the noun. They also specify some salient feature of the noun they refer to. In examples 25 and 26, the numeral classifiers refer to the shape of the noun referents, while in examples 27 and 28, they refer to the taxonomy of the noun referents.

(25) /tài kʰi-pʰlə/ egg two-CLF:ROUND ‘two eggs’

(26) /ŋà kʰi-bè/ fish two-CLF:FLAT ‘two fish’

(27) /tʰwì kʰi-dú/ dog two-CLF:QUADRUPED ‘two dogs’
There are numerous mensural classifiers in S’gaw Karen. Mensural classifiers group objects into a unit for enumeration. Mensural classifiers are specific units of measurement and can only be used with a limited number of nouns, see 29-31.

(29) əʔ əʔ
/pəʔəʔ/ kʰí-kʰê/
betel.nut two-CLF:SPLIT SECTION
‘two betel nut halves’

(30) əɾ əɾ
/mê kʰí-mê/
rice two-CLF:MOUTHFUL
‘two mouthfuls of rice’

(31) əɾ əɾ
/wə kʰí-kʰê/
.bamboo two-CLF:LATERALLY CUT SECTION
‘two halves of bamboo’

Repeaters are used for abstract nouns or nouns that are difficult to specify saliently with the inventory of numeral classifiers. In example 32, the noun /tʰìkɔ̏/ ‘country’ cannot be paired with a sortal numeral classifier in S’gaw Karen because its salient features cannot be easily specified. Likewise, /klɔʔ/ ‘language’ in example 33 is
an abstract noun and thus cannot be specified with the inventory of numeral classifiers in S’gaw Karen that generally refer to observable features of the noun referent, such as shape and size. In these instances, the noun itself is repeated and affixed to the numeral in the numeral classifier slot.

(32) /ˈtʰíkɔ̥/ kʰí-tʰíkɔ̥/
country two-CLF
‘two countries’

(33) /jɔ́ tɛ́ klɔʔ xɔləmí kʰi-klɔʔ/ I speak language pretty two-CLF
‘I speak two pretty languages’

A speaker’s choice of numeral classifiers seems to vary based on the age, social status, and the level of education of the speaker. For example, with the noun /lá/ ‘moon,’ consultants indicated that older speakers tend to use the numeral classifier /bé/ which specifies the flat nature of an object and is typically used for referring to flat objects, such as leaves and plates, whereas younger speakers that have emigrated to the United States and have been educated use the classifier /pʰlɔ/ which specifies the round
nature of objects. In S’gaw Karen, an individual can refer to a pregnant woman using /pʰlə̀/ which is used for round objects if the speaker wants to focus attention on the woman’s pregnancy rather than some other feature of the woman. A variety of mensural classifiers can be used for noun referents to specify their orientation or use. For example, /wá/ ‘bamboo’ can be paired with /bó/ if the bamboo is a long, straight section, /kʰɔ́/ if it is cut in half, or /kəpɔ́/ if it is cut into hollow, cylindrical sections.

3.2.3 FUNCTIONS OF NUMERAL CLASSIFIERS IN S’GAW KAREN

3.2.3.1 ENUMERATION

Numeral classifiers have an essential role in enumeration in S’gaw Karen and are obligatorily used with all nouns when a numeral greater than one is present in the

---

6 This example illustrates the importance of the speaker in choosing a numeral classifier. I speculate that older speakers use /bé/ because the moon looks like a flat object from earth, while younger, educated speakers use /pʰlə̀/ because they have learned that moons are spherical. Additionally, younger speakers are put into contexts in school where they are more likely to have to enumerate moons (e.g. the moons of Jupiter). Consultants also indicated that older speakers referred to airplanes using /bé/ because planes look flat when flying overhead, while younger, emigrant speakers (who have ridden on planes) use the classifier /bó/, which is used to specify the long, cylindrical features of noun referents.

7 A comprehensive list of classifiers used in the Mae Chaem dialect of S’gaw Karen appears in Ratanakul (2001b).
noun phrase. Numeral classifiers are optional for nouns when the numeral is over one hundred, see 34 a,b.

(34) a. ꝽꝽ ꝿ / tʰwì kʰí-dúu/
    dog   two-CLF:QUADRUPED
    ‘two dogs’

   b. ꝽꝽ Ᵹ Ᵹ Ᵹ Ᵹ / tʰwì takájá/
    dog   one.hundred
    ‘one hundred dogs’

Sortal classifiers are used to individuate noun referents so that they may be enumerated.

Sortal classifiers are limited in number and refer primarily to the shape of the noun referent as shown in example 35.

(35) Ᵹ Ᵹ Ᵹ Ᵹ Ᵹ Ᵹ Ᵹ Ᵹ / tʰymi kʰi-pʰlá/
    television   two-CLF:ROUND
    ‘two televisions’

Mensural classifiers also enumerate the noun but refer to temporary quantities or arrangements of the noun referents. Therefore, a noun may pair with a sortal classifier, example 36, or it may pair with a mensural classifier, example 37, depending on the discourse context.
(36) ɓɓ ɓɓ
/wà  kʰí-bó/
bamboo  two-CLF:LONG CYLINDRICAL

(37) ɓɓ ɓɓ
/wà  kʰí-kʰɔ́/
bamboo  two-CLF:LATERALLY CUT SECTION

3.2.3.2 ANAPHORIC USE

Numeral classifiers play a role in anaphoric constructions in S’gaw Karen. The noun may be omitted when a numeral classifier is present and the noun is recoverable from earlier discourse, as in 38 a,b. Numeral classifiers cannot be used in anaphoric constructions in the absence of a numeral, see 39.

