Sgaw Karen papers

Presented to Nimrod Andrew

Papers written for the course
LIN 4370/7911 Field Methods (Fall 2010)
Instructor: Marc Brunelle

University of Ottawa
April 2011

Foreword

This volume is a collection of papers written by the students of a Field Methods course on Sgaw Karen that took place at the University of Ottawa in the Fall 2010. The goal of the course was to uncover as much as possible about the grammar of Sgaw Karen with the help of a native speaking consultant, Nimrod Andrew. During the first half of the semester, students had to figure out the basic grammatical structure of the language (phonology, morphology, syntax, semantics), without consulting grammars or previous academic work. During the second half of the semester, students had to choose individual topics and to explore them through individual elicitation sessions with Nimrod Andrew and consultation of pre-existing scholarly resources.

The papers presented here are the result of this work. As the scope and difficulty of the topics and the level of the students (both undergraduate and graduate) greatly vary, some are more exploratory while others are exhaustive and almost publishable papers (a few students also chose not to include their papers in this volume). They are organized by sub-discipline and all begin with a short abstract that provides the reader with a non-technical summary of the content.

In my own name and in the name of all the students, I want to express my appreciation and gratitude to Nimrod Andrew. The course would not have been possible without his dedication, hard work and sharp intuitions. More importantly, his love of the language, his enthusiastic attitude and his unrelenting patience made this project an enjoyable and stimulating endeavour for all those who took part in it.

Marc Brunelle
Course instructor
May 31, 2011

TABLE OF CONTENTS

Phonetics

Morphology and Phrasal Syntax

Sentence Syntax and Information Structure

Text

On the Three-Way Plosive Contrast of Sgaw Karen

Dominike Thomas
Leonardo Alves-Soares

Sgaw Karen, a Sino-Tibetan language spoken in Burma, is traditionally described as having a three-way aspiration contrast in plosive consonants (Abramson 1995:155, Gilmore 1898:5). This contrast consists of voiced unaspirated plosives ([b, d]), unvoiced unaspirated plosives ([p, t, k]) and unvoiced aspirated plosives ([p^h, t^h, k^h]). Abramson (1995), the only other acoustic study on the subject, studied this contrast, in Pwo and Sgaw Karen, in isolated words and with very few tokens. This paper aims to further investigate the three-way contrast in Sgaw Karen, with an experiment designed for a statistical analysis of tokens in a carrier phrase.

1.0 Introduction

Sgaw Karen is a Sino-Tibetan language, belonging to the Tibeto-Burman branch, spoken in Burma (Abramson 1995: 156). Like other dialects of Karen (Phillips 2009 and Kato 1995 for different dialects of Pwo Karen), it is traditionally described as having a three-way contrast in plosives (Abramson 1995: 155, Gilmore 1898: 5). This contrast consists of voiced unaspirated plosives ([b, d]), unvoiced unaspirated plosives ([p, t, k]) and unvoiced aspirated plosives ([p^h, t^h, k^h]). The velar sound [k] is not phonologically voiced (Abramson 1995: 156). This contrast differs from the two-way contrast found in English. Gilmore (1898: 5) specifies that Karen has sounds that are "intermediate between ***p*** and ***b***" and "between ***t*** and ***d***".

Only one study (as far as we know) has been done acoustically regarding this three-way contrast in Karen. Abramson (1995) studied and compared the VOT (voice onset time) values of plosives in isolated words of Pwo Karen and Sgaw Karen. He was able to conclude that there is, indeed, a three-way contrast. VOT values for voiced consonants were negative, while the VOT values for aspirated plosives were much higher than for plain plosives. VOT values were therefore sufficient to delimit voiced, unaspirated and aspirated plosives (1995: 163). However, Abramson's (1995) study was mostly limited to words in citation forms, with very few tokens (sometimes as few as 1 per category).

We therefore designed this experiment to further investigate the three-way contrast in Sgaw Karen. Our experiment placed Karen words into in a carrier phrase. We decided to utilize such configuration so that we could properly assess the statistical significance of our results.

2.0 Methods

Our subject is a 27 year-old native speaker of Sgaw Karen who immigrated to Canada 6 years ago. He speaks English (which he learned in a refugee camp), Burmese (which he learned in school), and a little bit of Thai (learned informally) as second languages. Despite not having formal linguistic education, he teaches Karen and has also taken courses on second language teaching. We met the speaker in Field Methods (LIN7911, Fall 2010), a linguistics class offered to both graduate and advanced undergraduate students at the University of Ottawa in which we had to decipher the basic grammar of the Karen language.

For this study, we designed a word list based on class transcriptions of isolated words. This list included words with aspirated, plain, and voiced plosives. Initially, we had also included aspirated and plain alveo-dental fricatives. However, due to perception issues, our transcriptions of words containing such sounds were not always correct, and we decided not to use them in our experiment. Our participant corrected our word list and provided us with a Karen orthographic transcription, but due to time constraints, we were not able to re-organize the list in a more balanced way.

The words from the word list were included in a carrier phrase 'Say X again' in Karen ([tɛ ... kə dɔ tə blɔ]). The speaker's task was therefore to say the meaning of the word in English, the word in isolation in Karen, and the word in the carrier phrase in Karen. The target word is the only variable word in the sentence, whereas we used the word [tɛ] as a reference word to offset any inter-sentence variation.

Recordings were made in a sound booth using a Shure KSM44 microphone (positioned a few inches away from the mouth on a microphone stand) and a Symetrix SX202 dual mic preamp. A Mac Pro (Session 1) or Mac Book Pro (Session 2) was used, with Audacity, to record the sessions. The recordings were made in two different sessions, with some differences: both sessions were done in two different sound booths, and the first session made use of a Mac Pro while the second session was recorded on a Mac Book Pro. These differences are not thought to negatively affect our VOT measurements.

The recordings were then analyzed using Praat, a recording software used to analyze speech. The word [tɛ] was chosen as a reference word in order to calculate the effect of speech rate on segment duration. The closure and aspiration durations for the target words were divided by the duration of the reference word. As such, if the speech rate for a sentence was higher, both the target word and the reference word would be shorter, but the ratio would stay the same as for a sentence spoken more slowly.

The reference word duration (from release of the closure to the start of the following vowel) was segmented using Praat. The closure and aspiration part of the target sounds were also segmented in the same way, as shown in Figure 1:

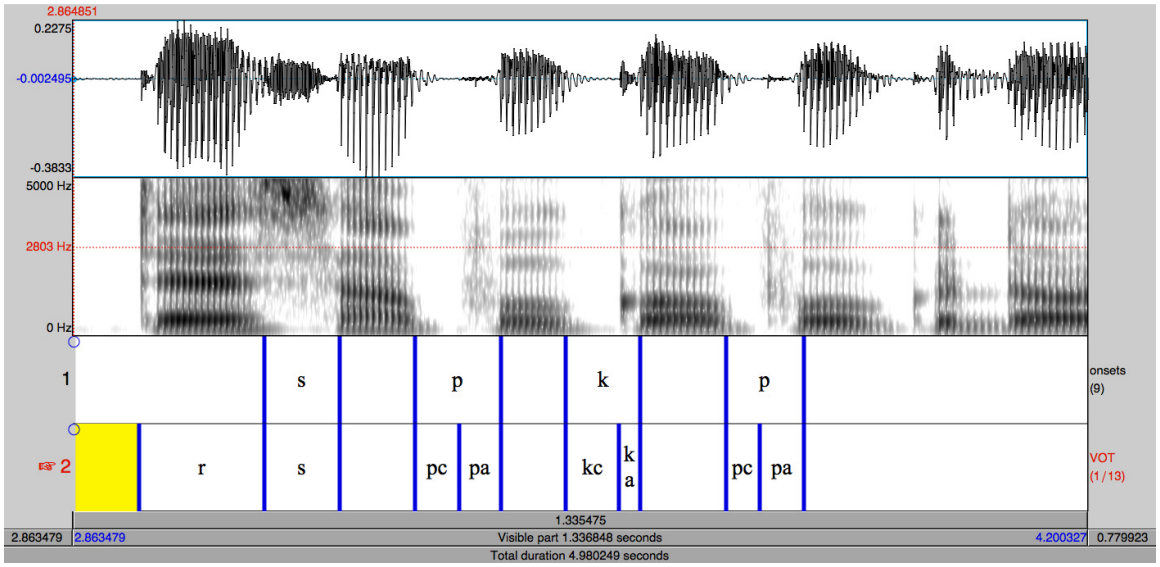


Figure 1 - Segmented sound file in Praat (sap^{h}okop^{h}o, 'animal')

The segments are represented by the blue lines. Reference word was coded (or identified) as 'r', sibilant [s] was coded as 's', the closure was coded as the segment +c (here, pc and kc), and the aspiration was coded as segment +a (here, pa and ka).

After coding was done, we used a Praat script to extract the duration values for the VOT tier. These values were imported into Excel, where the duration ratios were calculated and transformed into milliseconds, and where further specifications for laryngeal settings and place of articulation were input according to the phonological category inferred from orthography. For example, the value 'pc' was further specified as an aspirated bilabial, as per the writing system: q.zduD>zd, where z is [ph] and u is [k]. The writing system was deemed as a good representation of the phonological categorization as it is relatively recent, and so hasn't had time to evolve enough to create discrepancies between the written sound and the intended sound. It was assumed that the written sound is the target sound produced by speakers.

Segments that had no aspiration were further specified as having an aspiration of 0 ms. Segments whose phonetic and phonological values did not match were not included in the statistical analysis, but were presented in more detail below.

No aspiration was segmented or coded for voiced consonants. Voicing "lead" (Abramson 1995) was not possible to measure because of the carrier phrase context: voicing was rarely broken from the preceding word to the beginning of the target word, and so VOT values would always have been for the full closure duration. Since there is no phonologically voiced velar (Abramson 1995: 156), all velars were coded as unvoiced [k] and [k^{h}].

We made no distinction between word-initial and word-medial position for our targeted segments, as we expected that the carrier phrase neutralized any differences (word-initial segments following another word).

After the coding was done and the values checked, we proceeded to do statistical analysis (ANOVAS and T-tests) using SPSS 19.

2.1 Discrepancies between the phonetic and phonological categories

There were some issues while comparing the phonetic category (as we could see on the spectrogram) and the phonological category (as the writing system was telling us). For example, sometimes an unvoiced sound looked voiced on a spectrogram, such as in Figure 2:

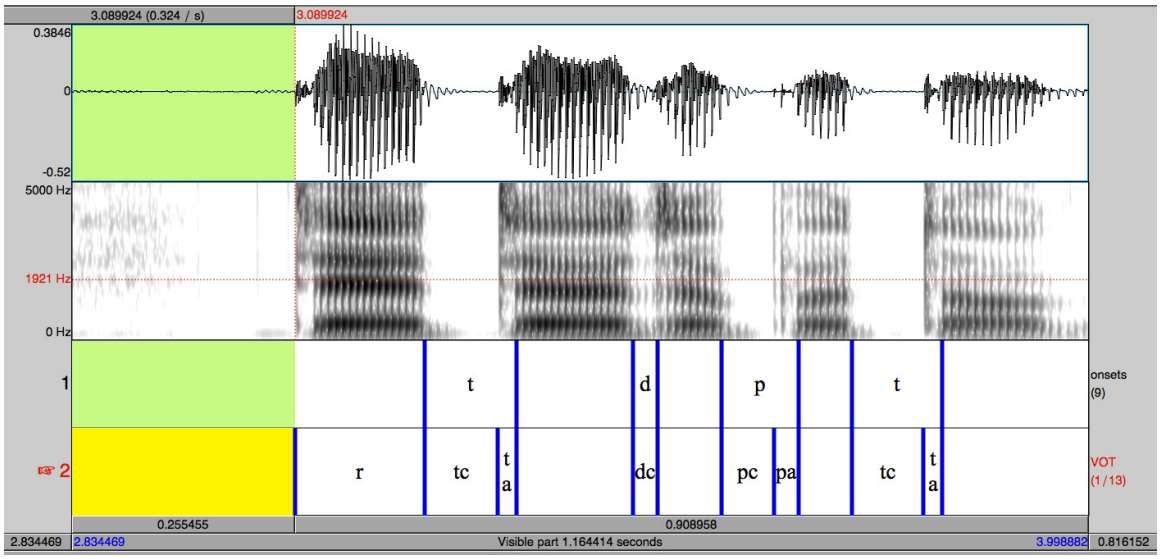


Figure 2 - Phonological segment /t/ was transcribed as [d] (/tɛtəpɔtɔ/ - [tɛdəpɔtɔ], 'first time/level')

The opposite also happened, as when a phonologically voiced segment appeared as unvoiced (Figure 3):

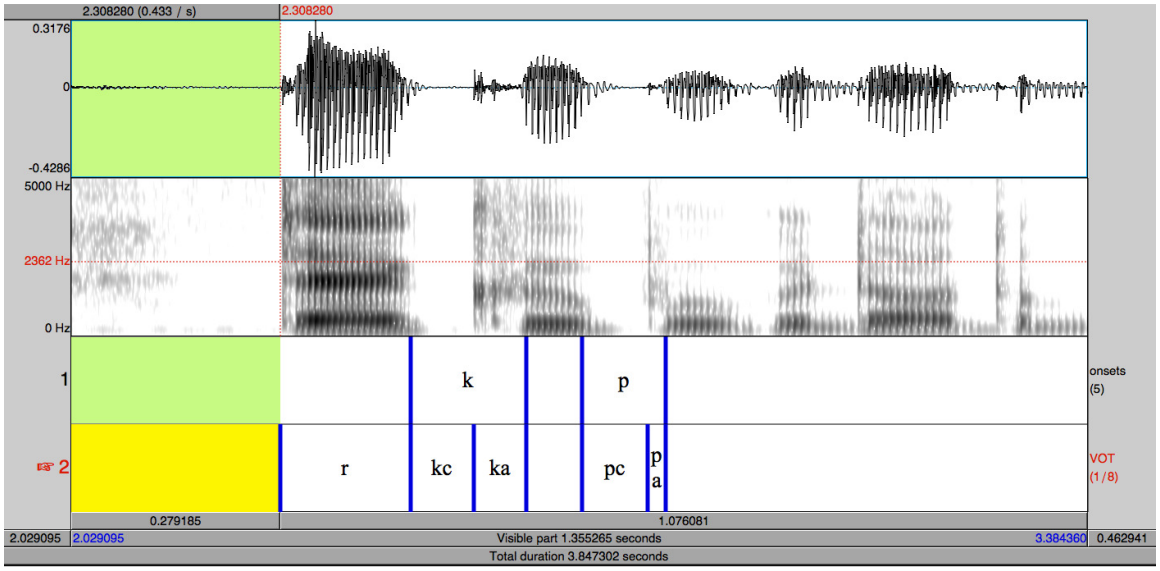


Figure 3 - Phonologically voiced /b/ is transcribed as [p] (/k^{h}ubu/ - [k^{h}upu], 'rice field')

We therefore discarded the problematic segments (but kept other non-problematic segments from the word, if applicable).

Other segments had to be discarded due to errors in the production (such as stuttering) or in the recordings (such as noise from paper during a segment). We also discarded any word where we had no clear basis for orthography (where we could not find the exact word in the dictionary).

3.0 Results and discussion

3.1 Effect of Speech Rate

We evaluated the difference in the ANOVAs between the normalized (in ratio) value for duration and the non-normalized (in ms) value for duration.

We compared significance levels for duration in milliseconds and duration in ratio to place of articulation and laryngeal setting, to see whether or not one measurement type had a different effect. We obtained the following table for place of articulation:

Table 1 - Linearly independent pairwise comparisons of place articulation

Descriptives				
		N	**Mean**	**Std. Deviation**
Duration (ms)	bilabial	248	57.11	32.500
	alveo-dental	474	45.85	32.300
	velar	304	54.06	24.290
	Total	1026	51.00	30.569
Ratio of Duration	bilabial	248	.3817	.21559
	alveo-dental	474	.2990	.20423
	velar	304	.3477	.15975
	Total	1026	.3334	.19790

Dependent Variable		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Duration (ms)	**Contrast**	2026.404	2	1013.202	3.557	.029
	Error	287140.707	1008	284.862		
Ratio of Duration	**Contrast**	.015	2	.008	.670	.512
	Error	11.438	1008	.011		

Both duration in milliseconds and duration in ratios give a non-significant difference. We obtain similar results (between milliseconds and ratio) when testing for laryngeal setting:

Table 2 - Linearly independent pairwise comparisons of laryngeal setting

Descriptives				
		N	**Mean**	**Std. Deviation**
Duration (ms)	**voiced**	148	37.48	40.285
	plain	574	47.38	30.931
	aspirated	304	64.44	16.404
	Total	1026	51.00	30.569
Ratio of Duration	**voiced**	148	.2367	.25312
	plain	574	.3056	.19583
	aspirated	304	.4329	.11372
	Total	1026	.3334	.19790

Dependent Variable		**Sum of Squares**	**df**	**Mean Square**	**F**	**Sig.**
Duration (ms)	**Contrast**	27273.950	2	13636.975	47.872	.000
	Error	287140.707	1008	284.862		
Ratio of Duration	**Contrast**	1.811	2	.905	79.792	.000
	Error	11.438	1008	.011		

Since in both cases, the significance results are similar (be they non-significant or significant), we decided to use the more intuitive durations in milliseconds for our statistical analysis.

3.2 Effect of transcriber and recording session

We found no statistically significant difference between recording sessions and between transcribers.

3.3 Descriptive Statistics

The number of tokens of each place of articulation (bilabial, alveo-dental, velar), laryngeal setting (voiced, plain, aspirated) and measurements (aspiration, closure) made is shown in Table 3:

Table 3 - Number of tokens

	Code	Value Label	N
Place of Articulation	1	bilabial	248
	2	alveo-dental	474
	3	velar	304
Laryngeal Setting	1	voiced	148
	2	plain	574
	3	aspirated	304
Kind of Measurement	1	aspiration	513
	2	closure	513

The values for the mean, standard deviation and number of tokens of each category, according to duration (ms) and to duration ratio, is detailed in Table 6 in the Appendix.

3.4 Mean Values

These results, for mean values, are compared below to the results for intervocalic consonants of Pwo Karen and Sgaw Karen from Abramson (table layout and data from Abramson 1995: 158-159):

Table 4 - Comparison of our results with Abramson (1995)

	p			**ph**		
	Pwo	**Sgaw**	**Our study**	**Pwo**	**Sgaw**	**Our study**
M	11	16	30.68	48	80	56.54
SD	5	6	18.58	14	13	16.98
N	3	3	51	9	9	53
	t			**th**		
	Pwo	**Sgaw**	**Our study**	**Pwo**	**Sgaw**	**Our study**
M	13	13	17.55	65	98	66.00
SD	2	6	6.33	6	-	18.18
N	3	5	136	3	1	47
	k			**kh**		
	Pwo	**Sgaw**	**Our study**	**Pwo**	**Sgaw**	**Our study**
M	21	14	26.56	128	109	65.51
SD	10	8	14.16	45	18	13.10
N	3	5	100	3	16	52

As we can see, our values for plain alveo-dentals are quite similar to both Pwo and Sgaw Karen from Abramson 1995. However, no other stops are consistent through results. Mean duration of aspiration for plain bilabials and velars is longer in our study than in Abramson's

(1995) for both Pwo and Sgaw Karen. Aspirated consonants have quite different results: aspirated bilabials are longer than Abramson's Pwo Karen results, but shorter than his Sgaw Karen results. Our results are very similar to his Pwo Karen results for aspirated alveo-dentals, but much shorter than his Sgaw Karen results. Aspirated velars are much shorter in our study than both dialects of Karen from Abramson (1995). To summarize, our results are very close to Pwo Karen for [t], [k] and [th], and very close to Sgaw Karen for [t]. There is therefore a difference (of which a degree of significance could not be calculated without all of Abramson's data) between our data and Abramson's. This difference is most likely due to the difference in total tokens, although the effect of individual speakers might also interact.