(38) a. ɓɓ ɓɓ: ɓɓ ɓɓ ɓɓ ɓɓ
/nə ʔɛ̀ʔtɗíʔ tɑ̃dɪ tɔnɔ̄ ɦá/
you want egg PL Q
‘Do you want eggs?’

b. ɓɓ ɓɓ: ɓɓ ɓɓ
/jɔ ʔɛ̀ʔʔíʔ kʰí-pʰlə/
I want two-CLF:ROUND
‘I want two.’

(39) * ɓɓ ɓɓ: ɓɓ
/jɔ ʔɔdɔ́ʔ ɗū/ I have CLF:QUADRUPED
* ‘I have a quadruped’
The basic structure of the noun phrase is the same when numeral classifiers and measure words are present, and measure words can likewise be used anaphorically, as in 40 a,b.

(40) a. ¿ njó kʰí nwí kʰwáʔ há/
   you drink water seven MW:glass Q
   ‘Do you drink seven glasses of water?’

   b. ¿ jó kʰí kʰwáʔ/
   I drink two MW:glass
   ‘I drink two glasses’

Because measure words and mensural classifiers occupy the same syntactic position in S’gaw Karen, it is easy to confuse a mensural classifier for a measure word. An important criterion for distinguishing them is that numeral classifiers can never be used as independent lexemes whereas measure words can be, as shown in example 41 a,b.

(41) a. ¿ júŋ tep ti kʰwáʔ/ ʔó
   /dú ʔi ʔawí/
   *CLF:QUADRUPED COP hungry
   * ‘The quadruped is hungry’

   b. ¿ jó pwā kʰí kʰwáʔ/
   I buy two glass
   ‘I buy two glasses’
3.2.3.3 DEFINITENESS & DISCOURSE PRAGMATICS IN S’GAW KAREN

S’gaw Karen numeral classifiers are used to indicate the importance of referents in discourse. When a referent is first introduced in the discourse, a numeral classifier is used even though numeral classifiers are generally not used with the numeral one, see 42.

(42) ṭəbló nè ṭəmíjís pʰó ?òwé tə-duú
one.time then cat little COP one-CLF
‘There was once a little cat.’

In these instances, the sentence is translated into English using the indefinite article, but one should not necessarily infer that the function of that numeral classifier in this instance is to mark indefiniteness. Likewise, when the cat is mentioned again later in the discourse, the speaker does not use a numeral classifier, and the sentence is translated into English using the definite article, see 43.

(43) ṭəmíjís pʰó siʔsʰóʔʔ ʔs
/ṭəmíjís pʰó siʔsʰóʔʔ ʔs/
cat little answers 3.SG.OBJ
‘The little cat answers him’

In these examples, the use of the numeral classifier to indicate new and old information overlaps with the similar function of the definite and indefinite articles in English. This
evidence suggests that the speaker is not marking definiteness of the noun referent; rather, the speaker uses the numeral classifier when introducing the referent to highlight the importance of the referent in the discourse. This function of numeral classifiers is likewise found in other languages with numeral classifiers (cf. Aikhenvald 322-326).

Numeral classifiers in S’gaw Karen can be used with referents that have already been introduced into the discourse. For example, in the tale of Ku Naw Lay and Naw Moo Aye from §A1.4, Naw Moo Aye is kidnapped by a python after her husband, Ku Naw Lay, loses a bet with the python. When the python is first introduced in the discourse, a numeral classifier is used with the noun. However, once the python reappears in the discourse to kidnap Naw Moo Aye, a numeral classifier is again used, as shown in example 44.

(44) คะ平面 นกขี้ม่วง:กิม:ต้ม คะต้ม คะกิม คะกิม
/ʔowé นóxéháʔkwi? báš'bá kaló tó-bó/
3.SG.NOM shoo but python one-CLF

คะที่ นั้น คะ คะ: คะ: คะ คะ: คะ คะ: คะ
/ʔí bixàʔ ?è tʰóʔ xíʔ nèʔ ʔelóʔ/
DEM tie 3.SG.POSS pig firm than before

‘She shooed but this python tied her pig tighter than before.’
Later in the discourse, Ku Naw Lay comes to rescue his wife, and he is topically more important than the python. For that reason, the numeral classifier is no longer used with the python, see 45.

(45) Ku Naw Lay knows that if he cuts his neck then he will die, so he cuts out his hand’s blood and gives it to this big python.

3.2.4 STRUCTURE OF THE NOUN PHRASE WITH NUMERAL CLASSIFIERS

3.2.4.1 THE CLASSIFIER CONSTITUENT IN THE NOUN PHRASE

In S’gaw Karen, when a numeral is present, the numeral classifier is always syntactically bound to the numeral, i.e. no elements may intervene between the numeral and the classifier. Adjectives and relative clauses may intervene between the noun and the classifier constituent. Demonstratives come at the end of the noun phrase after the classifier constituent.
The noun always precedes the classifier constituent, but the structure of the classifier constituent depends on the numeral used. When the numeral is below ten, the structure of the noun phrase is \([N – [\text{NUM} – \text{CLF}]],\) as in 46-48.

(46) \(\text{fish} \quad \text{three-CLF:FLAT}
\quad \text{‘three fish’}

\(\text{teacher} \quad \text{seven-CLF:HUMAN}
\quad \text{‘seven teachers’}

(48) \(\text{trunk} \quad \text{two-CLF:COLUMN}
\quad \text{‘two (tree) trunks’}

When the numeral is above ten, the structure of the noun phrase becomes \([N – \text{PL} – [\text{CLF} – \text{NUM}]].\) The numeral and the classifier are reversed, and the general plural marker /ʔá/ intervenes between the noun the classifier constituent, as in 49 and 50.

(49) \(\text{fish} \quad \text{PL} \quad \text{CLF:FLAT-twenty}
\quad \text{‘twenty fish’}
The use of a numeral classifier becomes optional when the numeral is one hundred or above. Consultants indicated that numeral classifiers are generally not used with large numbers, but that their use is not ungrammatical. This is typical of numeral classifier systems. For example, Telugu does not require classifiers for numerals over ten, Burmese for multiples of ten, and Thai for numerals over one thousand (Aikhenvald 2003: 100). When numeral classifiers are used for numerals over one hundred, the structure of the noun phrase is the same as when the numeral is greater than ten, see 51 a,b.

\[(51)\ a. \quad \begin{array}{lll} \text{bird} & \text{PL} & \text{CLF:FLAT-one.thousand} \\ /tʰò & ?á & kʰkæ̂tʰó/ \end{array} \]

\[\text{‘one thousand birds’}\]

\[\begin{array}{lll} \text{bird} & \text{one.thousand} & \text{‘one thousand birds’} \\ /tʰò & kʰkæ̂tʰó/ \end{array}\]
3.2.4.2 DEMONSTRATIVES IN THE NOUN PHRASE

As was discussed in greater detail in §3.2.2.3, numeral classifiers are generally not used with singular nouns unless they are being marked as important to the discourse. Since numeral classifiers are not obligatory when the numeral is one, demonstratives are primarily found with numeral classifiers when the numeral is greater than one. The demonstrative comes at the end of the noun phrase and follows the classifier constituent when present, as in 52.

(52) ]-'l 휴 Twig ʔi/
/kəlɔ̃  nwí-bó ʔi/
python  7-CLF  those
‘those seven pythons’

3.2.4.3 ADJECTIVES IN THE NOUN PHRASE

In S’gaw Karen, adjectives follow the nouns that they modify. When a numeral classifier is used, adjectives intervene between the noun and classifier constituent, as shown in 53 and 54.

(53) _interaction Twig ʔi/
/jə ʔødɔʔ thwi yə h²i-dtù/
I have dog good two-CLF
‘I have two good dogs.’
Relative clauses occupy the same position as adjectives when a classifier constituent is present, see 55.

Jones (1961: 45-46) argues that when there are more than two adjectives, the adjectives are placed at the end of the noun phrase, but still before the demonstrative as shown in example 56, where the classifier constituent follows directly after the noun. Example 57a illustrates the structure Jones argues is required when more than two adjectives are used in the noun phrase. Consultants indicated that while 57a is grammatically correct, the adjectives can also directly follow the noun, as in example 57b, and still be grammatically correct.
While consultants agreed that examples 57 a, b are grammatically correct, they suggested that it would be more typical for most speakers to modify the noun with a relative clause that follows the noun phrase, as shown in example 58.8

8 While Jones’s argument that moving an adjectival phrase to the end of the noun phrase when more than two adjectives are present might be better described as optional rather than obligatory, it does illustrate what Greenberg (1975) described as the close semantic relationship between the noun and the numeral classifier. Greenberg argued that, because of this close relationship, the numeral classifier is always found in close proximity to the noun it pairs with and can never be separated by too many elements. Solnit’s (1997) description of Eastern Kayah-Li in §3.2.4.4 suggests otherwise.

(57) a. /jə ?òdóʔ tʰwì kʰí-dúʔ ðúðú yẽ pʰáʔdò/ I have dog 2-CLF black good big

‘I have two black good big dogs’

b. /jə ?òdóʔ tʰwì ðúðú yẽ pʰáʔdò kʰí-dú/ I have dog black good big 2-CLF

‘I have two big black good dogs.’

(58) /jə ?òdóʔ tʰwì ðúðú ló ?ə yẽ pʰáʔdò/kʰí-dú/ I have dog black that are good big

‘I have two black dogs that are good and big.’
3.2.5 THE S’GAW KAREN NUMERAL CLASSIFIER SYSTEM WITHIN THE TIBETO-BURMAN FAMILY

Unfortunately, there is little research on the other Karenic languages, a notable exception being Solnit (1997). Duffin’s (1913) description of Pwo Karen lists some common numeral classifiers, which Duffin labels “numeral adjectives,” but Duffin’s analysis of numeral classifiers is limited to the function of enumeration. Like S’gaw Karen, numeral classifiers in Pwo Karen precede the numeral when the number is greater than ten and the plural marker /ʔə/ intervenes between the noun and the classifier constituent. Duffin (1913) provides the basic structure of the noun phrase when numeral classifiers are present, which like S’gaw Karen is [N – [ NUM – CLF ]], but he does not provide examples of noun phrases involving numeral classifiers with adjectives or demonstratives.