Lisker and Abramson (1968: 403) have found that two-way languages have three categories of aspiration: -125 to -75 ms for voiced consonants, and two categories of 0-25 ms and 60-100 ms for unvoiced consonants. Three-way contrast languages were described as having "ranges centering at -100, +10 and +75 msec." (Lisker and Abramson 1968: 403). For general consonants in Sgaw Karen (Table 6, Appendix), we found that voiced consonants had an aspiration value of 0, or negative VOT (since we were unable to calculate VOT due to the lack of sentence-initial segments), plain consonants had a mean duration of aspiration of 23.02 ms, and aspirated consonants had a mean duration of aspiration of 62.53 ms. It therefore seems that these values fit perfectly into the two voiceless categories of two-way contrast languages, but that they are a bit higher (plain) and lower (aspirated) than for three-way contrast languages.

In a comparison (imperfect because of comparisons between means and averages and between ranges and standard deviations) with Thai, a neighbouring language of Karen, we obtain the following results:

Table 5 - Comparison of our results with Lisker and Abramson (1968)

	p - plain		**p - aspirated**	
	Thai	Our study	Thai	Our study
Mean	6	85.46	64	64.51
Range	0:20	42:232	25100	32:141
SD	N.A.	27.21	N.A.	17.68
N	32	51	33	53
	t - plain		**t - aspirated**	
	Thai	Our study	Thai	Our study
Mean	9	68.32	65	68.75
Range	0:25	29:198	25-125	33:122
SD	N.A.	23.84	N.A.	18.17
N	33	139	33	44
	k - plain		**k - aspirated**	
	Thai	Our study	Thai	Our study
Mean	25	68.80	100	67.13
Range	0:40	31:128	50:155	42:90
SD	N.A.	17.50	N.A.	11.59
N	32	100	38	51

It therefore seems that for plain bilabials and alveo-dentals, our data shows that aspiration is longer than in Thai. However, results for plain velars are quite similar between the Thai average and our Sgaw Karen mean. In aspirated stops, aspirated bilabials and velars seem shorter than in Thai, while alveo-dental stops are very similar. However, these comparisons are very difficult to interpret, since averages and means can have different meanings.

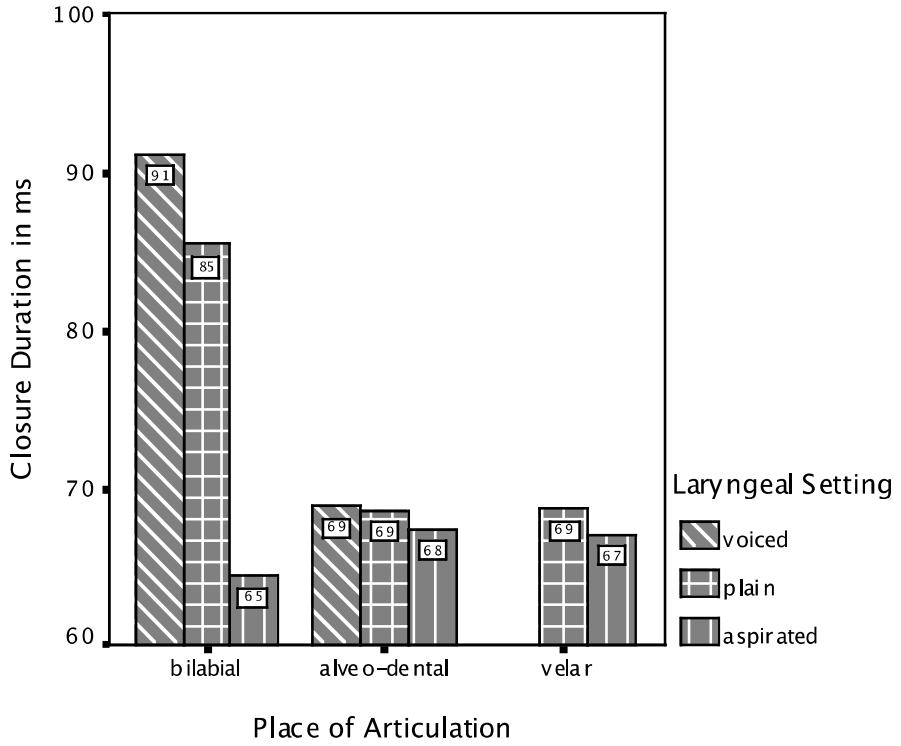
3.5 Closure duration by place of articulation

Our two-way ANOVA test revealed that there was a significant ($p<0.001$) effect for place of articulation. The Tukey HSD post-hoc test showed that, in general (combining plain, aspirated and voiced laryngeal settings), there was a significant difference ($p<0.001$) between bilabials and alveo-dentals, and between velars and alveo-dentals ($p<0.001$).

However, when looking at the Tukey HSD results for individual places of articulation, we found the following results: closure of voiced bilabials differed from that of all other places of articulations and laryngeal settings, including aspirated and plain bilabials ($p<0.001$, see Table 7), except for the closure duration of plain bilabials ($p=0.996$), which is as expected, since both sounds have no or very little aspiration. However, the closure duration for voiced alveo-dentals is significantly different from that of voiced bilabials and of plain bilabials ($p<0.001$) only; comparison with other laryngeal settings and places of articulation results in high non-significance (Table 7). Aspirated bilabials differ significantly only from other voiced and plain bilabials ($p<0.001$). Aspirated alveo-dentals and aspirated velars also differ significantly only with voiced bilabials and plain bilabials ($p<0.001$). The detailed results of every combination of factors from this post-hoc tests are shown in detail in Table 7.

The results for mean closure duration are summarized in the figure below (there is no voicing contrast for velars):

Figure 4 - Mean closure duration by place of articulation

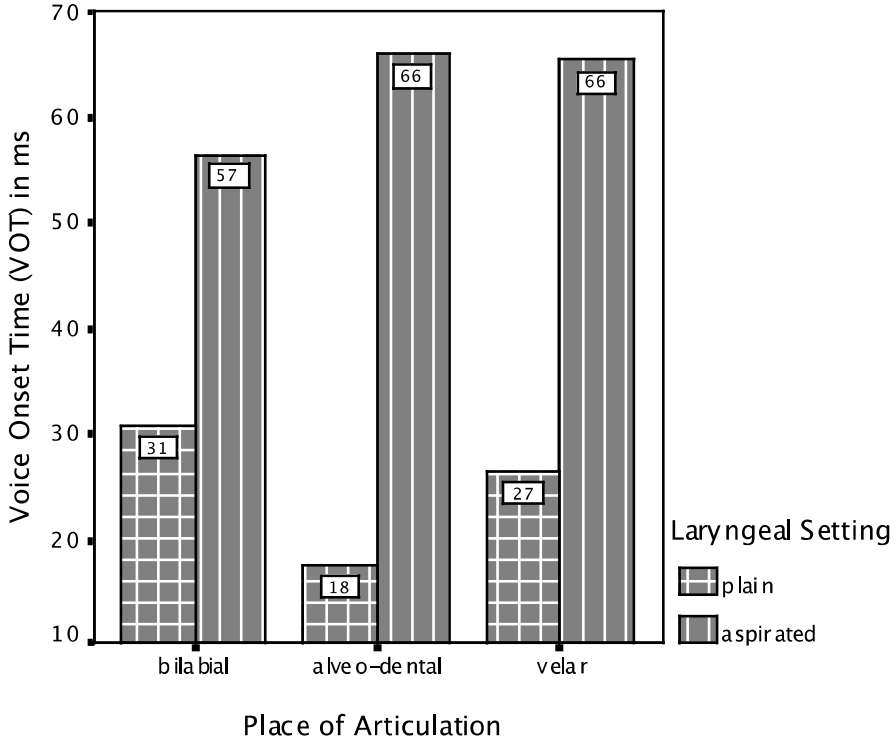


As seen in Figure 4, voiced and plain bilabials are greatly different from aspirated bilabials. This shows there is a relationship between closure duration and aspiration for bilabials. Aspirated bilabials may have a shorter closure duration because air builds up more quickly in the mouth cavity and is thus released more quickly.

3.6 Aspiration duration by place of articulation

Aspiration duration differs greatly between aspirated and plain consonants, according to place of articulation, as is shown in Figure 5:

Figure 5 - VOT by place of articulation (voiced consonants all have 0 ms of aspiration)



As shown in Figure 5(based on our two-way ANOVA test), the mean durations values of aspirated bilabials are not significantly different from aspirated alveo-dentals and aspirated velars. However, they are significantly different from all plain and voiced consonants ($p<0.001$). In the same way, the mean durations of aspirated dentals differ significantly from the plain and voiced stops, but not from other aspirated stops (Table 8). Aspirated velars also differ significantly from all plain and voiced consonants, but not from the other aspirated consonants.

Plain bilabials differ from all other places of articulation and laryngeal settings ($p<0.001$ except for plain alveo-dentals, $p<0.005$), except for plain velars. There is therefore no significant difference between the mean duration of aspiration for plain velars and plain bilabials ($p=0.990$). Plain velars also do not differ significantly from plain alveo-dentals ($p=0.038$), although they do differ from all other consonants ($p<0.001$, except for plain bilabials, $p<0.005$). There is no significant difference between voiced consonants, as they all have an aspiration value of 0 ms. They nevertheless significantly differ from their aspirated and plain counterparts, except for the plain and voiced bilabials, which are not significantly different ($p=0.196$).

4.0 Conclusion

This study has shown that the aspiration values for consonants are not, in general, statistically different from one another depending on place of articulation. However, there is a significant difference when we look at laryngeal setting and the duration of aspiration. We have shown the three-way contrast of Sgaw Karen plosives by controlling for context by placing the target words in a carrier phrase, controlling for speech rate by comparing durations in milliseconds with duration ratios, and by comparing closure and aspiration values for different places of articulation and different laryngeal settings.

As shown in this study, while there is no clear-cut difference for duration of closure, there is a statistically significant difference between plain, voiced and aspirated consonants for duration of aspiration. This study corroborates the three-way plosive contrast of Sgaw Karen.

5.0 Future Research

While this study has shown that the three-way contrast in plosives, as described in the literature, indeed exists in Karen, there are still a few unanswered questions. As described in Abramson (1995), another aspiration contrast exists in Karen: that of alveo-dental fricatives ([s] and [s^h]). Aspirated [s^h] is more difficult to identify in a spectrogram as the change in the noise pattern between the fricative and aspiration is very subtle, but it remains that the phonetic description of these sounds would be revealing, especially since, as Abramson (1995: 163) points out, an aspiration contrast in sibilants is very rare.

Abramson compared VOT values for plosives in isolated words, and had a passing reference to running speech. Our study deals with those same words in a carrier sentence, but we have not compared the words in the carrier sentences with the isolated words.

This study was also based on only one participant. While the presence of the contrast is unlikely to be unique to this participant, the closure and aspiration durations could be an effect of this particular speaker. Further research needs to be done with more participants to average out the closure and aspiration durations and to neutralize the (potential) speaker preference effect.

References

Abramson, A. (1995). Laryngeal Timing in Karen Obstruents. In F. Bell-Berti & L.J. Raphael (Eds.), *Producing Speech: Contemporary Issues* (155-165). New York: AIP Press.

Audacity Team. (2010). *Audacity 1.2.5* [computer program]. Retrieved from: http://audacity.sourceforge.net/

Boersma, P. & Weenink, D. (2010). *Praat: doing phonetics by computer* (Version 5.1.25) [computer program]. Retrieved from: http://www.praat.org/

Gilmore, D. (1898). *A Grammar of the Sgaw Karen.* Rangoon: American Baptist Mission Press.

IBM Corporation. (1995-2010). *IBM SPSS Statistics 19* [computer program]. New York: IBM Corporation.

Kato, A. (1995). The phonological system of Pwo Karen dialects. *Linguistics of the Tibeto-Burman Area* 18(1):63-103.

Lisker, L. & Abramson, A. (1964). A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements. *Word* 20 (3): 384-422.

Phillips, A. (2009). Omkoi Pwo Karen phonology and orthography. *Payap Working Papers in Linguistics* 5. [http://ic.payap.ac.th/graduate/linguistics/papers/Omkoi_Pwo_Karen_Phonology_and_Orthography.pdf]

Appendix A - Table of values

Table 6 - Descriptive statistics

	Place of Articulation	Laryngeal Setting	Kind of Measurement	Mean	Std. Deviation	N
Duration (ms)	bilabial	voiced	aspiration	.00	.000	20
			closure	91.22	18.891	20
			Total	45.61	48.036	40
		plain	aspiration	30.68	18.581	51
			closure	85.46	27.411	51
			Total	58.07	36.063	102
		aspirated	aspiration	56.54	16.982	53
			closure	64.51	17.679	53
			Total	60.53	17.710	106
		Total	aspiration	36.78	25.804	124
			closure	77.44	24.945	124
			Total	57.11	32.500	248
	alveo-dental	voiced	aspiration	.00	.000	54
			closure	68.93	17.725	54
			Total	34.47	36.806	108
		plain	aspiration	17.55	6.328	136
			closure	68.73	23.933	136
			Total	43.14	31.027	272
		aspirated	aspiration	66.00	18.184	47
			closure	67.52	18.225	47
			Total	66.76	18.123	94
		Total	aspiration	23.16	24.366	237
			closure	68.54	21.523	237
			Total	45.85	32.300	474
	velar	plain	aspiration	26.56	14.163	100
			closure	68.80	17.495	100
			Total	47.68	26.466	200
		aspirated	aspiration	65.51	13.098	52
			closure	67.18	11.483	52
			Total	66.34	12.285	104
		Total	aspiration	39.88	23.092	152
			closure	68.24	15.678	152

			Total	54.06	24.290	304
	Total	voiced	aspiration	.00	.000	74
			closure	74.96	20.501	74
			Total	37.48	40.285	148
		plain	aspiration	23.02	13.333	287
			closure	71.73	23.422	287
			Total	47.38	30.931	574
		aspirated	aspiration	62.53	16.655	152
			closure	66.35	15.975	152
			Total	64.44	16.404	304
		Total	aspiration	31.41	25.509	513
			closure	70.60	21.223	513
			Total	51.00	30.569	1026
Ratio of Duration	bilabial	voiced	aspiration	.0000	.00000	20
			closure	.5735	.09252	20
			Total	.2867	.29748	40
		plain	aspiration	.2072	.13453	51
			closure	.5677	.18017	51
			Total	.3874	.24050	102
		aspirated	aspiration	.3833	.11481	53
			closure	.4409	.13532	53
			Total	.4121	.12819	106
		Total	aspiration	.2490	.17749	124
			closure	.5144	.16228	124
			Total	.3817	.21559	248
	alveo-dental	voiced	aspiration	.0000	.00000	54
			closure	.4364	.11389	54
			Total	.2182	.23342	108
		plain	aspiration	.1157	.04467	136
			closure	.4452	.12583	136
			Total	.2805	.19008	272
		aspirated	aspiration	.4414	.12745	47
			closure	.4492	.10917	47
			Total	.4453	.11809	94
		Total	aspiration	.1539	.16441	237
			closure	.4440	.11963	237
			Total	.2990	.20423	474
	velar	plain	aspiration	.1658	.08649	100

			closure	.4305	.11152	100
			Total	.2982	.16587	200
		aspirated	aspiration	.4368	.09332	52
			closure	.4489	.08606	52
			Total	.4428	.08953	104
		Total	aspiration	.2585	.15646	152
			closure	.4368	.10359	152
			Total	.3477	.15975	304
	Total	voiced	aspiration	.0000	.00000	74
			closure	.4735	.12409	74
			Total	.2367	.25312	148
		plain	aspiration	.1494	.08902	287
			closure	.4619	.14116	287
			Total	.3056	.19583	574
		aspirated	aspiration	.4195	.11463	152
			closure	.4462	.11160	152
			Total	.4329	.11372	304
		Total	aspiration	.2079	.17253	513
			closure	.4589	.13070	513
			Total	.3334	.19790	1026

Table 7 – Closure Duration

Closure Duration (ms) vs. Place of articulation and Laryngeal Settings						
		Mean Difference (I-J)			95% Confidence Interval	
Factor 1	Factor 2		Std. Error	Sig.	Lower Bound	Upper Bound
voiced bilabial	voiced velar	22.29*	4.418	.000	7.11	37.46
	aspirated bilabial	26.71*	4.429	.000	11.49	41.92
	aspirated alveo-dental	22.47*	4.552	.000	6.84	38.10
	aspirated velar	24.09*	4.453	.000	8.79	39.38
	plain bilabial	5.76	4.453	.996	-9.54	21.05
	plain alveo-dental	22.90*	4.036	.000	9.04	36.77
	plain velar	22.42*	4.134	.000	8.22	36.62

voiced alveo-dental	voiced bilabial	-22.29*	4.418	.000	-37.46	-7.11
	aspirated bilabial	4.42	3.263	.994	-6.79	15.63
	aspirated alveo-dental	.18	3.428	1.000	-11.59	11.96
	aspirated velar	1.80	3.296	1.000	-9.52	13.12
	plain bilabial	-16.53*	3.296	.000	-27.85	-5.21
	plain alveo-dental	.62	2.706	1.000	-8.68	9.91
	plain velar	.14	2.850	1.000	-9.65	9.93
aspirated bilabial	voiced bilabial	-26.71*	4.429	.000	-41.92	-11.49
	voiced velar	-4.42	3.263	.994	-15.63	6.79
	aspirated alveo-dental	-4.24	3.442	.998	-16.06	7.59
	aspirated velar	-2.62	3.311	1.000	-13.99	8.75
	plain bilabial	-20.95*	3.311	.000	-32.32	-9.58
	plain alveo-dental	-3.81	2.725	.991	-13.16	5.55
	plain velar	-4.29	2.868	.983	-14.14	5.56
aspirated alveo-dental	voiced bilabial	-22.47*	4.552	.000	-38.10	-6.84
	voiced velar	-.18	3.428	1.000	-11.96	11.59
	aspirated bilabial	4.24	3.442	.998	-7.59	16.06
	aspirated velar	1.62	3.473	1.000	-10.31	13.55
	plain bilabial	-16.71*	3.473	.000	-28.64	-4.78
	plain alveo-dental	.43	2.920	1.000	-9.60	10.46
	plain velar	-.05	3.053	1.000	-10.54	10.44
aspirated velar	voiced bilabial	-24.09*	4.453	.000	-39.38	-8.79
	voiced velar	-1.80	3.296	1.000	-13.12	9.52
	aspirated bilabial	2.62	3.311	1.000	-8.75	13.99
	aspirated alveo-dental	-1.62	3.473	1.000	-13.55	10.31
	plain bilabial	-18.33*	3.342	.000	-29.81	-6.85
	plain alveo-dental	-1.19	2.763	1.000	-10.68	8.31
	plain velar	-1.67	2.904	1.000	-11.64	8.31
plain bilabial	voiced bilabial	-5.76	4.453	.996	-21.05	9.54

	voiced velar	16.53*	3.296	.000	5.21	27.85
	aspirated bilabial	20.95*	3.311	.000	9.58	32.32
	aspirated alveo-dental	16.71*	3.473	.000	4.78	28.64
	aspirated velar	18.33*	3.342	.000	6.85	29.81
	plain alveo-dental	17.14*	2.763	.000	7.65	26.64
	plain velar	16.66*	2.904	.000	6.69	26.64
plain alveo-dental	voiced bilabial	-22.90*	4.036	.000	-36.77	-9.04
	voiced velar	-.62	2.706	1.000	-9.91	8.68
	aspirated bilabial	3.81	2.725	.991	-5.55	13.16
	aspirated alveo-dental	-.43	2.920	1.000	-10.46	9.60
	aspirated velar	1.19	2.763	1.000	-8.31	10.68
	plain bilabial	-17.14*	2.763	.000	-26.64	-7.65
	plain velar	-.48	2.213	1.000	-8.08	7.12
plain velar	voiced bilabial	-22.42*	4.134	.000	-36.62	-8.22
	voiced velar	-.14	2.850	1.000	-9.93	9.65
	aspirated bilabial	4.29	2.868	.983	-5.56	14.14
	aspirated alveo-dental	.05	3.053	1.000	-10.44	10.54
	aspirated velar	1.67	2.904	1.000	-8.31	11.64
	plain bilabial	-16.66*	2.904	.000	-26.64	-6.69
	plain alveo-dental	.48	2.213	1.000	-7.12	8.08

Table 8 – Voice Onset Time (VOT)

VOT in milliseconds vs. Place of articulation and Laryngeal Settings						
		Mean Difference			95% Confidence Interval	
Factor 1	Factor 2	(I-J)	Std. Error	Sig.	Lower Bound	Upper Bound
aspirated bilabial	aspirated alveo-dental	-8.93	3.442	.404	-20.75	2.89
	aspirated velar	-9.04	3.279	.295	-20.31	2.22
	plain bilabial	25.86*	3.311	.000	14.48	37.23

	plain alveo-dental	37.78*	2.725	.000	28.42	47.14
	plain velar	29.98*	2.868	.000	20.13	39.83
	voiced bilabial	56.54*	4.429	.000	41.32	71.75
	voiced alveo-dental	56.54*	3.263	.000	45.33	67.75
aspirated alveo-dental	aspirated bilabial	8.93	3.442	.404	-2.89	20.75
	aspirated velar	-.11	3.442	1.000	-11.94	11.71
	plain bilabial	34.79*	3.473	.000	22.86	46.71
	plain alveo-dental	46.71*	2.920	.000	36.68	56.73
	plain velar	38.91*	3.053	.000	28.43	49.40
	voiced bilabial	65.47*	4.552	.000	49.83	81.10
	voiced alveo-dental	65.47*	3.428	.000	53.69	77.24
aspirated velar	aspirated bilabial	9.04	3.279	.295	-2.22	20.31
	aspirated alveo-dental	.11	3.442	1.000	-11.71	11.94
	plain bilabial	34.90*	3.311	.000	23.53	46.27
	plain alveo-dental	46.82*	2.725	.000	37.46	56.18
	plain velar	39.03*	2.868	.000	29.18	48.88
	voiced bilabial	65.58*	4.429	.000	50.37	80.80
	voiced alveo-dental	65.58*	3.263	.000	54.37	76.79
plain bilabial	aspirated bilabial	-25.86*	3.311	.000	-37.23	-14.48
	aspirated alveo-dental	-34.79*	3.473	.000	-46.71	-22.86
	aspirated velar	-34.90*	3.311	.000	-46.27	-23.53
	plain alveo-dental	11.92*	2.763	.002	2.43	21.41
	plain velar	4.13	2.904	.990	-5.85	14.10
	voiced bilabial	30.68*	4.453	.000	15.39	45.98
	voiced alveo-dental	30.68*	3.296	.000	19.36	42.00
plain alveo-dental	aspirated bilabial	-37.78*	2.725	.000	-47.14	-28.42

	aspirated alveo-dental	-46.71*	2.920	.000	-56.73	-36.68
	aspirated velar	-46.82*	2.725	.000	-56.18	-37.46
	plain bilabial	-11.92*	2.763	.002	-21.41	-2.43
	plain velar	-7.79*	2.213	.038	-15.40	-.19
	voiced bilabial	18.76*	4.036	.000	4.90	32.63
	voiced alveo-dental	18.76*	2.706	.000	9.47	28.06
plain velar	aspirated bilabial	-29.98*	2.868	.000	-39.83	-20.13
	aspirated alveo-dental	-38.91*	3.053	.000	-49.40	-28.43
	aspirated velar	-39.03*	2.868	.000	-48.88	-29.18
	plain bilabial	-4.13	2.904	.990	-14.10	5.85
	plain alveo-dental	7.79*	2.213	.038	.19	15.40
	voiced bilabial	26.56*	4.134	.000	12.35	40.76
	voiced alveo-dental	26.56*	2.850	.000	16.77	36.35
voiced bilabial	aspirated bilabial	-56.54*	4.429	.000	-71.75	-41.32
	aspirated alveo-dental	-65.47*	4.552	.000	-81.10	-49.83
	aspirated velar	-65.58*	4.429	.000	-80.80	-50.37
	plain bilabial	-30.68*	4.453	.000	-45.98	-15.39
	plain alveo-dental	-18.76*	4.036	.000	-32.63	-4.90
	plain velar	-26.56*	4.134	.000	-40.76	-12.35
	voiced alveo-dental	.00	4.418	1.000	-15.17	15.17
voiced alveo-dental	aspirated bilabial	-56.54*	3.263	.000	-67.75	-45.33
	aspirated alveo-dental	-65.47*	3.428	.000	-77.24	-53.69
	aspirated velar	-65.58*	3.263	.000	-76.79	-54.37
	plain bilabial	-30.68*	3.296	.000	-42.00	-19.36
	plain alveo-dental	-18.76*	2.706	.000	-28.06	-9.47
	plain velar	-26.56*	2.850	.000	-36.35	-16.77
	voiced bilabial	.00	4.418	1.000	-15.17	15.17

Tone and Phonation Types in Sgaw Karen

Joshua Finkeldey

Sgaw Karen possesses a writing system that represents 6 different tones, including one that is not explicitly marked. Preliminary acoustic analysis of recorded Karen speech indicates that each of the 6 written tones is acoustically distinct in terms of pitch, voice quality, and duration, though there may be enough similarity between some of them to cause perceptual confusion

1. Introduction

Typical of Southeast Asian languages, the Tibeto-Burman language Sgaw Karen has a rich inventory of tones, which are represented orthographically in the Burmese-derived script with six distinct tone markers. Weidert has described Sgaw Karen as possessing a register tone system, making use of both pitch and phonation mode distinctions, while noting that Karen languages spoken in contact with Tai languages may also have contour characteristics (1987: 317). This seems to be the case with the particular variety of Sgaw Karen spoken by our male informant, who has pointed out that "dialects" of Sgaw Karen are not necessarily mutually intelligible. In his variety, the 5 written tones, plus one unmarked tone, do correspond to six distinct acoustic counterparts, although their perceptual distinctiveness is not yet known. With the speaker exhibiting a fairly compressed range for fundamental frequency, but with significant differences in vowel length and voicing quality, it is not clear to what degree particular cues are important. Regardless of their perceptual weight, each tone type does have some characteristic *potential* cues that are amenable to measurement. For this reason, the description that follows errs on the side of giving potentially redundant information; a more formalized and categorical set of judgments would be more attractive but probably fail at this point to reflect reality.

2. Phonological Considerations

Sgaw Karen has a fairly rich inventory of basic onset consonants and also allows limited clustering with "medials". It does not allow syllable codas, with the exceptions of a few borrowed words and glottal stops, although these stops are phonologically linked to particular tones. As Andruski & Ratliff note is the case for Vietnamese and Green Mong, Sgaw Karen's voice quality contrasts are properties of particular tones, not consonants as in some other languages (2000: 38). This means that tonal variation in voice quality operates largely independently of the voicing of adjacent onsets, though it is not entirely unaffected by them. This relationship generally appears to hold for the other potential tone cues examined here, namely duration and F0 contour; varying the onset consonant of a word tends to have a blanket effect on tone types with all of them varying in the same direction, but some more than others.

3. Methodology

The subject of the data collection is a 27 year old native speaker of Sgaw Karen who also speaks Burmese and English. He is literate in Karen and English and assisted in the production of a word list in both languages for recording purposes. This word list consisted of 4 subsets divided by onset consonant type, each subset containing six sentences further targeting a tone type (represented in transcriptions by superscript numbers 1 through 6, corresponding partially in ordering to Karen writing conventions). The basic carrier sentence never varied except in its target CV(+tone) syllable, following this form:

jə p^ha^4 $li^3mɛ^2p^hlɤ^5$ _____ $lɤ^1$ $mɛ^2k^hro^5fo^5pu^6$
I read word _____ PREP microphone
I read the word _____ into the microphone.

Onset consonants for the target words in the word list were the dental/alveolar nasal and oral stops / n, d, t, t^h /, and the vowel was always /a/. Using all the possible combinations of onset, vowel, and tone produces a number of nonexistent Karen words in addition to real lexical items, but the speaker was comfortable producing both types.

The speaker was recorded in a sound treated booth in the sound patterns laboratory at the University of Ottawa. The recording was made using a Shure SM-10A microphone headset and a Marantz Professional PMD-660 compact flash recorder. Recording was accomplished in one session lasting approximately 45 minutes. The speaker was asked to read the set of 24 syllable types 6 times for a total of 144 uttered sentences. The high irregularity of vocal fold vibration in the vowel offsets of 3 tokens of tone 4 caused large artifacts in the calculation of mean F0. When all fourth tones were plotted individually, the three tokens were clear outliers. These tokens were excluded from the analysis, leaving 141 tokens.

It was requested that the sentences be produced at a comfortable and constant pace throughout the duration of the recording. Breaks were given to the speaker between each reading of 24 sentences, and short pauses were taken after each onset consonant group. The speaker made no noticeable errors during any of the six readings.

4. Acoustic Analysis

All recordings were transferred to a PC, maintaining the original sampling rate of 22,050 Hz, and then sectioned into sentence-length WAV files in preparation for analysis. These files were opened and given TextGrid labels using a script written by Marc Brunelle for the computer program Praat (Boersma & Weenink 2010). After labeling the targeted tone-bearing vowels, another Praat script was used to extract measurements of F0, formants, vowel duration, and voice quality at 10 equidistant points in the vowels.

F0 measurements were taken in order to create pitch contours, while the amplitude difference between the first and second harmonics (H1-H2) was measured to give an indicator of voice quality. Andruski cites Stevens (1998) to explain the usefulness of this measure; values of H1-H2 around +5 dB indicate breathy phonation, values around -5db indicate creakiness, and approximately equal values indicate modal voicing (393). However, higher harmonics in breathy vowels may be as much as 15 dB weaker than their modal counterparts (Stevens 2000: 89). Once the data was fully extracted to text form it was formatted with spreadsheet software and imported to IBM SPSS for restructuring and analysis.

5.0 Results - H1-H2 Differences

Although H1-H2 values are derived by indirect measurement, they have been shown to match well with the more direct method of measuring breathiness/creakiness via glottal airflow (Andruski & Ratliff 2000: 45). Since only the low vowel /a/ was examined in this project, it was not necessary to correct for F1 amplitude differences across vowel types. Figure 1 shows the average H1-H2 value for each tone, including all onset types.

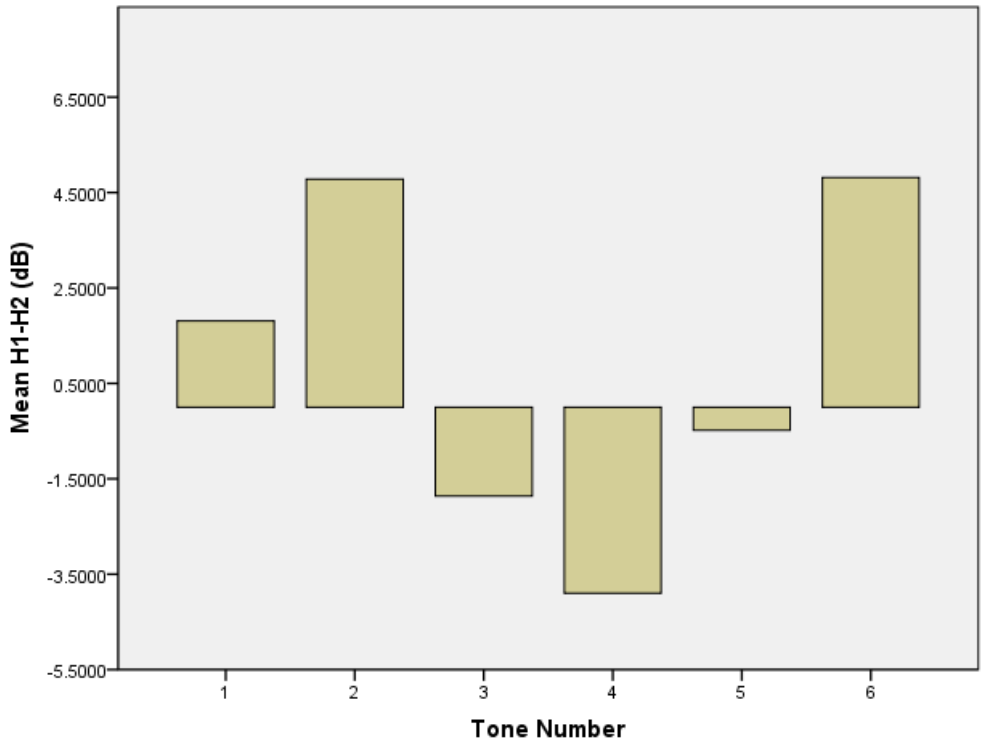


Figure 1. Mean intensity difference between 1st and 2nd harmonic, all tones and onsets.

Table 1. Mean H1-H2 values by tone number

Tone	Mean	N	Std. Deviation
1	1.808771	240	2.6419779
2	4.780949	240	2.0837790
3	-1.862113	240	2.8613482
4	-3.899004	240	3.4960207
5	-.483115	240	4.4139193
6	4.815162	240	2.2694615
Total	.860108	1440	4.4719974

While tones 2 and 6 have noticeably higher values, and tone 4 has a noticeably lower value, the contrasts are not as large as in previous studies. Andruski & Ratliff (2000) and Huffman (1985) both found average H1-H2 values close to 10 dB for breathy tones, while tones 2 and 6 here average 4.78 dB and 4.82 dB, respectively. However, there is much room for individual speaker variation in addition to average differences across languages; for example, Andruski & Ratliff's male Green Mong speakers had average values for breathy tones ranging over 15 dB (2000: 45). In any case, the difference between our Karen speaker's tones 2 and 6 were not found to be significant. The tone with the next highest average H1-H2 value, tone 1, does differ significantly from tone 2 ($p < 0.001$). It also differs significantly with the next highest, tone 5, which in turn differs from 3, which differs from 4 (all $p < 0.001$).

The relationships between these mean values are generally preserved when the data is sorted by onset consonant type; tones 2 and 6 remain at parity and out-value the others, tone 1 hovers between 0 and 1dB, and the others maintain negative values, with tone 4 always being the lowest. However, there are two primary exceptions. The first is the case of /n/ onsets, which muddy the difference between tones 3, 4, and 5, raising them all slightly. The values for tones 2 and 6 are also increased slightly. See Figure 2.

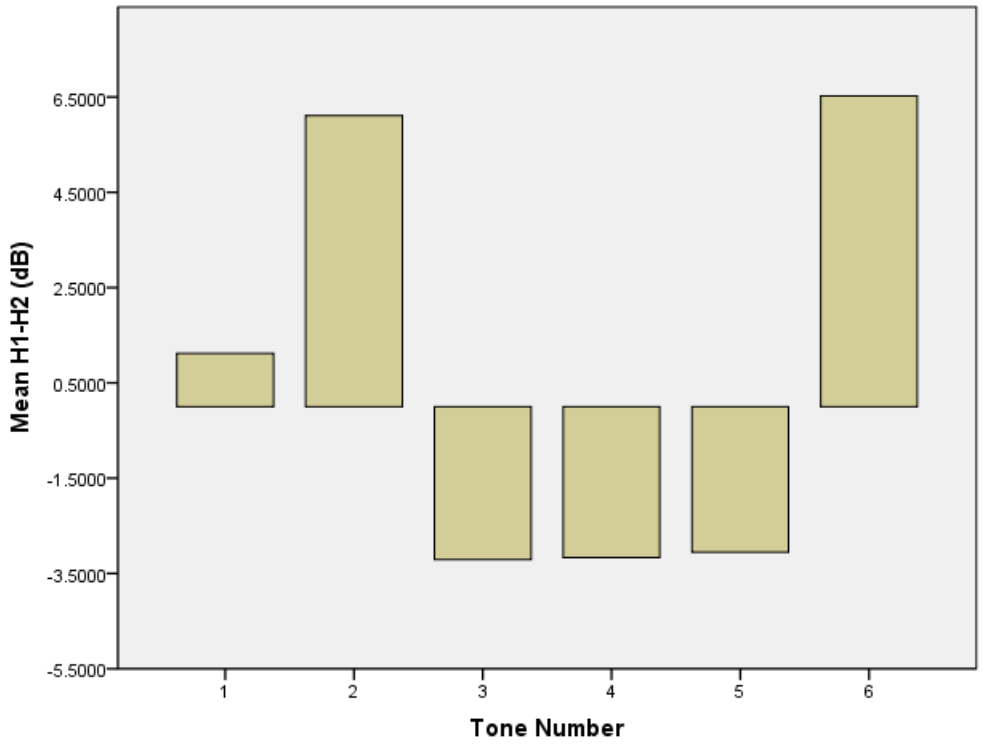


Figure 2. Mean intensity difference between 1st and 2nd harmonic, [n] onset.

The aspirated /t^h/ onset ([th] in charts and tables) is the second exception. This has the effect of biasing most tone types toward positive H1-H2 values, though tones 2 and 6 are not noticeably affected compared to the average. In this case, burst noise or lingering aspiration is suspected. All tones show an initial dip in H1-H2 value when plotted as a function of time (index point), and this effect is particularly pronounced after the aspirated stops. Figure 3 shows mean changes in H1-H2 for all tones and onsets (values up to measurement point 2 are distorted due to the effect of aspiration).