Solnit’s (1997) treatment of numeral classifiers in Eastern Kayah-Li focuses on the syntax of noun phrases and classifier phrases. Solnit (1997) also provides a list of common numeral classifiers in Eastern Kayah-Li and the nouns that they pair with. The structure of the classifier constituent is generally the same as in S’gaw Karen, [NUM –

---

9 Duffin only provides examples in the Pwo Karen orthography and his description of the phonological system of Pwo Karen precedes the creation of the IPA conventions.
CLF], but there are a number of classifiers which require the reverse order (Solnit 1997: 203-08). For example, in example 60, the classifier typically precedes the numeral when the classifier is used for a human. An important difference between S’gaw Karen and Eastern Kayah-Li is that the preferred position of the classifier constituent is clause-final.

(60) Eastern Kayah-Li (Solnit 1997: 7)

púcè ?e thō á di sí lwī
child eat finish PART rice CLF:HUMAN four

‘Four children have eaten’

According to Solnit (1997: 194), the classifier constituent is generally the final element in a clause, or it is the penultimate element if a sentence particle is present. In this case, the classifier does not need to be in the noun phrase with the noun that it is classifying, and Solnit (1997: 182) argues that the classifier constituent is part of the verb phrase instead. Moreover, Solnit (1997: 160-61) asserts that Eastern Kayah-Li is unique in this respect and postulates that the classifier constituent occupies a special syntactic position, which he terms the “extent expression,” otherwise only filled with sentence particles. It is interesting to note that measure words do not follow this syntactic pattern (Solnit 1997: 180). The reader is referred to Figure 2 in §1.2.1 on the Karenic language subfamily.
Of all the Tibeto-Burman languages, Burmese has received the most attention from scholars of numeral classifiers. Burmese classifiers are used similarly to classifiers in S’gaw Karen for individuation and enumeration, anaphoric use, and marking a referent as important in discourse. Burmese numeral classifiers are always found in the noun phrase, and the noun phrase with numeral classifiers is the same as S’gaw Karen:

\[ \text{[N – [NUM – CLF]]} \], as in 61.

(61) Burmese (Vittrant 2002: 135)

/lɛʔpʰɛʔ-ye niʔ Khweʔ/
drinking tea two- CLF:hollow container
‘two cups of tea’

3.3 DISCUSSION

This study examines four functions of numeral classifiers in S’gaw Karen. Numeral classifiers individuate the noun so that it may be enumerated. Numeral classifiers specify salient features of the noun referent. They may be used in referent tracking, allowing the noun to be omitted in anaphoric constructions when the noun can be recovered from previous discourse. Outside of obligatory uses, like in enumeration, speakers may optionally use numeral classifiers in order to highlight nouns that are important in discourse.
This last use of numeral classifiers has only been recently examined in numeral classifier languages. Previously, scholars thought that a general function of numeral classifiers was marking definiteness (cf. Goral 1978), but it now appears that this only happens in few languages (Aikhenvald 2003: 117). While definiteness can indicate whether or not a speaker or an object in discourse is new or has already been introduced, numeral classifiers can also be used with old information in discourse to highlight its importance. Furthermore, in these constructions, the use of numeral classifiers is optional. For these reasons and based on the examples in the current corpus, I propose that numeral classifiers do not mark definiteness in S’gaw Karen; rather, the functions of numeral classifiers in discourse and the function of definiteness in other languages overlap to some degree.

A more in depth examination of the pragmatic use of numeral classifiers in discourse would require a corpus of spoken S’gaw Karen. I plan on building a corpus of spoken S’gaw Karen as I continue my studies. Such a corpus would allow for an examination of a variety of features that I encountered in this project. The S’gaw Karen diaspora makes the creation of a spoken S’gaw Karen corpus problematic. On one hand, dialects of S’gaw Karen that previously may have been identified by the
geographic region of its speakers can no longer be distinguished on this basis because many speakers have fled to Thailand and much of Eastern Burma has been destroyed in war. Many resettled S’gaw Karen speakers were born in Thai refugee camps, educated in those camps in standard S’gaw Karen, and spent as much as two decades within the fences of the camp. Thus, in building a spoken corpus, one must take into account the refugee camps that an individual has lived in, whether or not the individual was born there, the dialects or languages spoken by the individual’s parents and family, and where in Burma the individual’s family lived prior to the refugee camps. While such a corpus would not reflect the S’gaw Karen language and dialects spoken in the hills of Eastern Burma, it could result in interesting findings on the effects of language change as a result of diaspora and the influence of the numerous languages spoken in the refugee camps. Furthermore, even though S’gaw Karen has an estimated four million speakers, third-country resettlement, continuing ethnic conflict in Burma, and low rates of L1 literacy threaten the survival of S’gaw Karen, and a spoken corpus of S’gaw Karen would help to preserve this endangered language and culture.
BIBLIOGRAPHY


Sociolinguistics in Southeast Asia 5.109-21.


BISANG, WALTER. 1996. Areal typology and grammaticalization: Processes of grammaticalization based on nouns and verbs in East and mainland South East


Baptist Mission Press.


GREENBERG, JOSEPH. 1972. Numeral classifiers and substantive number: Problems in the

GREENBERG, JOSEPH. 1975. Dynamic aspects of word order in the numeral classifier.

*Word order and word order change*, ed. by Charles N. Li. 27-46 Austin:

University of Texas Press.

GRINEVALD, COLETTE. A morphosyntactic typology of classifiers. *Systems of nominal
classification*, ed. by Gunter Senft. 50-92. Cambridge: Cambridge University
Press.


Baptist Mission Press.

HAUDRICOURT, ANDRÉ-GEORGES. 1953. À propos de la restitution du Karen Commun.


*Southeast Asian tribes, minorities, and nations*, ed. by P. Kunstadter. 93-124.


LÖBEL, ELISABETH. 2000. Classifiers vs. genders and noun classes: A case study in Vietnamese. *Gender in grammar and cognition*, ed. by Barbara Unterbeck and


MYANMAR LOCATION MAP. 2010. Courtesy of Uwe Dedering. Online: http://upload.wikimedia.org/wikipedia/commons/3/30/Myanmar_location_map.svg


TAI, JAMES. 1992. Variation in classifier systems across Chinese dialects: Towards a
cognition-based semantic approach. *Chinese languages and linguistics 1: Chinese
dialects*, ed. by James Tai and Frank Hsueh. 587-608. Taipei, Taiwan: Academia
Sinica.

Clarendon Press.

THOMAS, DOMINIKE, AND MARC BRUNELLE. 2010. The tale of Kinole and Nomie. *Sgaw
Karen Papers: Presented to Andrew Nimrod*, ed. by Marc Brunelle. Ottawa:
University of Ottawa, MS.

THURGOOD, GRAHAM. 1985. Pronouns, verb agreement systems, and the subgrouping of
Tibeto-Bruman. *Linguistics of the Sino-Tibetan area: The state of the art*, ed. by
Graham Thurgood, James Matisoff, and David Bradley. 376-400. Canberra:
Australian National University.


ZHANG, NIINA NING. 2012. Numeral classifier structures. Minxiong, Taiwan: National Chung Cheng University, MS.
APPENDICES
‘A Rabbit per Bush.’

‘Once upon a time, there are many rabbits in a forest.’

‘They have a king.’
‘Their king, if he tells something then they listen, and if he asks them to do something, then they do it out of peace and love.’

‘If their enemy comes, then they resist with the same heart.’
A1.1.5

so  enemy  PL  come  do  close  rabbit

‘So enemies are not able to come close to the rabbits.’

A1.1.6

food  if  have  also  give  each. other  and

‘If they have food, they also give it to each other and eat the same.’

A1.1.7

when  3.PL.POSS  king  punish  then  3.PL.OBJ  for

‘When their king punishes them, then they agree and accept.’
A1.1.8

In this forest, there is a tiger.

A1.1.9

One day when this tiger is walking around in the forest, he sees some rabbits.

A1.1.10

Because he wants to eat a rabbit, the tiger goes slowly.
‘But those rabbits, when they know their enemy is coming, they pick up their ears, growl, and listen.’

‘They have prepared themselves to face their enemy.’
A1.1.13

This tiger, when he sees rabbits from afar, he is not able to go close.

A1.1.14

He looks at the rabbits pick up their ears; to him they look like sharp spears so he is very scared of them.
'The tiger wants to eat the rabbit’s meat until his spit drips and he cannot control himself, so he plots like he will eat the rabbit’s meat.'

'The tiger slowly moves himself close to the rabbits’ place.'
'He says to the rabbits, “Friends, I look at you all and I see that you have only one king even though you are all powerful; you all possess strength.’
‘If you all have many kings, your power will be bigger than it is, and you will rule this forest.’

‘If the tiger sees a rabbit, he speaks out his words like this, and if he sees one later, he says it again.’
The tiger does this for a long time, and the rabbits’ hearts are not similar anymore.’

‘They think too proudly of themselves and only want to become king.’

‘If their king tells them something, they do not listen, do not follow, and do not do what is right anymore.’
A1.1.23

Later, the rabbits go live among many bushes, one rabbit per one bush.

A1.1.24

If their enemy comes, then they call out for each other, but they do not hear and do not help each other anymore.

A1.1.25

If they hear another ask for help, they do not go help anymore.
In happiness, the tiger goes, catches, and eats one rabbit each day, and soon the rabbit generation will be gone.
'Bad plans.'

'Once upon a time, there are two kinds of animals in a garden: a wildcat and a little chicken.'

'The wild cat himself schemes to catch and eat the little chicken all the time.'
'The little chicken knows that the wildcat wants to eat him so he sleeps among leaves so that you all can’t see him well.'

'One night, the little chicken hides inside a dried gourd.'
'When it is night, the wildcat comes to search for the little chicken in order to eat him.'

‘The little chicken is greatly scared and is shaking so that the gourd falls and breaks.’
The gourd falls loudly and startles the wildcat. While running, he trips over a log, hits his head on a rock, and dies.
A1.3

A cat that dumb one

‘One stupid cat.’

A1.3.1

A one.time then cat little COP one

‘There is a little cat.’