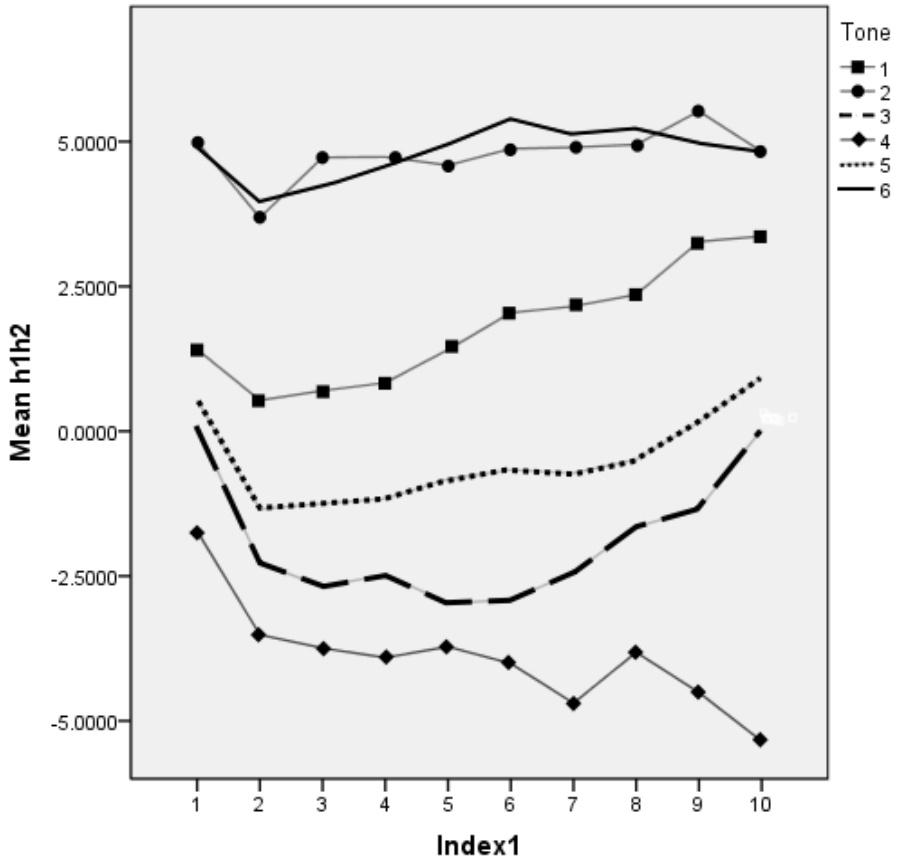


Figure 3. Average H1-H2 values over time

5.1 Results – Duration

Andruski & Ratliff show that Green Mong words with creaky voice are shorter than words with breathy or modal voice, noting that this was also found by Huffman's study (2000: 48). The data from our Karen speaker shows the same pattern. The tone with the lowest average H1-H2 value (most creakiness) is also the tone with the shortest average duration. This is shown in figure 4. A second observation from Andruski & Ratliff is partly repeated. They note that in Green Mong, the creaky tone is basically syllable final, resulting in clearly voiced syllables than end with a "slower glottal stricture" (48). This is considered to contrast with earlier descriptions of Mong syllables as "checked", that is, containing an abrupt final glottal stop. Tones in Sgaw Karen have been described as either open or glottal-stopped (again "checked"), and it is interesting to see that they too have a more gradual increase in phonation irregularity (Weidert 317). Figure 4 shows the mean vowel duration of all tones, and figure 5 shows an example of a wideband spectrogram of tone 4.

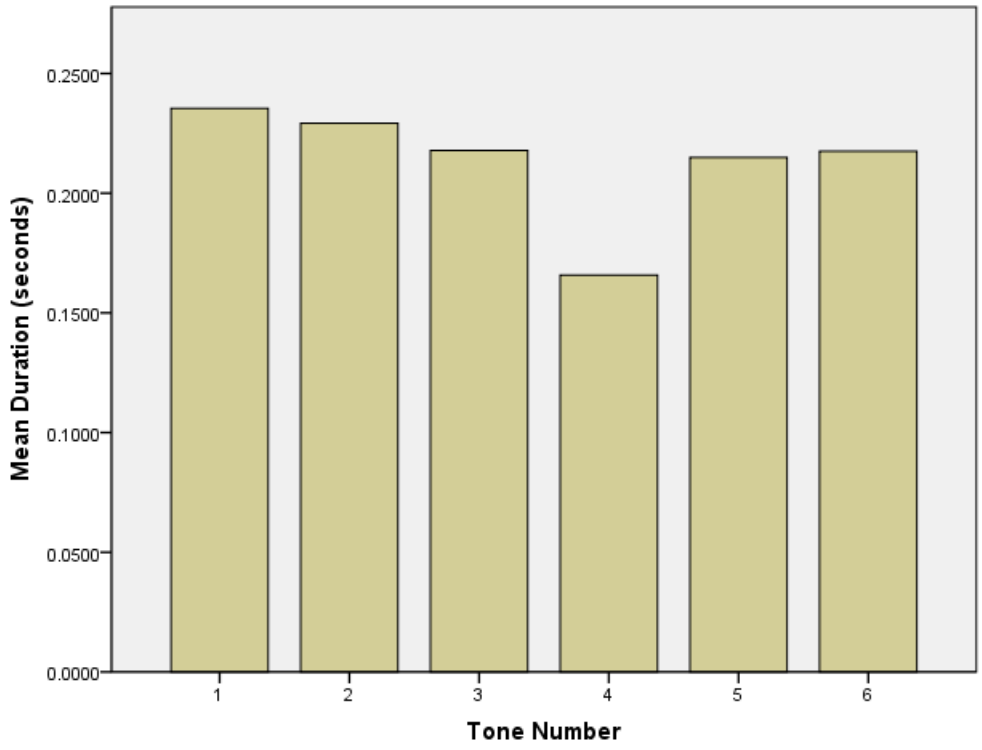
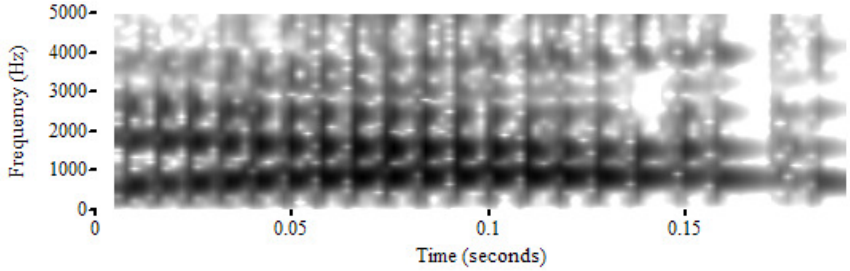


Figure 4. Mean Vowel duration for each tone and all consonant onset types (144 tokens)

Figure 5. Wideband spectrogram of tone 4 showing increase in creak



Differences in mean duration between tones 3, 5, and 6 are not significant. The durations of tone 1 and 2 also only differ randomly. However, comparing tones 1 and 3 with 2 yields significant results ($p < 0.001$, $p < 0.02$). Even if the difference is nonrandom, an average difference of 12 milliseconds does not sound like a promising cue for perceptual discrimination.

Varying the onset consonant alters some of the duration relationships. In particular, syllables beginning with /t^h/ are shorter than the rest, with a mean length of 0.19 seconds, compared to the overall mean of 0.22 seconds. At least part of this is the result of labeling methods; vowel onsets were not marked as beginning until burst noise was largely absent, to avoid artifacts in other parts of the data.

5.2 Results - F0 Contour/Height

Figure 6 shows a plot of mean fundamental frequency for each tone, measured at ten points from onset to offset. Three tokens of tone 4 have been excluded due to errors in calculating F0 during the creaky offset of the vowel.

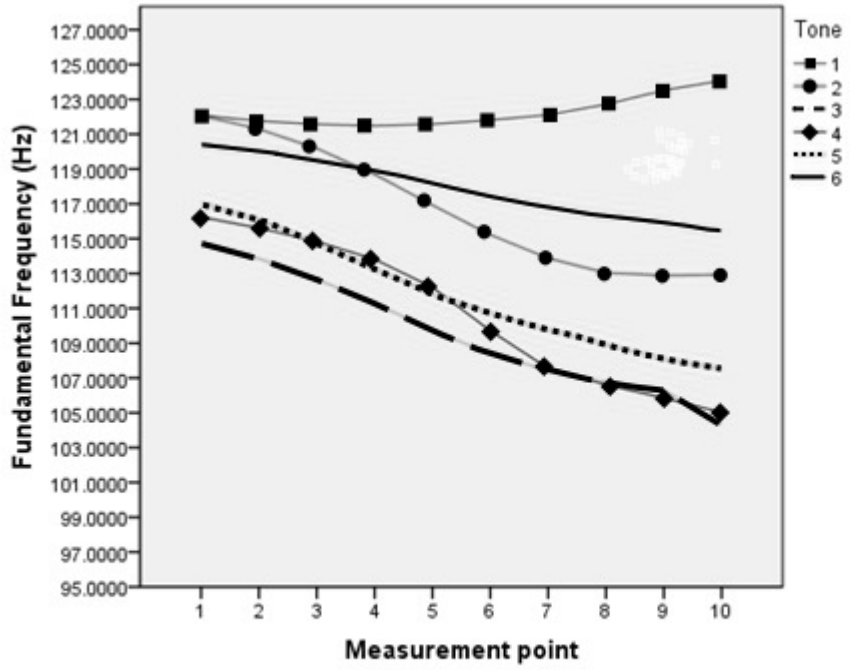


Figure 6. Mean F0 contours for all 7 Sgaw Karen tones for one male speaker.

The overall tone space utilized by the speaker is quite small. Even at the tone offsets, where the greatest difference is found, there is only a range of about 20 Hz between the highest and lowest tones. In addition, about a third of this difference is cause by tone 1, which is the only tone with an average rise in frequency at offset instead of an average fall. Tone 1 has the highest onset pitch, at 122 Hz on average, and it rises to 124 Hz at offset.

Tone 2 also has a high pitch onset, not differing significantly from tone 1, but it drops to 113 Hz at offset. Although the onset and overall pitch of tone 2 does not differ significantly from tone 6, tone 2 does have a significantly lower offset and a more abrupt transition to a lower frequency. Impressionistically, and to a non-speaker of Karen, tone 2 sounds falling in many cases, whereas tone 6 sounds level (and possibly breathier, despite having objectively similar H1-H2 values).

Tones 3, 4, and 5 have mean onsets of approximately 115 Hz, 116 Hz, and 117 Hz, respectively, and are not found to be significantly different. They also have a combined mean offset frequency of 105 Hz, none differing significantly in pairwise comparisons. Upon observing this, it becomes tempting to group these tones into a clear “low” onset category and tones 1, 2, and 6 into a “high” onset category, but unfortunately the two closest tones from each speculative

category, tones 5 and 6, do not differ significantly. If considered on the basis of tone onset minus tone offset, and comparing with tone 2, tones 3, 4, and 5, should all be called “falling tones”. They have average drops in frequency of 10.4 Hz, 11.2 Hz, and 9.5 Hz, compared to the 9.2 Hz range of tone 2. Recalling figure 4, tones 3, 4, and 5 are also shorter on average than tone 2, so they must drop more rapidly to reach the same offset frequency. However, these lower tones do not have more level or rising tones as nearby in the tone space as tone 2 does, so without an immediate contrast, the drop may not be particularly meaningful.

Some things get messier when the F0 data is filtered by onset consonant type. With the onsets /n, t, d/ , the onset of tone 4 rises, even surpassing the onset of tone 1, as in figure 7.

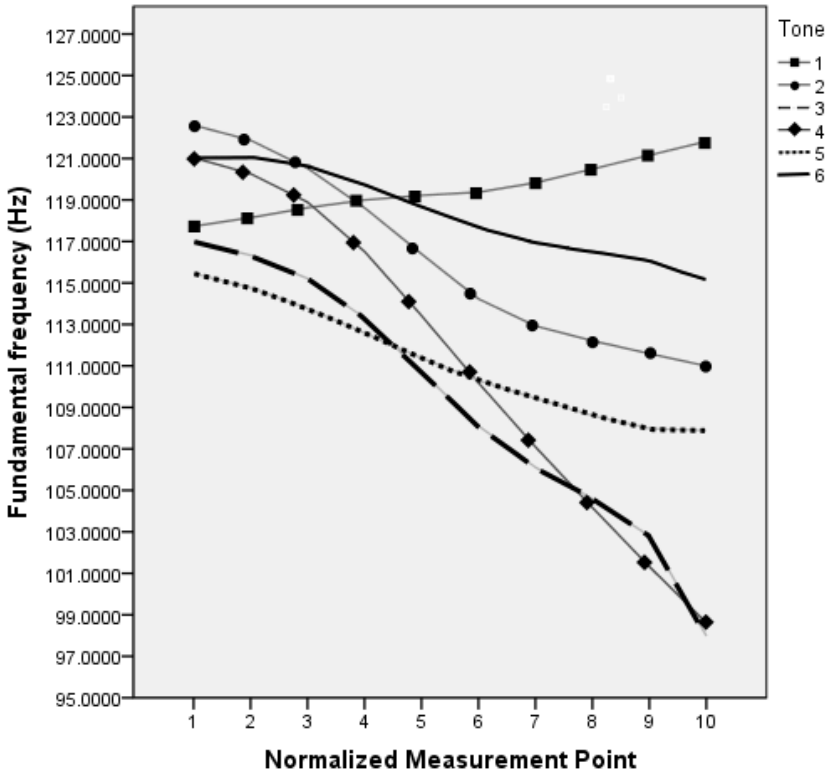


Figure 7. Mean F0 contours for all 6 tones with onset [d]

The gap in offset frequency between tones 5 and 3, seen marginally in the overall average, grows larger to a significant difference. However, the opposite effect is apparent with onset consonants /n/ and /t/.

In no case does the offset frequency ranking of (tone 1 > tone 6 > tone 2) change. For /t^h/ onsets, tones 2 and 6 are at parity, differing insignificantly in offset frequency, but tone 2 retains its more abrupt downward contour.

6. Discussion and Conclusion

Here is a tentative arrangement of the tones by their acoustic qualities. I will not attempt to renumber them, since the list of overlapping attributes is presumably confusing enough already.

F0 Height-Contour	**Voice Quality**	**Duration**	**Number (class convention)**
High-rising	modal	long	1
High-falling	breathy	long	2
Mid	breathy	long	6
Low	modal/creaky	long	5
Low	creaky	long	3
Low	creaky	short	4

As seen in the acoustic report above, the reality of these characteristics is not so categorical. The tone space is compressed and its constituent contours shift around when provoked by different onsets. The range of H1-H2 values is also not great, and the allegiance of specific qualities to specific tones is hard to see outside a large sample of tokens. The discrete arrangement seen in the table above is probably driven at least partially by the author's own prejudice that different functional categories *ought* to be distinguishable; e.g. a voice quality difference allows the differentiation of the two low, long tones. In practice, the overlap in apparent voice quality for the two low, long tones probably just means that they aren't particularly distinct without contextual semantic and syntactic information. Furthermore, it is possible that distinguishing characteristics of the tones were unconsciously exaggerated by the speaker, since it the purpose of the task presented to him was probably transparent. I expect that tokens with tones 2 and 6 and particularly tones 3 and 5 would have low rates of successful identification if played to native speakers in ambiguous contexts. Considering degrees of variability and gradient distinctions, perceptual testing following recordings with more speakers should be interesting.

References

Andruski, J. E. (2006). Tone clarity in mixed pitch/phonation-type tones. *Journal of Phonetics*. 34: 388-404.

Andruski, J. E. & M. Ratliff (2000). Phonation types in production of phonological tone: the case of Green Mong. *Journal of the International Phonetic Association* 30 (1-2): 37-61.

Boersma, P. & D. Weenink (2010). *Praat: doing phonetics by computer* (Version 5.1.43) [computer program]. Retrieved from: http://www.fon.hum.uva.nl/praat/ on September 22nd, 2010.

Huffman, M. K. (1987). Measures of phonation type in Hmong. *Journal of the Acoustical Society of America*. 81 (2): 495-504.

Johnson, K. (2003). *Acoustic and Auditory Phonetics*. 2nd ed. Malden, MA: Blackwell Publishing Ltd.

Kato, A. (2007). Pwo Karen. In G. Thurgood & R. J. Lapolla (Eds.), *The Sino-Tibetan Languages* (632-648). New York: Routledge.

Pike, K. L. (1967). *Tone Languages*. Ann Arbor, MI: University of Michigan Press.

Stevens, K. N. (2000) *Acoustic Phonetics*. Cambridge, MA: MIT Press.

Weidert, A. (1987). *Tibeto-Burman Tonology: a Comparative Account*. (E. F. K. Koerner, Ed.) Amsterdam: John Benjamins Publishing Co.

Structural and Prosodic Similarities between Compounds and Lexicalized Phrases in Sgaw Karen

Jeffrey Wong

The following paper attempted to distinguish two types of words that are very obviously distinct in English but not in Sgaw Karen. These types are compounds, the combination of two words to form a new word such as *eggbeater*, and lexicalized phrases, expressions that look like sentences but that have a meaning that is not literal such as *shit hit the fan*.

1.0 Introduction

The transparent nature of the English language can be seen through the identification and analysis of compounds and lexicalized phrases. For starters, both are easy to identify. For example, in the sentence "John has two leaf-blowers" we know that leaf-blower is a compound and that *blower* provides the meaning of the word because it gets pluralized and not *leaf*. As for lexicalized phrases, unlike a normal sentence, their lexical nature makes it impossible to decipher the meaning from the parts that make them up. For example, in the sentence J*ohn is a couch potato* the lexicalized expression describes John as someone who rarely moves and not as some type of potato. The difficulty in identifying the meaning of the sentence is thus due to its idiomatic nature. This means that if you attempt to substitute the words in a lexicalized phrase, you risk removing its idiosyncratic meaning; c*ouch tomato* does not have the same meaning as *couch potato*. However, not all languages show this transparency. For example, Noyer (1998) argues that in Vietnamese there are no syntactic structural differences between compounds and lexicalized phrases. He posits a continuum in which words have meanings that are either available or unavailable. Following this trend, Sgaw Karen, a Sino-Tibetan language spoken by over one million speakers in Burma and Thailand, also exhibits this structural opacity between compounds and lexicalized phrases. Therefore, this brings up the question of whether in fact there are syntactic or even prosodic differences between compounds and lexicalized phrases in Sgaw Karen. To answer this question, I propose the following two-pronged solution: The inexistence of syntactic differences is explained by the continuum posited by Noyer, while the absence of prosodic differences is due to the pauses between words, and the lack of vowel reduction and intonational differences.

In order to prove the validity of this research question and its solution, a description of potential compounds and lexicalized phrases in this language will be given, which will be followed by a description of Noyer's continuum. Finally, an explanation on the application of the continuum to Sgaw Karen and on the absence of prosodic differences will round-out the components of this analysis.

2.0 Compounds and Lexicalized Phrases in Sgaw Karen

Based on the data collected, the main categories of potential compounds and lexicalized phrases found within the language were: professions, objects and periods in the day.