A1.3.2

A 3.SG.NOM stand by fire place beside and bake

A eat crab little

‘He stands by the fireplace and bakes and eats a little crab.’
'During this, a monkey comes to where he is and says, “Can I help you?”'

A1.3.4

\text{θàmíjò} \hspace{0.5em} pʰó \hspace{0.5em} s̥ʔá\text{hò} \hspace{0.5em} ?è \hspace{0.5em} bàmənúü \hspace{0.5em} ?óxó

cat \hspace{0.5em} little \hspace{0.5em} answers \hspace{0.5em} 3.SG.OBJ \hspace{0.5em} why \hspace{0.5em} because

\text{təθé}

NEG.can

‘The little cat answers him, “why not?”’
A1.3.5

say thank for offer help

because

“I say thanks to you because of your offer to help me.’’

A1.3.6

need help

“I need your help.”

A1.3.7

like that pull up crab little from fire

place inside and open

If you say so, pull the little crab from the fireplace and me, I will open your little crab’s shell.’’
112

A1.3.8

The little cat believes the little monkey so he does as the monkey tells him so.

A1.3.9

But that crooked-heart monkey stands behind the little cat.
The monkey takes the crab from the little cat’s hand then opens the shell and eats one piece after another himself.’

‘But the little cat doesn’t know anything about it.’
In the end, the little cat never eats any crab and the front of his little hand is scorched by the fire.
A1.4
Ku Naw Lay and Naw Moo Aye

‘Ku Naw Lay and Naw Moo Aye.’

A1.4.1
one.time person couple COP two-CLF 3.PL.POSS name

mèwé kʰùmölé dɔʔ nɔmùʔé
COP Ku Naw Lay and Naw Moo Aye

‘One time there is a couple named Ku Naw Lay and Naw Moo Aye’

A1.4.2
3.PL.NOM have 3.PL.POSS dog little one-CLF 3.PL.POSS

tʰóʔ to-dú dɔʔ tʰõlwĩ pʰó tə-bè
pig one-CLF and dove little one-CLF

‘They have one dog, one pig, and one dove.’
A1.4.3
tənī nē ?ə wā lē lō pwā pū dō?
one.day then 3.SG.POSS husband go to forest in and

tənī nē oonē lō ə pōli dō:
then 3.SG.POSS 3.SG.REFL go to

The snake who changes himself to a human speaks to Ku Naw Lay like this, “you want to wrestle with me?”
2.SG.NOM if lose then 2.SG.NOM must give

1.SG.OBJ 3.SG.POSS wife

‘If you lose, then you must give me your wife.’

1.SG.TOP 1.SG.NOM if lose then something 2.SG.NOM

want then ask 1.SG.OBJ 1.SG.NOM FUT give 2.SG.OBJ

‘Me, if I lose, then something you want, ask me, and I will give it to you.’”

and 3.PL.NOM start wrestle 3.PL.RELF and

Ku Naw Lay lose so return to 3.SG.POSS house

‘And they start to wrestle together and Ku Naw Lay loses, so he returns back to his house.’
Another day at dawn, Ku Naw Lay tells his wife, “I will go trade in a foreign land.’

‘Stay in the house, be careful.’

‘If you feed the pig food, don’t step down on the ground.’”
A1.4.11

Pour it down from inside the house.’

A1.4.12

And he commands his wife and goes.’

A1.4.13

‘Naw Mu Aye stays in the house with her pig, dog, and little dove.’
A1.4.14

Later on the next day, Naw Mu Lay feeds the pig food and does as her husband commands her.

A1.4.15

When she pours the pig food, a big python comes and wraps around her pig.
‘She shooed but this python tied her pig tighter than before.’

‘Then when she drops her comb, she steps down and picks it up, then this python comes to where she is and pulls her back to his hole.’
'The little dove sees all these things that happen and goes fly to where Ku Naw Lay is and chirps like this “the python constricted Naw Mu Aye; Ku Naw Lay come back quickly.”'

A1.4.19

when Ku Naw Lay hear 3.SG.Poss dove little chirp

then 3.SG.Nom come back quick~quick to 3.SG.Poss house

‘When Ku Naw Lay hears his little bird chirp, he comes back quickly to his house.’
A1.4.20

but  3.SG.NOM  NEG  see  3.SG.POSS  wife  no.more

‘But he didn’t see his wife anymore.’

A1.4.21

so  prepare  3.SG.REFL  that  FUT  chase  back

‘He worries about his wife very much so he prepares himself to get his wife back.’

A1.4.22

‘He calls his dog and goes to get his wife back.’
When his little dog barks at a place, he digs and reaches the snake’s hole and he sees his wife and the python for a moment and his wife and python disappear.

Because he worries so much, he doesn’t know himself to be tired anymore.
A1.4.25

\[ \text{He digs in two or three places, then he sees his wife sitting by the python beside her.} \]

A1.4.26

\[ \text{‘Then python tells him “if you want your wife back, then you must cut out your neck’s blood for me.”} \]
Ku Naw Lay knows that if he cuts his neck then he will die, so he cuts out his hand’s blood and gives it to this big python.

But the python tells him “this is not your neck’s blood, so I will not give your wife back.”
‘Ku Naw Lay cannot do anything anymore.’