2.1 Professions

With professions, we notice that these constructions have a very particular structure in which an action pertaining to a job is nominalized. Here are some examples:

1) pwa˩ tusi˅ (pʰo)
person to farm person
"farmer"

2) pwa˩ ku˅ θe˅ (pʰo)
person to cut tree person
"tree cutter"

3) pwa˩ ka˩ʔ ta`˩ (pʰo)
person hunt or shoot marker person
"hunter"

4) pwa˩˽ ma˩˽ ɲa˅ (pʰo)
person work small fish person
"fisherman"

5) pwa˩ nɔ θole˅ (pʰo)
person to drive car person
"driver"

6) pwa˩˽ ʔɔ˅ hu˅ ta`˩ (pʰo)
person eat steal marker person
"robber"

Based on these examples, we can see that professions are derived using *pwa˩* (see last page of the paper for the description of the different symbols used for tones) which means *person* followed by a verb to designate the action described by the profession and the optional word *pʰo*. Furthermore, these examples bring up the question of whether there is a difference in meaning between the words *pwa˩* and *pʰo* which both mean *person*. The only visible difference between the two seems to be that the second word is optional. Before moving on to the reasoning behind *pwa˩* and *pʰo*, let us examine an interesting characteristic concerning some words in this language. As can be observed with the word *robber*, in example 6, the verb *eat* accompanies the verb *steal*. This is because the description of certain verbs in this language requires that the goal of the action be included. In the case of stealing, the goal is to eat. Getting back to *pʰo*, it is an optional word that can be omitted in the construction of most professions or when referring to a profession. For example:

7) ʔawɛ pwa˩tusi˅(pʰo) te lɛ lə ʔakʰɨt
this farmer past go his farm
"this farmer went to his farm"

In this sentence, the reference is made directly to the farmer and not to some person acting as a farmer. This means that the word *pwa˧˩tusi˅*, with or without *pʰo*, can only mean *farmer* in this sentence. The following sentences, on the other hand, distinguish the context in which *pʰo* is obligatory:

8) ʔawɛ po˧˩kwa me`˧˩ pwa˧˩˽ ma˧˩˽ ɲa˅ pʰo
this man is person to work small fish person
"this man is a fisherman"

9) ʔawɛ po˧˩kwa me`˧˩ pwa˧˩˽ ma˧˩˽ ɲa˅
this man is person to work small fish
"this man is fished"

Therefore, the only time *pʰo* is required is if the word refers to a person practicing a profession and there is a distinction in the meaning that can be made by its omission. In example 8, the man is the doer of the action of being a fisherman.

In example 9, the man is the receiver of the action of fishing. This seems to indicate that in some instances *pʰo* acts as a nominalizer changing the meaning of *ma˧˩˽ ɲa˅*, from the verb *to work with fish*, to *fisherman* and also changing the role of the subject from receiver to doer of the action.

2.2 Objects

Even though objects are referred to as a category here, this is done in a very general sense since there are many different sub categories such as fruits and vegetables and temporal objects such as the sky which have been grouped together. Here are some relevant examples of words referring to objects in Sgaw Karen:

10) kʰa˧˩ʔ kwe˧˩ li˧˩
piece furnit. write book
"desk or piece of furniture to write book"

11) kʰa˧˩ʔ se˅nɔ˧˩˽
piece furnit. sit
"chair or piece of furniture to sit"

12) mɨ kʰo˅
above head
"sky or above head"

13) hɔ˅ kʰo˅
under head
"earth or under head"

14) hiˇ kʰoˇ
house head
"roof or house for head"

15) kʰoˇ sʰuˇ→ hair
head hair
"hair or head of hair"

16) ta`ɬ sʰa hiˇ
marker pain house
"hospital or house of pain"

17) ta`ɬ θi hiˇ
nom. die house
"funeral home or house to die"

18) takʰɔɬ θaɬ
mango fruit
"mango"

If we examine the words with *head,* (examples 12-15) we notice an interesting fact related to how the Karen language represents the world. The uses of the word head in each of these constructions seem to indicate that the language attributes a certain representational power to this word; it refers to more than just the cranial structure attached to the body. It is also used to describe where an object is situated as in the word *sky* and also creates a symbolic representation just like in the word *roof*. As for the two words referring to places, examples 16 and 17, they contain *ta*`ɬ just like in the word *robber*, thus potentially demonstrating the marker's multi-functionality. Whereas in the word *robber*, found in example 6, it specified the main action verb, in the word *funeral home*, it acts as a nominalizer for the verb *die*, and in *hospital*, it takes on some other unidentified function.

2.3 Periods of the day

Words that describe periods of the day provide examples of some of the most interesting traits in this language. Here are some examples:

19) mu lɔɬ niɬ
sun down go
"sunset or sun down go"

20) mu he θɔˇ
sun comes up
"sunrise or sun comes up"

21) mu sa ka
sun bright time
"day time or time of sun bright"

22) mu na˩˩ ka
sun dark time
"night time or time of dark sun"

23) mu xi lɔ˩˩
sun run down
"afternoon or sun run down"

With these five examples we can see how periods in time in the Karen language are represented. They all refer to the sun doing different actions or taking on certain properties. Even in the case of night time, where instead of utilizing the word moon *la,* Karen opts to use a dark sun instead. Another interesting observation can be made when comparing afternoon and sunset. Even though *lɔ˩˩*, down, seems to assume the same function in both words, its position within the word is different. This may be because the ordering affects the meaning of the construction.

2.4 Summary

Overall, the examples of potential compounds and lexicalized phrases that have been provided here demonstrate the difficulty in distinguishing or even identifying compounds and lexicalized phrases in this language. Each example that was given can take on a double meaning, either the literal meaning or a collective overall meaning. For example, the expression *mu xi lɔ˩˩*, can either mean *sun run down* or *afternoon.* In the next section, an answer to this problem will be provided by examining the continuum posited by Noyer.

3.0 Noyer's Vietnamese Continuum

In his article entitled "Vietnamese Morphology and the Definition of a Word" (Noyer 1988), Noyer observes that within this language, two concepts affect compounds and lexicalized phrases to the point where they become indistinguishable. These concepts are availability and idiomaticity. The first concept refers to the productivity of parts of words and the latter refers to the transparency of the meaning of the word. He gives the example of the word for garden *vuòn tuoc* (Noyer 1988, p.74) to demonstrate these concepts. *tuoc* is glossed as meaning garden and *vuòn* is an essential word that has no meaning. Thus, we have a word with one unproductive and unidentified term on which the overall meaning is dependent upon.

This is an example of a word in which the terms have limited availability and the meaning is completely idiomatic. In contrast, Vietnamese also has words such as *dê quôc* (Noyer 1988, p.73), meaning *empire* which are much more transparent. *dê* is glossed as emperor while *quôc* is glossed as nation. Both terms are productive, they can occur in many words, and the meaning is less idiomatic, you can decipher the meaning of the word based on the parts that compose it. Using these two concepts, Noyer is able to posit a continuum, in which the availability of the meaning of a word, in this case compounds and lexicalized phrases, is affected by factors such as

the denotation of the expression. Fossilization may also be an influential factor since in the case of the word garden, *vuòn* has no meaning on its own but is required in the word. Therefore, it is possible that over time this word has lost its meaning but is still part of the word as a whole. Noyer calls the process by which the availability of the individual words is subjected to external factors, activation. Thus, it can be stated that the overall meaning of the word will only activate the individual meanings if it is required by the lexical context provided by the word. The following examples demonstrate the two possible options in relation to activation:

24) núi lửa
mountain fire
"volcano" (Noyer 1988, p.76)

25) nhà cửa
house door
"habitable home" (Noyer 1988, p.79)

In the first example, the activation of both words is required to convey the overall meaning of a volcano which is a mountain of fire. While in the second example, the individual meanings of *house* and *door* are compromised in order to form the idiosyncratic meaning of a habitable home.

The following section will deal with the application of the continuum as a means of explaining the structural similarities between potential compounds and lexicalized phrases.

4.0 The Continuum Applied to Sgaw Karen

Now that the research question has been outlined and Noyer's continuum has been defined, we will now proceed in applying the continuum to potential compounds and lexicalized phrases in Sgaw Karen. Although there exist similarities between the structural ambiguity found in compounds and lexicalized phrases in Vietnamese and Sgaw Karen, the distinction that can be made between the two is in relation to the nature of the context.

While in Vietnamese, the meaning of the word is dependent on the lexical context, in Sgaw Karen the meaning is influenced by the phrasal context. In the following examples which are from Sgaw Karen, these facts pertaining to the language are exhibited:

26) ʔawɛ mu te xi lɔ
the sun past goes down
"the sun went down"

27) jə ti mɨ kʰo
I see above head
"I see the sky"

In the first example, the individual meanings must be activated because of the contextual requirements imposed by the past marker in the first sentence. In other words, the past marker makes it impossible for the idiosyncratic meanings of the words to be activated, since you would

have a past marker followed by the noun *afternoon* which would make the sentence ungrammatical. In the other example however, the denotation of the words is compromised in favor of the idiosyncratic collective meaning which is conveyed due to the nonsensical nature of the literal meaning.

To conclude this analysis of compounds and lexicalized phrases, the explanation surrounding the inexistence of prosodic differences will be given in the next section.

5.0 Absence of Prosodic Differences

After analyzing the pitch and loudness of each word using Praat, a linguistic sound analysis and recording program, in isolation, within a construction and within a sentence, it was determined that due to the relatively similar results, there was on average a difference of 1 or 2 points, there were no prosodic differences between compounds and lexicalized phrases. This was due to pauses and an absence of vowel reduction and intonation differences. However, it must be noted that my understanding of the use of this program is very basic and may in fact be why prosodic differences were absent.

6.0 Strange Potential Compounds

Before finishing the analysis, it is important to consider the following last few examples which relate to potential compounds:

28) tʰi pu˩˩˦
water inside (hole)
"well"

29) -tʰi də
water pregnant (body)
"water bottle"

These examples might be compounds but a more logical theory would be to say that words like *pu˩˩˦* and *də* are not words at all but types of affixes, parts of words, and acting as markers just like *ta˨˩*. The reason why stating they are affixes is more logical has to do with the fact that the words *də* and *pu˩˩˦* are productive and appear in other words as well. It is particularly interesting that *də* literally means *pregnant,* but its use in certain contexts such as the word water bottle refers to the fact that water is contained within a body which in this case is a bottle.

7.0 Conclusion

To conclude, lexicalized phrases and compounds in Sgaw Karen showed no structural or prosodic differences. The latter was due to pauses and a lack of intonational differences and vowel reduction. Structural differences were also absent and this was due to the fact that these words apply the two concepts featured in Noyer's continuum which are availability and

idiomaticity which along with the influence of the phrasal context, dictate whether the individual meanings of words get activated or whether the collective meaning of the word forces the personal denotations to be compromised. Another factor that may have an effect is the fact that the language is also composed of mostly monosyllabic words. Hence, almost every word is one syllable in length and the words are independent from one another which impedes on processes which help distinguish whether words are pronounced together such as co-articulation. It is also possible that the reason for a lack of transparency between compounds and lexicalized phrases in this language is due to the fact that these two categories do not exist. Sgaw Karen has a very particular way of representing the world through words which may not require the use of complex structures such as compounds and lexicalized phrases. The lack of these structures may be compensated by an abundance of words with multiple meanings such as *head*, a very impressive classifier system and a series of markers that provide extra information to the context of a sentence. Another point that must be factored in is the fact that there was some uncertainty in relation to the identification of words, and this was due to the lack of intonational differences and the speaker's intuition which was alien to linguistic concepts such as affixation which could represent one of the steps required to fully answering this question.

Also, this task was further made complicated by the lack of concrete information concerning lexicalized phrases and compounds in Sgaw Karen in the literature. While many references were used in writing this paper, for most the part they only acted as sources for background information, looking up words and doing comparisons. Even when a source pertaining to compounds or lexicalized phrases in this language could be found such as Suriya Ratanakul, 1986a (check bibliography for complete source), there were too many inconsistencies. These inconsistencies created contradictions between what the literature was saying were compounds and what the consultant was implying were compounds. For example, in Ratanakul 1986a, an example of a Sgaw Karen compound that is given was the verb *to increase*. However, when the speaker provided the word for the action, *ʔaθɔ˩˩˦* it was very clear that this could not be a compound since this word is composed of a marker *ʔa* and the word up, *θɔ˩˩˦*. Hence, there is still a lot of work that needs to be done on this topic in order to eliminate these inconsistencies. Therefore, in future research, more data will need to be collected in order to complete the following objectives: Find further proof for the existence of lexicalized phrases and compounds in Sgaw Karen, establish a more concrete and stable syntactic structure for lexicalized phrases and compounds, and determine if an affixation hypothesis is a potential solution for explaining words with quirky functions such as *pu˩˩˦*, *də*, *ta* `˩ and $p^{h}o$.

References

Noyer, R. (1998). Vietnamese Morphology and the Definition of a Word. *UPENN Working Papers in Linguistics*. 5 (2): 69-92.

Other references that were used as secondary sources for background information, looking up words, and doing comparisons:

Blackwell, G. E. (1954). The Anglo-Karen dictionary. *Rangoon: Baptist Board of Publications*. 1-344.

Brown, N. (1854). Comparative vocabulary of the Sgau and Pwo Karen dialects. *Journal of the American Oriental Society.* 4:317-326.

Gilmore, D.C. (1896). A Karen grammar. *Rangoon: American Baptist Mission Press*.

Suriya R. (1986)a. A study of compound nouns in Lawa, White Meo, and Sgaw Karen. *Languages in global perspective: Papers in honor of the 50th anniversary of the Summer Institute of Linguistics*. 549-556.

Wade, J. (1842). The grammar of the Sgaw and Pgho Karen language. *Tavoy: American Baptist Mission Press.*

Wade, J. (1849). The vocabulary of the Sgau Karen language. *Tavoy: American Baptist Mission Press.*

<u>Tone Chart</u>

Tone 1 (high+ rising): default

Tone 2 (low + slightly falling):`ɬ

Tone 3 (low):ɬ

Tone 4 (mid-high+ glottal stop):ɬʔ

Tone 5 (falling + rising): ˇ

Tone 6 (low + breathy):ɬ̤

Classifiers in Sgaw Karen

Guilherme Cervo

This paper deals with Classifiers in Sgaw Karen Language, dialect spoken in Burma. Sgaw Karen is a Sino-Tibetan language that makes use of lexical items to classify the referent of a noun according to its meaning. Classifiers can be 'classifiers by nature' or noun-derived classifier. They can refer to the shape of objects (round, flat, long), to their taxonomy (fruit, animal, person) or can be used to quantify arbitrary measures (money, weight, speed). The objective of this paper is to enumerate the classifiers of Karen language and also their position in the sentence. Do they all follow the same structure or are there differences within the same system?

To answer this question I have presented a number of visual stimuli on a computer screen to a native speaker of the language and recorded how he says them and what classifiers he uses. The idea behind this type of procedure is to force him to say classifiers that cannot easily be explained. Many classifiers have been found, as well as, two types of classifier constructions based on the number of referents, one for constructions with numerals below 10 and another for constructions with numerals above 10.

1. Introduction

South East Asia is a geographical area that is extremely rich from a linguistic point of view, a region where a wide variety of language-families meet and interact with one another. One special characteristic of languages of this part of the world is the use of classifiers and quantifiers. Sgaw Karen, a Tibeto-Burmese language of Burma, makes some particular use of these particles. Classifiers are suffixes that attach to lexical items so that the noun to which they refer can be classified. The main function of a classifier is to make count nouns enumerable by individualizing and classifying them (Bisang, 1999). Although quantifying is the main function of classifiers, it is not the only one. In some languages, such as Vietnamese, classifiers go beyond counting and can even display the function of individualization and classification without being primarily involved in counting. In addition to being a referent and individualizer, classifiers also rise in contexts of possession and relative clauses.

In this paper the classifiers of Sgaw Karen will be examined. Suriya (1998) did and extensive and exhaustive research on the classifiers of "Sgaw Karen" but in her paper she relied on the dialect that is spoken in Thailand (according to our consultant). It seems that in this dialect of Karen, the spoken language has been largely influenced by Thai, and other native languages, spoken in the same region, or perhaps that the language just drifted apart from its vernacular variety. This would explain why many of the classifiers presented in her paper are phonologically different from those on the dialect under study, and why they are used in different contexts. The objective of this paper is, therefore, to compare the classifiers presented by Suriya and determine whether the classifiers match those of the dialect studied. Another objective of this paper is to briefly talk about their position in the sentence and in which environment they occur most frequently. I will not go into much detail about the syntax and semantics/pragmatics of classifiers in Karen since that is the work of my colleague Jean-Bruno.

2. Methods

In order to determine the classifiers of Karen, I asked our consultant to name random pictures shown on a computer screen. Since a good number of the classifiers presented by Suriya (1998) have different tones/are phonologically different of the dialect studied, I relied on her paper simply to create computer images of several situations in which classifiers are used, and that she grouped together in her article. For example, one of the classifiers on her paper was *pɔ*^{2} 'section of bamboo'. Images of bamboo were then created, or downloaded from the Internet, saved as a slideshow and then presented to our consultant. I followed Suriya's list of classifiers until more than 100 pictures had been put together so that the experiment could begin. When interviewed, our consultant was asked to look at each picture and pronounce them using numbers between one and 10. Given that it is extremely hard to ask someone to enumerate and name abstract objects, as well as conceptualize non-concrete ideas, this approach proved to be very efficient. Before we proceed, I'd like to make a comment on the tones used by our consultant. He wasn't very consistent with tones 0 and 1, or when there was a schwa present so I apologize in advance for any mistakes concerning them. I am confident however that our colleague doing the paper on tones will be able to explain their use thoroughly.

3. Numerals

Here is a list of the most relevant numerals in Karen. They will help you to understand some of the examples that will be used later in this paper. te^{0}

1. $k^{h}i^{0}$
2. $θe^{0}$
3. lui^{5}
4. je^{2}
5. $hɨ^{0}$
6. nui^{0}
7. $hɔ^{0}$
8. $x^{w}i^{0}$
9. $te^{0}si^{1}$

100. $te^{0}kə^{0}ɣa^{6}$ 1000.$te^{0}kə^{0}t^{h}o^{0}$

4. Type of Classifiers

Classifiers in Karen can be divided into two categories: noun-derived classifiers and unit classifier. Suriya (1998) has also mentioned verb-derived classifiers however, when tested on 'verb-derived' constructions, our consultant used either classifiers that fall under the category of unit classifiers or noun-derived classifiers, that is, those derived from noun-phrases.

4.1 Unit Classifiers

These morphemes are classifiers by nature. Their only function is to classify nouns and cannot be used as nouns themselves.

1. [dɨ^{1}]: this classifier is used with animals that walk with four legs: ‘dog’ *$t^{h}wi$*, ‘cat’ *$\theta a^{5} mi^{1} ɲɔ^{6}$*, and ‘monkey’ *$ta^{6} ʔɨ^{4}$*.
2. [ɣa^{6}]: classifier used for humans: ‘friend’ *$\theta a^{1} ko^{4}$*, ‘men’ *pwa^{6}*, ‘baby’ *$po^{1} \theta a^{5}$*, ‘president’ *$gɔ^{2} ko^{5}$*, ‘prime minister’ *$go^{1} to^{3} ko^{5} glə^{2}$*, ‘colonel’ *$\theta u^{4} ko^{5} glə^{2}$*, ‘dictator’ mɔ^{6}so^{1}ro^{6}
3. [k^{h}a^{1}]: classifier meaning item, thing: ‘one item’ *$ta^{2} ʔe^{5}$* *$tə^{1} k^{h}a^{1}$* , ‘two languages’ *$glo^{1} k^{h}i^{1} k^{h}a^{1}$*
4. [tɔ^{6}]: classifier used with school grades and dams. E.g., ‘grade 4’ *$lui^{5} tɔ^{6}$*
5. [p^{h}ɔ^{4}]: classifier used with ‘flowers’ *pho1ta6s^{h}ɨ^{4}* (or just *$p^{h}o^{1}$*)
6. [tu^{5}]: classifier used with ‘tree’ *$t^{h}e^{5}$*
7. [ɣɔ^{6}]: classifier meaning ‘time in space’: ‘next time’ *$lə^{0} k^{h}i^{0} tə^{0} ɣɔ^{6}$*
8. [ko^{5}]: classifier used with terrestrial vehicles: motorcycle, trains, bus, cars.
9. [lo^{5}]: classifier used with objects such as ‘toilet paper’ *$ti^{1} s^{h}ɨ^{1} ta^{1}$* and some root vegetables such as yam and potatoes.