‘Finally, he cuts out his neck’s blood and gives it, and the big python lets his wife go.’
'But his wife sees that her husband dies for her and she feels bad, so when people burn her husband’s body, she also steps into the fire and dies.'
APPENDIX TWO: THE S’GAW KAREN WRITING SYSTEM

A2.1 THE S’GAW KAREN WRITING SYSTEM

The S’gaw Karen writing system was developed by Rev. Jonathan Wade in 1834. Wade borrowed the writing system from Burmese and adapted it for S’gaw Karen. The Burmese writing system is itself based on Old Mon, which developed from Brāhmī (Myanmar-Language Commission iv-viii). The Burmese and S’gaw writing systems both feature circular symbols because dried palm leaves were used as writing material beginning in the 17th century, and straight lines would result in the palm leaves being torn by the writing stylus; however, early Burmese graphemes were square rather than circular since most writing was still inscribed in stone (Lieberman 2003: 136).

Because the S’gaw Karen writing system was purposefully created, as compared to one that developed over a long period of time, it is a fairly accurate representation of the spoken language. Additionally, there is neither an authority responsible for maintaining a standardization of the language nor is there an organized system of public education in S’gaw Karen that could maintain orthographical conventions that do not reflect the spoken language. Those readers interested in the S’gaw Karen writing system are referred to Saw Lar Baa’s (2001) master’s thesis The Phonological Basis of a Northwest
Karenic Orthography, which includes a suggested method of expanding the S’gaw Karen writing system to also represent dialectal variations of S’gaw and other Karenic languages.

The S’gaw Karen writing system is an example of an abugida. Abugida writing systems are well-suited to represent languages that have a consonant-vowel syllable structure. Abugidas are similar to abjad writing systems, like those used in Hebrew and Arabic, in which the main written unit of a syllable can only represent a consonant. While abjad systems do not include vowel symbols, which are inferred by the reader, abugidas indicate vowels through diacritic marks affixed to the consonant symbol.

Abugidas differ from both alphabets in that vowels are not represented by a symbol that can be used as a main syllable (rather than as a diacritic) and from syllabaries, like those used in Cherokee and Japanese, in that consonant symbols and vowel diacritics can take many combinations rather than there being a different symbol representing each consonant and vowel combination.

In S’gaw Karen, the central symbol of a syllable represents a consonant. An optional medial consonant can be written under the main consonant symbol, forming a consonant cluster (see §2.2.4 and §A2.3 below). Vowels are represented with diacritic
marks around the consonant symbol, including above, below, and to the right of the
consonant symbol. Tone values are represented with a separate symbol to the right of
the consonant/vowel combination, with the exception of the mid-high tone which is not
representing orthographically (see Figure 6).

![Figure 6: The Structure of the S’gaw Karen Syllable for /kʰwi/ ‘saw’](image)

Much of the literature on S’gaw Karen available in the United States and on the
internet provides transliteration guides between the S’gaw Karen writing system and
English. These guides illustrate the transliteration conventions established by Reverend
Wade, and they are not IPA transcription guides since Wade created his writing system
before the creation of the IPA. Baa’s (2001) thesis is, to my knowledge, the only
publication which provides a correlation between the S’gaw Karen writing system and
the IPA for each symbol. As Baa’s work can be difficult to access in the United States,
I have provided the equivalences between the S’gaw Karen writing system and IPA.
symbols according to my proposed phonological system (See §2.2 for differences between my proposal and Baa 2001).

A2.2 S’GAW KAREN CONSONANTS

Below are the equivalences between the S’gaw Karen writing system and IPA for consonants (see Table 4):
Table 4: Written Representation of S’gaw Karen Consonants

<table>
<thead>
<tr>
<th>Stops</th>
<th>Fricatives</th>
<th>Nasals</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>θ</td>
<td>m</td>
</tr>
<tr>
<td>pʰ</td>
<td>s</td>
<td>n</td>
</tr>
<tr>
<td>b</td>
<td>sʰ</td>
<td>ŋ</td>
</tr>
<tr>
<td>t</td>
<td>ʃ</td>
<td>ŋ</td>
</tr>
<tr>
<td>tʰ</td>
<td>x</td>
<td>η</td>
</tr>
<tr>
<td>d</td>
<td>θ</td>
<td>ʍ</td>
</tr>
<tr>
<td>k</td>
<td>ʃ</td>
<td>η</td>
</tr>
<tr>
<td>kʰ</td>
<td>ʃ</td>
<td>ʍ</td>
</tr>
<tr>
<td>?</td>
<td>ʃ</td>
<td>ʍ</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricates</td>
<td>Lateral</td>
<td></td>
</tr>
<tr>
<td>tʃ</td>
<td>l</td>
<td></td>
</tr>
<tr>
<td>dʒ</td>
<td>ʃ</td>
<td>ʍ</td>
</tr>
</tbody>
</table>

A2.3 S’GAW KAREN CONSONANT CLUSTERS

Below are the equivalences between the S’gaw Karen writing system and IPA for consonant clusters (see Table 5):
Table 5: Written Representation of S’gaw Karen Consonant Clusters

<table>
<thead>
<tr>
<th>Consonant Cluster</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>pw</td>
<td>_pedido</td>
</tr>
<tr>
<td>ml</td>
<td>pedido</td>
</tr>
<tr>
<td>lw</td>
<td>pedido</td>
</tr>
</tbody>
</table>

A2.4 S’GAW KAREN VOWELS

Below are the equivalences between the S’gaw Karen writing system and IPA for vowels (see Table 6). Vowels are never syllable initial and must be preceded by a consonant. Often one finds claims that vowels may be syllable initial, but in those
instances vowels are actually preceded by a glottal stop. Note that the equivalences are all shown with a mid-high tone, which is the default tone when a tone phoneme is otherwise not explicitly written (see §A2.5 and §A2.6).