4.2 Noun-derived Classifiers

Noun-derived classifiers are words that when used alone are nouns but that when used with a numeral function as a classifier. Noun-derived classifiers are used to quantify measures, time in space, groups of things and also simple units. As we shall see, the fact of being a ‘noun-derived’ classifier makes it hard for new learners to understand the choice for classifiers for certain things. Even though the standard shape of a building’ is a quadrilateral polygon, the classifier used with this noun derives from the word ‘round’ *$plə^{5}$* and can induce one to think that the building is round. This classifier is also used with ‘house’ and ‘boat’ even if the shape is not in accordance with what the classifier means when used by itself as a noun. English nouns like mile, pound, kilometer, gram, gallon, second, minute, Celsius, Fahrenheit, etc., have been lexicalized and are used as classifiers when following a numeral. When by themselves, they function as noun. I will not be listing these classifiers since they are borrowed words from English. It is good to keep in mind though that even if our consultant says that Karen language doesn’t have classifiers to quantify mass noun (such as nouns to weight objects), and that they use English words instead, that it would be unlikely that such classifiers don’t exist. It is more likely that these classifiers got lost in time and were replaced by more accessible nouns by the speech community.

Here is the list of noun-derived classifiers. Numbers 1 to 22 are noun-phrase units that function as classifiers when classifying, or qualifying, an object.

1. [bo^{0}]: word meaning cylindrical, long, string-like. When following a numeral, it is used as a classifier for either animate or inanimate objects: ‘python’ *$kə^{0} lɔ^{6}$*, ‘dragon snake’ *$pə^{0} ju^{6}$*, ‘ray’ *$mu^{2} ʔa^{1} je^{6}$*, ‘hair string’ *$ko^{5} \theta u^{5}$*, ‘cigarette’ *$mo^{5} thu^{0}$*, ‘ring’ *$pə^{1} si^{4}$*. Also used with crocodile, whale, shark, and catfish.
2. [be^{5}]: word meaning flat used as a classifier after numerals: ‘bed’ *$lo^{1} mi^{5}$*, ‘book’ *li^{1}*, ‘shirt’ *$s^{h}ɛ^{4} ka^{1}$*, ‘knife’ *$dɔ^{1}$*, ‘scissors’ *$ta^{4} ri^{4}$*, ‘sword’ *na^{4}*, ‘window’ *$tə^{0} rɛ^{6}$*, ‘wall’ *$da^{2} du^{5}$*, ‘picture’ *$ta^{2} ɣɔ^{6}$*, ‘teeth’ mɛ^{1}, ‘blade’ *$lə^{1} wa^{4}$*, ‘mirror’ *$mɛ^{5} thi^{1} kəla^{6}$* ‘stamp’ *$ta^{2} ɣɔ^{6} ko^{5}$* ‘airplanes’ *$kə^{1} bɔ^{1} ju^{6}$* , ‘boat’ kli^{1}. This classifier is also used with the words: ship, plate, flag, fish and birds.
3. [plə^{5}]: as a noun it mean ‘round’. After a numeral it becomes as classifier for ‘building’ *$da^{2} \theta ut^{h}ɔ^{5}$*, ‘house’ *$hi^{5}$*, ‘hat’ k^{h}o^{1}plɔ^{1}, ‘rocket/missile’ glo^{1}kli^{1}, ‘star’ s^{h}a^{5}, ‘stone’ le^{1}. Also used with TV, computer, sewing machine and instruments, such as guitar and piano, independently of the shape.

4. [de^{0}] : noun meaning ‘bottle’. It becomes a classifier when following a numeral: *t^{h}i-k^{h}i-de^{0}* ‘2 water bottles’, *beer-tə-de^{0}* ‘one beer bottle’ (according to our consultant, Karen speakers use the same noun as in English to refer to ‘beer’).

5. [ke^{1}tru^{2}] : noun meaning volume, collection. Used as a classifier after a numeral. E.g., our consultant was shown a picture of an encyclopedia with eight books in it, *li^{3} təke^{1}tru^{2}*. The same word was used when shown a compilation of CDs and DVDs.

6. [blɔ^{0}] : noun meaning ‘time’. Used as a classifier after numerals. E.g., ‘five times’ *lui^{5}blɔ^{0}*

7. [kɔ^{0}] : noun meaning ‘side’, used as a classifier after numerals. It can refer to any side of on object, such as left/right side. In the passage read by our consultant there was an example of ‘side’ as in a fight (it is easier to think of a ‘fight round’) in which the husband had to fight the python and win at least ‘one side of the fight’ *[…me^{2}tə^{0}nə^{6} tə^{0}kɔ^{0}…]*

8. [s^{h}ɔ^{6}]: as a noun it mean ‘hand’ and as a classifier it means ‘handful’: ‘two handful of rice’ *bu^{1}θa^{4} k^{h}i^{1}s^{h}ɔ^{6}*

9. [k^{w}a^{4}]: noun meaning for cup, mug, bowl. Used as classifiers after numerals. E.g., ‘two tea cups’ *la^{3}p^{h}a^{4}t^{h}i^{1} k^{h}i^{1}k^{w}a^{4}*

10. [kə^{1}do^{5}]: noun meaning ‘bunch’. Used as a classifier after numerals as in ‘two bunches of flowers’ *p^{h}o^{1}ta^{6}s^{h}ɨ^{4} k^{h}i^{1} kə^{1}do^{5}*

11. [ro^{1}]: noun meaning ‘roll’ used as classifier after numerals

12. [kə^{1}bli^{4}]: as a noun it mean ‘petal’ and it is used as a classifier after numerals: ‘three flower petals’ *p^{h}o^{1}θe^{0}kə^{1}bli^{4}*

13. [pa^{1}] : noun meaning ‘pile’ used as a classifier when referring to books, papers.

14. [tə^{1}pu^{1}] : noun meaning ‘group’. As a classifier it is used to refer to a group of people and a group of objects, things. E.g., ‘banana cluster’ *tək^{w}i^{0}θa^{5} tə^{1}pu^{1}*.

15. [ni^{6}]: noun meaning ‘day’ used as a classifier after a numeral

16. [ni^{5}]: noun meaning ‘year’ used as a classifier after a numeral

17. [nui^{0}]: noun meaning ‘week’ used as a classifier after a numeral

18. [na^{6}]: noun meaning ‘night’ used as a classifier after a numeral

19. [b^{w}e^{1}] noun meaning ‘penny’ used as a classifier after a numeral.

20. [θi^{4}]: noun meaning ‘quarter’. Used as a classifier after a numeral

21. [pa^{6}]: noun meaning ‘fifty cents’ used as a classifier after numeral

22. [ba^{4}]: noun meaning ‘loonie’ used as a classifier after a numeral

5. Word order of Classifiers

The word order of sentence with classifiers in Sgaw Karen is similar to the one in Thai. The order is Noun + Adjectival Phrase + Numeral + Classifier + Demonstrative. This is the standard order for declarative constructions with nine or less referents. When the numeral goes above 10, the word order changes and becomes: Noun + Adjectival Phrase + Genitive *ʔa^{1}* + Classifier + Numeral (10 or above) + Demonstrative. It is possible to notice that, in this type of construction, the classifier changes positions and receives a possessive grammatical prefix ʔa^{1}. This prefix is grammatically similar to the preposition *of* in English, and it is most likely a genitive case marker. Here are some examples of constructions involving classifiers:

A. [ʔa^{1}wɛ ʔɔdɔ mɛ^{5}t^{h}i1kəla^{6} tə^{1}-be^{5}]
3sg have mirror num-class
"He has a mirror"

B. [Jə-ʔɔ tək^{h}ɔθa^{5} k^{h}i^{1}-plə^{5}]
1sg-eat mango two-class
"I eat two mangos"

C. [t^{h}e^{5} hɔ^{0}-tu^{5}]
tree num. eight-class
"eight trees"

D. [t^{h}e^{5} ʔa^{1}-tu^{5} te^{0}kə^{0}t^{h}o^{0}]
tree poss-class num. thousand
"one thousand trees"

E. [pho^{1}ta^{6}s^{h}ɨ^{4} ʔa^{1}-ɣɔ Θe^{0}-p^{h}ɔ^{4} ʔi]
roses poss-red num-class dem
"These three red roses"

F. [pho^{1}ta^{6}s^{h}ɨ^{4} ʔa^{1}-ɣɔ ʔa^{1}-p^{h}ɔ^{4} te^{0}si^{1} ʔi]
roses poss-red poss-class num.ten dem.
"These ten red roses"

6. Discussion

Classifiers are lexical (not grammatical) items that attach to other lexical items so that the noun in which they refer to can be classified. In Sgaw Karen, classifiers can be classifiers by nature (meaning they only function as a classifier) or they can be derived from a noun. The standard word order for constructions with classifiers is Noun + Adjectival Phrase + Numeral + Classifier + Demonstrative. This construction is only used with numeral from one to nine. Once classifiers are quantifying nouns with more than 10 referents, the word order of structures with these morphemes change to Noun + Adjectival Phrase +the genitive *ʔa^{1}* + Classifier + Numeral (10 or above) + Demonstrative, as in the example F above

Suryia (1998) did a great job enumerating all the classifiers of the dialect she studied. Some of her classifiers are somewhat similar to those found in this research but they tend to differ phonologically. It seems however that she failed to extend her research to classifiers quantifying more than ten referents. She only mentioned one word order but there are evidences of two types of constructions. Unit classifiers in Karen tend to be category specific and refer to their taxonomic feature, while noun-derived classifiers are not. Although noun-derived classifiers tend to imply the meaning of the noun it derives from, their semantic meaning should be disregarded when it is used in the classifier position, that is, one should not think of all buildings being round or of all boats being flat. Also, when classifiers are attached to the numeral 1, it doesn't necessarily mean that the classifier is referring to one object only. It seems that in this case it is referring to objects in the general sense. For future research, it would be interesting to look at the semantic meaning of certain structures involving classifiers and their choice in regards to certain

abstract objects. Also, it would be interesting to spend more time with our consultant trying to establish what classifier is the default one for novel words (which in my opinion is either *be*^{5} or *plə*^{5}).

References

Becker, A. J. 1975. *"A linguistic image of nature: the Burmese numerative classifier system"*. *Linguistics* 165: 109-121.

Bisang, W (1999a.) Classifiers in East and Southeast Asian languages: Counting and beyond, in: Gvozdanovic, Jadranka. ed. Numeral types and changes worldwide, 113 - 185. Berlin: Mouton de Gruyter.

Greenberg, J. (1972). Numeral classifiers and substantival number: problems in the genesis of a linguistic type. Working Papers on Language Universals (Stanford University, dept. of Linguistics) 9.1:1-39.

Jones, R. B. (1970). Classifier constructions in Southeast Asia. Journal of the American Oriental Society 90.1:1-12

Ratanakul, S. (1997). Numeral classifiers in Sgaw Karen. Mon-Khmer Studies 28:101-113

La syntaxe et la sémantique des classificateurs numéraux en Karen Sgaw

Jean-Bruno Chartrand

This work deals with the system of numeral classifiers that is present in Karen Sgaw, a Sino-Tibetean language. In this language, when a noun is modified by a numeral entity, a classifier must adjoin to denote its shape or aspect. This work focuses on the syntactic structure of such constructions as well as their semantic scope both in simple cases and in longer stretches. This paper aims to shed light on questions of the sort: when and where do numeral-classifier constructions appear? What is their semantic significance? Can a noun appear without a classifier or simply why do numeral-classifiers exist? This paper however does not give an exhaustive list of all the classifiers since this is the project of another paper. Its gives an insight on what classifiers represent in Karen Sgaw.

1.0 Introduction

La majorité des langues d'Asie et d'Asie du sud-est ont cette particularité de posséder un système de classificateurs numéraux riche et complexe. Le Karen Sgaw ne fait pas exception à cette généralisation alors qu'il exhibe lui aussi un système plutôt développé de classificateurs numéraux. Ceux-ci apparaissent lorsqu'un nom est modifié par un nombre ou un mot indiquant la numération. Plusieurs études ont été menées sur les classificateurs mais peu fournissent une analyse détaillée du système en Karen Sgaw. Dans ce travail je me concentrerai sur la syntaxe des classificateurs numéraux, je traiterai aussi de ce qui force leur apparition en plus d'inclure une section sur la pragmatique de ces derniers.

2.0 Structure syntaxique simple des classificateurs numéraux

Nous savons désormais que le Karen Sgaw est une langue SVO (du moins dans les *root clauses*) qui exige un sujet mais laisse volontiers tomber les objets dans ses constructions. Ainsi, un nom ou un pronom doit toujours apparaître dans un énoncé. Lorsqu'un nom apparaît, il peut généralement être compté et si tel est le cas, un classificateur doit apparaître et s'adjoindre à la suite du nombre pour permettre de donner une information précise sur le nom dont on traite. Nous appellerons ces constructions des constructions de classificateurs. Elles sont donc composées selon l'ordre nom+nombre (ou mot marquant la numération) +classificateur. La liste exhaustive de ces derniers est donnée par un autre élève de la classe et donc je lui laisse le soin de dresser une liste complète de ceux que l'on peut retrouver en Karen Sgaw. Il est important de noter que l'ordre est toujours celui-ci mais que le nom peut être séparé de l'entité nombre+classificateur lorsque, par exemple, un verbe ou un adjectif apparaît. L'entité nombre+classificateur ne peut cependant jamais être décomposée. Ainsi on peut retrouver:

[t^{h}ɥi^{5} ki^{1}dɨ^{1}]
chien 2 cl.animal
'Deux chiens'

Mais aussi:

[tʰɥi⁵ ha⁴ ki¹di¹]
chien marcher 2 cl.animal
'Deux chiens marchent' ou 'Deux chiens qui marchent'

Il est alors intéressant de vérifier si le classificateur doit toujours s'adjoindre à une entité numérale ou encore si l'ordre est bel et bien toujours nom+nombre+classificateur ou si l'entité nombre+classificateur peut se retrouver devant le nom.

3.0 Structure syntaxique des classificateurs dans les énoncés plus complexes

Il a été mentionné plusieurs fois déjà qu'un classificateur doit apparaître et s'adjoindre à un nombre ou à une entité numérale si tel est le cas. Nous précisons ici 'une entité numérale' car le classificateur ne s'adjoint pas toujours à un nombre. Si le nombre n'est pas précisé mais que l'on sait que l'on parle de plusieurs noms, un mot se référant à la numération est présent, alors le classificateur apparaît et s'adjoint à un terme qui désigne la pluralité de l'élément nominal:

[tə¹kwi¹ta⁵ tə¹pʰa⁵ ʔa¹plə⁵]
banane un cl.groupement plusieurs cl.rond
'Un régime de bananes'

Dans un énoncé comme celui-ci, deux classificateurs apparaissent bien qu'un seul nombre ne soit présent. Le premier nombre s'adjoint au classificateur qui est un terme de mesure et qui précise qu'il y a un groupement de bananes, puis un deuxième classificateur doit apparaître pour préciser la forme de ce qui est décrit, mais comme le classificateur ne peut apparaître seul, un terme se référant à la numération apparaît, dans ce cas-ci plusieurs, et donc le classificateur peut être utilisé. Il importe aussi de préciser que pour un seul terme de numération, un seul classificateur peut apparaître. Encore une fois, il importe de préciser que s'il est vrai qu'un seul classificateur ne peut apparaître pour un nom, ce nom peut avoir différents classificateurs potentiels. Les langues à classificateurs numéraux possèdent donc cette particularité, contrairement aux langues européennes dont la flexion est toujours précise et univoque pour un terme, d'avoir différent classificateurs potentiels pour des noms dont l'aspect n'est pas nécessairement fixe. On peut ainsi retrouver:

[ko¹ ʔa¹sʰə¹ tə¹be⁵]
biscuit sucré un cl.plat
'Un biscuit sucré (plat)'

[ko¹ ʔa¹sʰə¹ tə¹plə⁵]
biscuit sucré un cl.rond
'Un biscuit sucré (rond)'

On remarque que la traduction littérale de ces deux énoncés donne le même résultat bien que le même nom puisse être représenté par deux classificateurs différents. C'est que le nom peut

prendre différents aspects. Ainsi, lorsque le locuteur natif devait traduire l'énoncé 'Un biscuit sucré', il demandait qu'elle était la forme du biscuit pour adjoindre le classificateur approprié. On peut donc se demander s'il existe un classificateur par défaut, ou un classificateur général, qui ferait surface dans les cas ou le locuteur ne connaît pas avec certitude l'aspect du nom à décrire. De tels classificateurs existent dans d'autres langues dont le Vietnamien qui adjoint le classificateur général *cái* pour désigner un objet inanimé ou encore *con* pour représenter un animal non-humain (Nguyen 1957). Il semblerait ici que la langue cherche à tout prix à intégrer les noms dans son système de classification même si le classificateur qui s'adjoint reste très général et peut ne pas représenter vraiment l'aspect du nom en question. Le Karen Sgaw diffère cependant du Vietnamien en ce sens qu'il ne possède pas vraiment de 'classificateur par défaut', bien que le classificateur [be^{5}], qui désigne quelque chose de plat, apparaît régulièrement dans les données même lorsque le lien sémantique entre le classificateur et le nom n'est pas explicite. Ainsi, ce classificateur est utilisé, entre autres, avec un ordinateur ou encore une télévision, des objets qui ont été intégrés plus tard au lexique de la langue et qui doivent toutefois obéir aux règles de classification, le fait que l'on utilise le classificateur pour les objets plats pourrait nous permettre d'affirmer que ce classificateur est le classificateur par défaut de la langue. Le locuteur soutient toutefois que le classificateur est plutôt utilisé pour représenter la forme de ces objets qui, selon lui, sont de plus en plus minces! Il ne semble donc pas y avoir de classificateur par défaut en Karen Sgaw.

Puis, si un adjectif apparaît pour décrire le nom qui est compté, le classificateur est toujours présent, l'adjectif peut s'insérer entre le nom et le groupement nombre+cl. ou encore peut se poster devant le nom et donc l'ajout d'un adjectif, quelque soit sa nature, ne modifie pas le système établi. Il est aussi intéressant de vérifier ce qui advient lorsqu'un démonstratif est inséré dans un énoncé. Suriya Ratanakul (1999) est un des rares chercheurs qui s'est intéressé à la structure des classificateurs numéraux en Karen Sgaw. Son article *Numeral classifiers in Sgaw Karen* présente une analyse intéressante des classificateurs pour le dialecte Mae Chaem de la langue, qui est un dialecte différent de celui de notre locuteur. Néanmoins, plusieurs constats sont similaires et il est intéressant de vérifier si l'ajout d'un démonstratif donnait les mêmes résultats que ceux répertoriés par Ratanakul. Dans la phrase nominale de base, l'ordre syntaxique se trouve à être, en Mae Chaem: Nom+Adjectif+Nombre+Classificateur+Démonstratif, par exemple:

[ʃe²ka² ɣɔ² sə³be¹ ʔi²][1]
chandail rouge 3cl.plat Dém.
'Ces trois chandails rouge'

On remarque alors que l'élément démonstratif se trouve à la fin de la phrase nominale. Cependant, si nous retrouvons une phrase nominale démonstrative sans nombre, alors le classificateur est optionnel (à noter que comme pour le dialecte étudié en classe, l'adjectif peut être repoussé en fin d'énoncé, son emplacement est plutôt aléatoire):

[koʔ¹ (be¹) ʔi² bɛ²ɣe²]
gâteau cl.plat dém. délicieux
'Ce délicieux gâteau'

1 À noter que cet exemple est tiré de l'article de Ratanakul et donc que la transcription diffère de celle de notre locuteur

Ceci semble être expliqué par le fait qu'un démonstratif dénote une entité en particulier et donc permet de relier un terme à une seule entité dans le monde. Cependant, comme une entité numérale n'est pas présente, le classificateur ne doit pas obligatoirement être présent, il devient facultatif. Ceci n'est pas le cas pour le dialecte de Karen Sgaw que nous avons étudié en classe, le classificateur doit s'adjoindre à une entité numérale que la phrase nominale soit démonstrative ou non, l'ordre que l'on retrouve pour ces constructions est habituellement démonstratif+nom+adjectif+nombre+classificateur avec l'adjectif pouvant être pré-nominal dans certaines constructions:

[ʔa¹wɛ¹ po³kwa¹ pʰa⁴tʰɔ ki¹¹ɣa⁶]
dém. homme grand 2 cl.humain
'Ces deux grands hommes'

De plus, si l'entité marquant la numération est enlevé de l'énoncé, le classificateur disparaît aussi et ce même si l'on sait que l'on traite d'un homme en particulier. Si un mot apparaît et sert à marquer qu'un nom est, ou pourrait être compté, alors un classificateur apparaît. Il semblerait que seuls les termes marquant la numération au pluriel peuvent permettre à un classificateur de s'adjoindre, les articles définis comme 'le' ou 'la' parlant d'un nom spécifique ne semblent pas avoir ce pouvoir. Ces articles sont, lorsque traduits, remplacés le plus souvent par un nombre.
On peut observer cet exemple additionnel provenant de l'histoire traduite en classe pour valider notre point:

[dɔ⁴ ʔa¹wɛ¹θe⁵ ʔɔ⁵wɛ¹ da⁵lə¹ θə¹wɔ ¹ tə¹plə⁵ ʔa¹pu⁶ne⁵lɔ⁶]
et ils vivre 3pp village un cl.rond intérieur fin phrase
'Et ils vivaient à l'intérieur du village'

On remarque que même dans un énoncé plus complexe qu'une simple phrase nominale, notre système préalablement établi est respecté. Le nom village est compté, on précise qu'il y en a un, le nombre suit et comme notre nom est compté, un classificateur précisant la forme du nom s'adjoint au nombre, dans ce cas-ci on utilise le classificateur [plə⁵].

4.0 La classification des noms, leur relation avec les classificateurs

La nature des noms semble aussi avoir une incidence sur le système de classificateurs et vice-versa. Il faut porter une attention méticuleuse à la distinction entre ce que nous pouvons appeler des noms de masse et des noms pouvant être comptés (*mass nouns and count nouns*). Une hypothèse concernant ces types de noms a été avancée concernant les langues d'Asie, il s'agirait de reconnaître tous les noms de ces langues comme étant des noms de masse[2]. Un nom de masse, est un nom qui doit être modifié par un élément de mesure ou un quantificateur pour être interprété. Par exemple, en anglais, on ne dit pas "a water" mais bien quelque chose comme "a litre of water", ces noms sont des noms de masse, ils désignent des entités qui doivent être modifiées pour pouvoir être comptées. L'anglais exhibe cependant les deux types de noms mais qu'en est-il des langues asiatiques? Les travaux de Cheng et Sybesma (1999) sur les

2 Plusieurs travaux traitent de cette hypothèse dont Allan (1977)

classificateurs numéraux en mandarin et en cantonnais nous permettent de soutenir la thèse inverse que celle des noms de masse *(mass noun hypothesis)*. Selon eux, certains noms possèdent dans leur bagage sémantique quelque chose qui fait d'eux des unités distinctes (des noms pouvant être comptés) et d'autre non et doivent alors être regroupés (des noms de masse). Ainsi, à la lumière des ces considérations, comment analyse-t-on la classification des noms en Karen Sgaw? Comme un classificateur apparaît seulement lorsqu'un nom est compté, et donc est modifié par une entité numérale, devons nous penser que ces noms ne sont que des noms de masse, car s'ils ne sont pas modifiés, ils ne sont pas comptés. Ou encore qu'il existe deux types de noms et que, bien qu'ils doivent tous posséder un classificateur lorsqu'il sont comptés, ils diffèrent en ce sens qu'il possèdent tous un élément dans leur bagage sémantique qui précise s'ils peuvent représenter des entités individuelles ou alors doivent obligatoirement être regroupés. Contrairement à l'anglais, notre système de classificateurs ne cherchera pas uniquement à rendre des noms qui ne peuvent pas être comptés, comptables, mais aura une autre portée que nous examinerons à la fin de ce travail concernant leur lien avec la perception du monde. Cette distinction étant apportée, nous n'affirmons pas que le Karen Sgaw ne possède pas de *massificateurs*[3] (classificateurs exprimant la masse d'un nom donné), seulement que son système nominal ne peut pas être analysé comme ne possédant seulement que des noms qui fonctionnent comme les termes 'water' ou 'sand' en anglais. Il en reste que les classificateurs en Karen Sgaw peuvent être distingués en ce sens qu'ils peuvent être des classificateurs d'unités nominales, des classificateurs dérivés d'une action ou des classificateurs dérivés d'un verbe. Aussi, certains classificateurs sont en effet des termes indiquant la mesure (*massifiers*) et d'autres cherchent à définir la forme ou l'aspect d'un objet ou d'une entité, je vous réfère au travail de mon collègue en ce sens. Dans ces classifications nominales, la langue possède toute une série de noms exprimant uniquement des regroupements d'entités comme c'est le cas avec ce que nous avons analysé comme des classificateurs pour des nombres à deux unités ou plus. Il existe aussi toute une classe de classificateurs qui désignent les mesures. Ratanakul distingue des classificateurs unitaires, des classificateurs de groupes, des classificateurs d'action, de mesure et enfin des classificateurs dérivés de verbes. Le dialecte de Karen Sgaw étudié en classe semble aussi respecter une structure similaire et pour cette section, une distinction importante entre classificateurs d'unités, donc des noms qui peuvent être comptés, et des classificateurs de groupes devait être apportée. L'exemple que je donne ici concerne le terme 'meat'. En anglais, 'meat' est un nom de masse, on ne dit pas 'a meat' ou 'two meats' mais plutôt 'a pound of meat' ou 'a piece of meat'. En Karen Sgaw, cette différence est traduite par la présence d'un différent classificateur:

[tə²ɲa⁵ tə¹lo⁵]
viande un cl.morceau
'Un 'morceau' de viande' désigne litt. 'Un viande'

[tə²ɲa⁵ tə¹pʰɔ²]
viande un kilo (unité de mesure)
'Un 'kilo' de viande'

Ainsi, dans les deux cas, c'est le classificateur qui précise si l'on traite d'un nom de masse, ou encore d'un nom pouvant être compté. On remarque aussi que dans les deux cas, un classificateur

3 Terme emprunté à Cheng et Sybesma (1999)

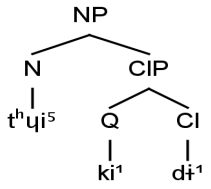
doit apparaître pour indiquer la mesure du nom 'viande', un exemple de ce type milite donc en faveur d'une analyse comme celle de Cheng et Sybesma, il faut analyser des noms comme étant de masse et d'autres, comptables, puis d'autres qui dépendent du classificateur utilisé. Le fait qu'il existe aussi des classes de classificateurs qui cherchent avant tout à modifier des unités et d'autres (comme les classificateurs pour le temps, les nombres ou les mesures) qui modifient un groupement supporte aussi la thèse énoncée précédemment. Notre analyse n'écarte toutefois pas entièrement l'hypothèse que tous les noms sont en fait des noms de masse et que le classificateur doit apparaître, comme c'est le cas en anglais, pour permettre de les compter. Du fait que l'entité nombre+ classificateur soit indécomposable et que lorsqu'un nombre apparaît, il exige la présence d'un classificateur. La distinction sémantique des noms en Karen Sgaw supporte cependant la thèse de Cheng et Sybesma, les noms peuvent eux-mêmes être classés entre noms qui représentent des masses d'entités et des noms qui peuvent être comptés.

5.0 Les classificateurs numéraux lexicalement et leur position dans l'arbre syntaxique

Une autre question doit être adressée dans ce travail, quelle sorte d'entité lexicale est un classificateur numéral? Si l'on examine l'anglais, on peut analyser les cl. numéraux comme étant une sorte de quantificateur (ou encore des *massifiers* pour emprunter la terminologie de Cheng et Sybesma), par exemple les mots 'litre' ou 'pound' dénotent bien une quantité et permettent à un nom de masse de pouvoir être compté. Mais nous avons vu en Karen Sgaw que les cl. numéraux ne s'adjoignent pas seulement à des noms de masse et qu'il faut tenir compte de la nature du nom. En ce sens, ils ne dénotent pas toujours un quantificateur. Ce qui a été proposé, c'est d'analyser les classificateurs comme une classe grammaticale à part entière et de définir la formation nombre+ classificateur comme une construction de classificateur. Nous retrouverions alors un arbre syntaxique de ce genre pour ces constructions simples:

[t^hɥi^5 ki^{1}dɨ^1]
dog 2 class.animal
'Two dogs'

Notre construction de classificateurs est donc formée d'un nombre, un quantificateur, plus un classificateur. Cette construction est cependant simple et les constructions complexes présentent aussi des arbres complexes mais l'essentiel est de définir le classificateur comme une catégorie et de dire que son adjonction à un nombre crée une construction de classificateur.



6.0 La pragmatique et les rôles des classificateurs numéraux

" Chaque langue, écrit Whorf, est un vaste système de structures, différent de celui des autres langues, dans lequel sont ordonnées culturellement les formes et les catégories par lesquelles l'individu non seulement communique mais aussi analyse la nature, aperçoit ou néglige tel ou

> tel type de phénomènes et de relations, dans lesquelles il coule sa façon de raisonner, et par lesquelles il construit l'édifice de sa connaissance du monde [...]. Nous disséquons la nature suivant des lignes tracées d'avance par nos langues maternelles. "[4]

Je débute la dernière section de ce travail par une courte citation de Whorf qui a une incidence directe avec notre système. Cette thèse stipule que le langage est ce qui détermine notre perception et notre vision du monde. À travers mon étude des classificateurs en Karen Sgaw, j'ai aussi été confronté à cette relation entre langage et perception du monde et en suis venu à remettre en question, en quelque sorte, la thèse Sapir-Whorf. Ainsi, ce serait les outils linguistiques que nous possédons qui devraient sculpter la vision que nous avons des choses et notre perception du monde. L'étude de la structure des classificateurs apporte cependant des éléments qui peuvent permettre de soutenir la thèse inverse. En Karen Sgaw, comme dans plusieurs langues d'Asie et de langues qui fonctionnent avec un système de classificateurs numéraux, les classificateurs existent avant tout pour exprimer une propriété du nom qui est décrit. Cette propriété tourne souvent autour de l'aspect ou de l'apparence de l'objet décrit. Ainsi, la relation entre classificateur et objet repose sur un fort lien sémantique car on classifie avant tout les objets selon ce que les sens nous dicte. Rosch (1973) a été un des chercheurs les plus influant à s'opposer à la thèse Sapir-Whorf par son étude du système des couleurs en Dani, une langue de Nouvelle-Guinée. Cette langue ne possède que deux mots pour représenter des contrastes entre les couleurs, un pour les couleurs pâles et chaudes et un pour les couleurs foncées et froides. Il a réussi à prouver que les locuteurs réussissaient à faire des distinctions extralinguistiques, ils avaient les mêmes capacités que les locuteurs de langues avec un vocabulaire plus riche pour les couleurs pour manipuler des concepts relatifs aux couleurs. Ils pouvaient même incorporer des noms 'artificiels' de couleurs pour manifester des distinctions perceptuelles. Ainsi, on retrouve l'idée inverse, celle de créer des entités linguistiques pour rendre état de distinctions perceptuelles que le système existant ne permet pas de distinguer. Je crois que nous retrouvons quelque chose de similaire pour les langues à classificateurs numéraux dont le Karen Sgaw bien que la distinction ne soit pas aussi évidente. À première vue, le système de classificateurs numéraux semble soutenir la thèse Sapir-Whorf car il semblerait que des éléments soient intégrés dans des classes qui ne les définissent pas vraiment sémantiquement, comme si la langue voulait absolument intégrer tous les éléments dans le système même si des classes linguistiques manquent pour traduire les contrastes. Nous pouvons donner en exemple le cas du mot 'village' qui est associé au classificateur 'rond' ou celui de 'téléphone' associé au classificateur 'plat'. Il en reste qu'un lien sémantique peut quand même être établi, un village peut être organisé selon une forme ronde et un téléphone peut avoir un aspect plat, ce sont sans doute les classes qui se rapprochent le plus de ce que la langue permet de décrire . Si la langue ne crée pas de nouvelles classes pour intégrer des termes pour lesquels les classes existantes ne sont pas appropriées, comment peut-on soutenir que le système s'oppose à la thèse de Whorf? C'est qu'en fait, à la base, ce serait la perception du monde qui nous amènerait à structurer notre langage de telle ou telle façon. Le français ou l'anglais ne possèdent pas de systèmes de classificateurs numéraux comme celui du Karen Sgaw, il faut alors penser que la langue reflète des différentes visions du monde. En Karen Sgaw, l'ajout d'entités marquées dans le lexique prend la place de ce que nous appelons flexion et la remplace par un terme qui donne une information sur ce qui est décrit. Ainsi, lorsque le locuteur était appelé à intégrer dans son lexique un mot dont il ne

4 Extrait des notes de linguistique française 2008 (citation de Benjamin Lee Whorf)

connaissait pas la signification et donc ne pouvait pas associer un lien perceptuel, il ne l'intégrait tout simplement pas affirmant qu'il ne pouvait traduire un mot sans connaître ce à quoi le mot ressemble. Aussi, lorsque appelé à traduire un mot pouvant avoir différents classificateurs, comme le mot 'biscuit', il demandait avant tout d'expliquer la forme du biscuit, l'intégration linguistique passe alors par une association entre un 'signifiant' et un 'signifié'. J'inclue cette partie comme simple hypothèse que je pourrai développer au cours de ma maîtrise et cette section reste une simple hypothèse car il est évident que pour la soutenir, il faudrait la tester expérimentalement pour vérifier s'il est vrai que le système linguistique du locuteur repose avant tout sur sa perception du monde et que ce n'est pas au contraire les outils linguistiques qu'il possède comme une sorte de *built-in* sémantique qui lui façonne sa vision de monde et traduit une perception qui est en réalité bien différente de ce qu'il veut vraiment représenter.

7.0 Discussion et conclusion

Ce travail a présenté une étude variée de la syntaxe des classificateurs numéraux en Karen Sgaw, de leur rôle dans la langue et de l'implication de ces derniers pour la sémantique et la pragmatique. Cette étude m'a aussi permis d'aborder le thème de la relation entre langage et perception du monde et de proposer une possible corrélation entre les langues à classificateurs numéraux et cette relation. Il en reste que plusieurs études expérimentales restent à être conduites sur le sujet des classificateurs numéraux pour permettre d'opposer à la thèse Sapir-Whorf un schéma différent que celui établi précédemment. D'autres sphères du langage peuvent aussi bénéficier de l'étude des classificateurs numéraux comme la syntaxe ou encore la morphologie. Il sera intéressant d'observer si d'autres chercheurs se pencheront sur le sujet dans les années à venir.

Références

Aikhenvald, A. Y. (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford Universty Press

Allan, Keith. (1977).*Classifiers*. Language, 53, 2, 285-311.

Becker, A. J. 1975. *"A linguistic image of nature: the Burmese numerative classifier system"*. *Linguistics* 165: 109-121.

Kay, Paul and Kempton, Willet.1984. *What is the Sapir-Whorf hypothesis?* American Anthropologist, News Series, Vol.86 no.1 .pp.65-79

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought.* Chicago, Illinois: The University of Chicago Press.

Nguyen, Dinh Hoa. *Classifiers in Vietnamese*. Word 13: 124-152.

Ratanakul, Suriya. *Numeral-classifiers in Karen Sgaw.* Mon-Khmer Studies 28: 101-113

Reilly, Ronan. *Sandy ideas and coloured days: Some computational implications of embodiment*. Artificial Intelligence Review. Volume 8. 1995

Rosch, E. (1973). *Natural categories*. Cognitive Psychology, 4, 328-350.

Cahier des notes de linguistique française 2008.

Notes des entretiens avec un locuteur natif du Karen Sgaw pour le cours LIN7911

Negation in Karen Sgaw

Catherine Gibb

Karen Sgaw is a tonal language spoken in Burma by the Karen people. It has a simple, subject-verb-object word order, with basic phonology, morphology and syntax. Negation in this language is relatively simple. It consists of two morphemes that affix to the verb in order to add negation to the phrase. While the first morpheme is obligatory in all contexts, the second morpheme does not seem to be obligatory in all contexts, as examined below. This paper examines negation in varying contexts, constructions and circumstances.

1.0 Introduction

Karen is a language spoken in eastern Burma by the Karen people. Karen Sgaw is a particular dialect of the Karen family of languages. It is a tonal language, with 6 tones representing contrasts in the language. The sound system is relatively simple, as is the morphological and syntactical structure. The language uses a basic SVO word order.

Most of the research on Karen Sgaw has been done into the sound system of the language (see Gilmore 1898, Manson 2009). However, some research has also been done into the use of adverbs (Ratanakul 2001) and the use of numeral classifiers (Ratanakul 1998). Limited research has been done in the syntax of Karen Sgaw, although some research on the sound system also references syntax briefly.

In particular, Gilmore's (1898) review of Karen Sgaw briefly discusses negation. His analysis concludes a simplistic review of the morphology and phonology associated with negation. Generally, he concludes that /ba^{5}/[5] is used at the end of negative phrases and that /tɤ^{1}/ is affixed to the verb. Gilmore also makes note of a morpheme used in negative imperative constructions, namely tɤ^{1}ɣe^{6}. T he following paper supports his findings; however, a more in depth analysis of the properties of negation in Karen Sgaw is required.