Table 6: Written Representation of S’gaw Karen Vowels

<table>
<thead>
<tr>
<th>S’gaw</th>
<th>IPA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ᵃí</td>
<td>ə̃́</td>
</tr>
<tr>
<td>ᵃú</td>
<td>ə̃</td>
</tr>
<tr>
<td>ᵃé</td>
<td>ə̃</td>
</tr>
<tr>
<td>ᵃë</td>
<td>ə̃</td>
</tr>
<tr>
<td>ᵃó</td>
<td>ə̃</td>
</tr>
</tbody>
</table>

A2.5 S’GAW KAREN TONES

Below are the equivalences between the S’gaw Karen writing system and IPA for tones (see Table 7). Note that the mid-high tone is not represented by a symbol (see §A2.6).
Table 7: Written Representation of S’gaw Karen Tones

<table>
<thead>
<tr>
<th>Tone Description</th>
<th>Symbol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mid-high</td>
<td>n/a</td>
</tr>
<tr>
<td>Mid</td>
<td>ᶜ</td>
</tr>
<tr>
<td>Mid-low</td>
<td>ʢ</td>
</tr>
<tr>
<td>Low</td>
<td>ʢ</td>
</tr>
<tr>
<td>High falling checked by a glottal stop</td>
<td>ʢ</td>
</tr>
<tr>
<td>Low falling checked by a glottal stop</td>
<td>ʢ</td>
</tr>
</tbody>
</table>

A2.6 /a/ WITH NO TONE PHONEME AND THE MID-HIGH TONE

It is not clear why Wade (1861) chose not to create a symbol for the mid-high tone. Likewise, the vowel /a/ is frequently not represented with a diacritic. As a result, in certain instances the tone or the vowel must be inferred. Below are three rules used to write a syllable in instances when the vowel is /a/ without a tone phoneme, when a vowel has a mid-high tone phoneme, or when the vowel is /a/.

1. If the vowel of a syllable is /a/ and has no tone phoneme, then a consonant symbol is written alone with neither a vowel diacritic nor a tone symbol.

2. If a syllable has a mid-high tone phoneme, then only the consonant symbol and vowel diacritic are written and the tone is not represented by a symbol. It is only
in this instance, when the tone is mid-high, that /a/ is represented with a
diacritic.

3. If the syllable has the vowel /a/ and any tone phoneme other than mid-high, then
only a consonant symbol and tone symbol are written but /a/ is not represented
with a vowel diacritic.
IRB EXEMPTION

MIDDLE TENNESSEE
STATE UNIVERSITY

12/5/2013

Collin Olson
English
Collin.olson@mtsu.edu

Protocol Title: A S’gaw Karen Corpus, Phonology, and Discussion of Some Grammatical Features
Protocol Number: #14-164

Dear Investigator(s),

Your study has been designated to be exempt. The exemption is pursuant to 45 CFR 46.101(b)(2). Educational Tests, Surveys, Interviews, or Observations.

We will contact you annually on the status of your project. If it is completed, we will close it out of our system. You do not need to complete a progress report and you will not need to complete a final report. It is important to note that your study is approved for the life of the project and does not have an expiration date.

The following changes must be reported to the Office of Compliance before they are initiated:
- Adding new subject population
- Adding a new investigator
- Adding new procedures (e.g., new survey, new questions to your survey)
- A change in funding source
- Any change that makes the study no longer eligible for exemption.

The following changes do not need to be reported to the Office of Compliance:
- Editorial or administrative revisions to the consent or other study documents
- Increasing or decreasing the number of subjects from your proposed population

If you encounter any serious unanticipated problems to participants, or if you have any questions as you conduct your research, please do not hesitate to contact us.

Sincerely,

Kellie Hilker
Compliance Officer
615-494-8918
COPYRIGHT PERMISSION

Collin Olson  
1904 Cross Creek Dr  
Murfreesboro, TN 37127  

December 6, 2013  

Dear Tim Moffatt of the Drum Publication Group,  

I am a graduate student at Middle Tennessee State University. I would like your permission to translate four short stories in S’gaw Karen from one of your publications in order to build a corpus with which I can study various elements of the S’gaw Karen language. I will not reproduce any of the text of this publication in S’gaw Karen, but I will use the translated English version.  

The publication that I would like to use is:  


Please indicate your approval of this permission by signing the letter where indicated below and returning it to me as soon as possible. My email address is CollinSOlson@gmail.com. Your signing of this letter is to confirm that you or your company, Drum Publications, owns the copyright to the above described material.  

Thank you very much.  

Sincerely,  

________________________  
Collin Olson  

PERMISSION GRANTED FOR THE USE REQUESTED ABOVE:  

________________________  
Tim Moffatt  
Drum Publication Group  
P.O Box 66  
Kanchanaburi 71000  
Thailand