2.0 Basic Verb and Adjective Constructions

The word order in Karen Sgaw is SVO, and subject pronouns are obligatory. An example of basic word order can be seen in (1).

(1) t^{h}wi^{5} ʔɔ^{5} s^{1}pi^{4}θa^{5}
dog eat grape
'the dog eats grapes'

[5] Rev. Gilmore's research is cited in the traditional Karen script. However, to aid in clarity and understanding, I will use the IPA to refer to his work.

There is evidence that adjectives function in the same way as verbs, as there is no verb present between the noun and the adjective, as seen in (2). Similarly, adjectives can receive tense markers that modify the time at which the adjective will apply to the noun, as seen in (3).

(2) ʔa^{1}we^{1} po^{3}kwa^{1} θa^{4}kɨ^{0}

this man happy
'this man is happy'

(3) ʔa^{1}we^{1} po^{3}kwa^{1} kɛ^{1}θa^{4}kɨ^{0}
this man happy.future
'this man will be happy'

3.0 Negation in Basic Verb Structure and Adjective Constructions

Karen Sgaw has a generally regular pattern to construct negative phrases, as both verbs and adjectives function the same way in the language. The negative construction consists of the morpheme tɤ^{1}- affixed to either the main verb or adjective in the phrase and the free morpheme ba^{5} placed at the end of the phrase. Basic negative construction with one verb is shown in (4), with two verbs in (5), and with adjectives in (6) and (7).

(4) jɤ^{1} tɤ^{1}ha^{4} ba^{5}
I neg.walk neg.
'I am not walking'

(5) jɤ^{1} tɤ^{1}ʔe^{5}dɔ^{4} ha^{4} ba^{5}
I neg.want walk neg.
'I do not want to walk'

(6) jɤ^{1} tɤ^{1}θa^{4}kɨ^{1} ba^{5}
I neg.happy neg.
'I am not happy'

(7) nɤ^{1} tɤ^{1}θa^{4}ʔu^{4} ba^{5}
you neg.sad neg.
'you are not sad'

In (4), there are two verbs in the construction, and tɤ^{1}- is affixed to the main verb, and ba^{5} occurs at the end of the phrase. Adjectival phrases are negated in the same fashion. A more complex version of a negative adjective phrase is shown in (8). Here, there is a more complex subject and the morpheme tɤ^{1}- is affixed to the adjective, followed by ba^{5}.

(8) pɔ^{1} ʔa^{1}we^{1}ne^{5} tɤ^{1}pɔ^{4} ʔi^{6} tɤ^{1}xwi^{1}la^{6} ba^{5}
flower that one 1.flower prep. neg.pretty neg
'that flower is not pretty'

4.0 Negation in Subordinate Clauses and Double Negation across Clauses

When there is negation in subordinate clauses, the negative morpheme tɤ^{1} affixes to the verb that is being negated (in the subordinate clause) and ba^{5} falls at the end of the phrase. When negation appears in two clauses, one of which is subordinate, the morpheme tɤ^{1} is affixed to both verbs and ba^{5} appears at the end of phrase. An example of negation in a subordinate clause is shown in (9), and an example of negation across clauses is shown in (10).

(9) jɤ^{1} ba^{5}θa^{4} t^{h}wi^{5} tɤ^{1}dɨ^{1} lɤ^{1} ʔa^{1} tɤ^{1}ʔe^{5}da^{2} ba^{5}
I like dog 1.animal prep. it neg.bite neg.
'I like the dog that doesn't bite'

(10) jɤ^{1} tɤ^{1}ba^{5}θa^{4} t^{h}o^{5} tɤ^{1}be^{5} lɤ^{1} ʔa^{1} tɤ^{1}ju^{6} ba^{5}
I neg.like bird 1.flat prep. it neg.fly neg.
'I don't like the bird that doesn't fly'

There is no evidence that multiple negative markers can be included in one phrase. The data suggests that when negation exists embedded in subordinate clauses as well as across clauses, the morpheme tɤ^{1} negates the verb, and ba^{5} only occurs at the end of the entire phrase, rather than at the end of each individual clause.

5.0 Negative Questions

Negation constructions in questions are similar to the constructions shown above. Two examples are shown in (11) and (12).

(11) ba^{5}mɤ^{1}nɨ^{6}ʔa^{1}ho^{1} lɤ^{1} nɤ^{1} tɤ^{1}ɣe^{6}kɤ^{1}li^{3} ba^{5} le^{5}
why prep. you neg.dance neg. question
'why aren't you dancing?'

(12) me^{1}da^{6} tɤ^{1}ɣa^{6} lɤ^{1} ʔa^{1} tɤ^{1}ʔɔ^{5} pɛ^{1}ʔi^{6} le^{5}
who 1.pers prep. 3sg. neg.3sg. here question.
'who is not here?'

In (11), tɤ^{1}- is affixed to the verb in the same was as in (4) and (5). Questions in Karen Sgaw are denoted by a question marker, in this case le^{5}. In negative question constructions, ba^{5} precedes the question marker, in this case immediately following the negative verb, as in (11). In (12), tɤ^{1}- is contracted into ʔɔ^{5}, which suggests that it does not require a verb or an adjective to attach to for the phrase to be grammatical. ʔɔ^{5} is a morpheme that marks the 3rd person singular. The negation of ʔɔ^{5} is discussed further with negative polarity items.

Another item worth noting is that in (12), the morpheme ba^{5} is dropped from the construction. This appears in other situations as well, which suggests that, of the two negation morphemes, tɤ^{1}- is required and ba^{5} is not obligatory. When pressed, the consultant stated that ba^{5} is meant to be used in (12); however, it is evident that it may be omitted due to casual speech. Another example where ba^{5} is omitted is shown in (13).

(13) nɤ^{1} tɤ^{1}t^{h}i^{5} ʔɔ^{6} lɤ^{1} tɤ^{1}nwi^{1} ʔa^{1}tɔ^{2} pu^{6}
you neg.see her prep. 1.week in period
'you will not see her for a week'

Negative tag questions show predictable structure based on the previous discussion and examples. The morpheme tɤ^{1} is affixed to the verb, ba^{5} follows the phrase, and the phrase is ended by a question marker, which in this case is a̱^{1}. An example of a tag question is shown in (14).

(14) nɤ^{1} tɤ^{1}t^{h}i^{5} ʔɔ^{6} ba^{5} a̱^{1}
you neg.see her neg. question
'didn't you see her?'

6.0 Negative Imperatives

Negative imperative constructions (or commands) in Karen Sgaw are very different from negation in basic verb constructions. Both morphemes are required and contracted together, as well as a new construction at the end of the phrase. The examples in (15) and (16) show two similar constructions of imperatives in Karen Sgaw.

(15) tɤ^{1}ba^{5} xwi^{2} lɤ^{1} hi^{5}pu^{6} tɤ^{1}ɣe^{6}
neg. run prep. house neg.
'don't run in the house'

(16) tɤ^{1}ba^{5} ʔɔ^{1} la^{3}p^{h}a^{4}t^{h}i^{1} tɤ^{1}ɣe^{6}
neg. drink tea neg.
'don't drink tea'

Here, tɤ^{1} and ba^{5} are contracted together to form a negative morpheme used only in imperatives. Another morpheme is added at the end of the phrase, which includes the morpheme tɤ^{1}- and functions to enforce the meaning of the imperative phrase. It is not found elsewhere in the data collected thus far.

7.0 Negative Polarity Items

Negative polarity items function unusually in Karen Sgaw. For example, see the phrase in (17)

(17) pwa^{6} tɤ^{1}he^{1}so^{5}ni^{2} ʔɔ^{6} da^{2}he^{5} nɔ^{1} tɤ^{1}mi^{6} ba^{5}
people neg.bring her gift none 1.thing neg.
'nobody brought her a gift'

As seen in (17), there is no item that is specifically identified as the negative polarity item. The negative morpheme tɤ^{1}- is contracted into the verb 'to bring', and ba^{5} follows the phrase at

the end. Another morpheme is also added between the noun and the quantifier. This pattern can be seen in (18), although the negative quantifier does not follow the noun directly.

(18) pwa^{6} θa^{4}sa^{2} tɤ^{1}ʔo^{5} pɛ^{1}ʔi^{6} nɔ^{1} tɤ^{1}ɣa^{6} ba^{5}
people young neg.live here none 1.person neg.
'there are no young people here'

The examples here, as well as example (12), show that personal pronouns can be negated. When the negation is of a negative polarity item, as in (17) - nothing - and (18) – no young people - a negative quantifier, nɔ^{1} is used to show that there is none of the noun being quantified. When the negation is of a personal pronoun, as in (12), the negative morpheme tɤ^{1} is affixed directly to the pronoun. This pattern results in a type of double negation. The first negation is of the verb, and the second negation is of the noun. However, this double negation is understood as one negative meaning, rather than negation of the first negation.

Another negative polarity item that functions differently than the ones discussed above is shown in (19).

(19) nɤ^{1} tɤ^{1}kɹɤ^{4} hɔ^{5} nɔ^{1}tɤ^{1}blɔ^{1} ba^{5}
you neg.should cry never neg.
'you should never cry'

In the above example, the negative polarity item is represented by a singular word as well as having the morpheme tɤ^{1}- contracted into the verb. As can be seen in (19), the morpheme nɔ^{1} is shown in the word 'never'. This suggests that there is some compounding related to the negation of temporal aspects, with a negative morpheme perhaps being contracted into a quantifier. The literal translation of the compound could be 'no one time'. However, it is not clear if the morpheme tɤ^{1} functions as the quantifier for 1 or as a negative morpheme. Another example of the use of negation of a temporal aspect is shown in (20).

(20) bɤ^{1} tɤ^{1}le^{6} tɤ^{1}ɣa^{6} nɔ^{1}tɤ^{1}blɔ^{1} ba^{5}
we neg.go 1.person never neg.
'we never go alone'

In this example, the meaning for 'alone' is represented by the quantifier tɤ^{1}ɣa^{6}, which represents one person. Another way to represent the meaning of 'alone' is shown in (21). This example was found in isolation, and therefore is shown in isolation, not in the context of a phrase.

(21) t^{h}ɛ^{1}tɤ^{1}ɣa^{6}o̤^{6}
alone
'alone'

The above examples in (19) and (20) suggest that there is negation of the verb, as well as a negative polarity item in the phrase. This, along with the examples in (12), (17) and (18), show that double negatives are often present in the language, however, they do not hold the same meaning as they do in English. In Karen Sgaw, the double negatives have a singular negative meaning, and not two negative meanings that negate one another.

8.0 Conclusion

Negation in Karen Sgaw functions very similarly across different verb constructions, by affixing tɤ^{1}- to the verb, and completing the phrase with ba^{5}. The construction is the same in questions, as well as in subordinate clauses and across clauses. The morpheme ba^{5} does not seem to be required based on certain examples in the language, (12) and (13), and may be dropped due to casual speech. Finally, negative polarity items function with an additional morpheme, nɔ^{1}, which may or may not be contracted into the negative polarity item. Similarly, there is a significant amount of double negation, although it does not function in the same way as in English. This double negation is particularly evident when negative polarity items are used in the phrase. Future research may look at negative compounding, or negation in more complex structures, as well as double negation of negative polarity items and negation of personal pronouns.

References

Gilmore, D. (1898). A Grammar of the Sgaw Karen. *American Baptist Missionary Press*.

Manson, K. (2009). Prolegomena to reconstructing Proto-Karen. *La Trobe Working Papers in Linguistics* 12: 26.

Ratanakul, S. (1998). Numeral classifiers in Sgaw Karen. *Mon-Khmer Studies* 28: 101-13.

Ratanakul, S. (2001). Adverbs in Sgaw Karen. *Mon-Khmer Studies: a Journal of Southeast Asian Linguistics and Languages* 31: 127-134.

Adverbs of Intensity in Sgaw Karen

Muna Nassri

The paper will investigate the use of functional adverbs in Karen discourse. Specifically, it will investigate intensity adverbs (more or less intensity) such as very, a lot, almost, and quite. It will first examine how these adverbs modify verbs and adjectives; for example, how an intensity adverb modifies an adjective in a sentence like 'I am very happy' versus how it modifies a verb in a sentence like "I went to school a lot". Furthermore, it will explore the position of these adverbs in sentences and how they are constructed. For example, some adverbs such as 'almost' are infixed in the words they modify whereas others follow or precede the word. It will also test these adverbs with sentences with negation.

1.0 Introduction

Adverbs of intensity are adverbs that modify verbs, adverbs, and adjectives and express "how much", "how intensely", or "to what extent". Adverbs of intensity exist in Sgaw Karen as well, and they are found to modify verbs, adjectives, and adverbs. There are adverbs that are intensive and adverbs that denote smallness or fewness. There also exist adverbs that are irregularly formed such as 'not at all' which are combined with negatives. The paper will also explore the use of quantifiers and how they are contrasted with intensity adverbs. Since the regular word order in Karen sentences is subject-verb-object, these adverbs usually follow the verbs or adjectives they modify. This paper will investigate the semantics of these adverbs and their syntactic structure in Karen phrases.

2.0 Inventory for adverbs of intensity in Karen

Adverbs of intensity in English fall on a scale expressing strong, medium, or weak intensity. There are numerous words that express intensity on these three scales:

Strong	**Medium**	**Weak**
extremely	rather	slightly
completely	pretty	a bit
totally	quite	hardly
absolutely	fairly	not at all
really		
very		
a lot		

In comparison with English, adverbs that express intensity in Karen are found in a much lesser extent and the inventory is not very big:

Strong	Medium	Weak
do^{5}ma^{4}	təs^{h}e^{4}	təs^{h}e^{4} p^{h}o
p^{h}a^{4}do^{5}	do^{5}ma^{4}	təba^{5}
ʔablə/tɔbo		ʃa^{6}blə
s^{h}u^{5}		ʔaʔa (ba)

The adverbs noted above consist of particles that are found to be redundantly combined with different particles to express meanings of intensity used in different contexts. For example, do^{5} is found in both words **do^{5}**ma^{4} and p^{h}a^{4}**do^{5}** which alone denotes the meaning 'big'. The same idea is found in the words **ʔa**blə and **ʔaʔa,** where the particle ʔa alone denotes the meaning 'very' or 'a lot'. It can also be combined with the negation particle ba^{5} to denote the meaning 'not very' or 'not a lot'.

2.1 Strong -intensity adverbs

2.1.1 [do^{5}ma^{4}]

do^{5} (big) + ma^{4} (very) – this adverb is the most commonly used adverb that denotes strong intensity. It generally modifies adjectives, verb, and adverbs. Although English has a different selection of words that express strong intensity, such as extremely, totally, completely, and too, Karen does not have a very big selection. All these words can be expressed in this single adverb in Karen.

2.1.2 [p^{h}a^{4}do^{5}]

p^{h}a^{4} (very) + do^{5} (big) – this adverb is found to modify adjectives. It gives the implication 'very big' and it can be used as an intensity adverb or as the adjective 'big'. When used as the adjective 'big', it cannot be modified by an intensity adverb.

2.1.3 [ʔablə]

ʔa (a lot) + blə (time) – this adverb is found to modify verbs. It gives the meaning 'many times' and it is used to express how often an action x occurs. It "intensifies" the verb by modifying its meaning to express that it happens a lot.

2.1.4 [tɔbo]

This adverb also modifies verbs and it expresses the meaning that an action x happens 'always'.

2.1.5 [s^{h}u^{5}]

This adverb modifies verbs and it implies the meaning 'a lot'. Similar to the adverb ʔablə, it is used to indicate how often the action occurs. However, this adverb has more limited uses and it can only be used with a few verbs such as cry, talk, walk, and eat. The main and more general use of this adverb is to imply quantity which will be discussed later.

2.2 Weak-intensity adverbs

2.2.1 [ʃa^{6}blə]

This adverb is found to modify verbs. It is the opposite of the adverb ʔablə and it denotes the meaning 'not very often' or 'a few times'. It gives the verb a "weak intensity" implication by expressing that action x happens a few times.

2.2.2 [təs^{h}e^{4} p^{h}o]

This adverb is also found to modify verbs. It conveys the meaning 'a little' or 'very little'. Similar to the adverb ʃa^{6}blə, it implies that action x happens a few times.

2.2.3 [təba^{5}]

This adverb denotes negation and it modifies adjectives, adverbs, and verbs. The first particle tə is the 'not' in the phrase, and it is always followed by the negation particle ba^{5} which comes at the end of the phrase. It also functions to negate sentences with the presence of an intensity adverb such as 'a lot' or 'very'. For example, this adverb can be compared in a phrase like "I do not eat" versus "I do not eat a lot".

2.2.4 [ʔaʔa ba^{5}]

This adverb is only used when combined with negation where ba^{5} is the negation particle. It implies the meaning 'not a lot'.

2.3 Medium-intensity adverbs

Unlike English, which has several adverbs denoting medium or moderate intensity, Karen does not have a specific inventory of words that express medium degree or intensity. For this category, there is an overlap between words expressing little intensity and words expressing strong intensity. That is, when testing medium-intensity adverbs used in English such as 'quite' or 'moderately', these words were expressed with the same adverbs used for strong or weak intensity adverbs in Karen. For example, when testing the adverb 'quite' in English, it was found that in some cases this adverb was expressed as do^{5}ma^{4} and in other cases it was expressed as təs^{h}e^{4} təs^{h}e^{4} (a little bit).

3.0 Quantifiers

Quantifiers are words that express quantity or amount of something. They are words that answer the questions "how many?" or "how much?". Quantifiers that express amount such as 'a few', 'a little', or 'a lot' are used to modify nouns and they are contrasted with adverbs that express frequency or intensity and which modify verbs:

3.1 [s^{h}u^{5}]

This adverb means 'a lot'. It is mainly used to describe the quantity or amount of something, but at the same time it can be used as a degree adverb to express 'how often' in the same sense as the adverb ʔablə. It can be combined with negation to negate that there is 'a lot' of something.

3.2 [p^{h}aʔa]

This adverb is also used to express the quantity or amount of something in a verbal sentence, and it gives the meaning 'a lot'. It can also be combined with the adverb of negation.

təs^{h}e^{4} (p^{h}o): even though this adverb is used to express intensity to give the meaning 'little/not very often', it is also used as a quantifier that expresses a 'little' of something in verbal sentences.

3.3 [bu^{4}kɨ li]

This quantifier modifies adjectives and it is translated as 'almost' in English. It is composed of three particles: bu^{4}kɨ which means "near to" and the particle li which usually comes at the end of the sentence being modified and it conveys the meaning 'near to end'.

4.0 Modifying adjectives

Unlike English where intensity adverbs precede the adjective, intensity adverbs in Karen usually follow the adjectives they modify.

4.1 Testing with strong-intensity adverbs

In a simple sentence with a linking verb containing a subject, adjective, and the intensity adverb, the sentence follows the order subject-adjective-adverb. Linking verbs such as 'am' or 'is' are not represented:

(1) "I am **very** happy":
jə θa^{4}kɨ **do^{5}ma^{4}**
1sg happy **very**
AGENT-ADJECTIVE-ADVERB

(2) "I am **too** tired":
jə lɔ^{6}bwi^{4} **do^{5}ma^{4}**
1sg tired **too**

Or**:**

(3) jə lɔ^{6}bwi^{4} **pa^{4}do^{5}**
1sg tired **too**
AGENT-ADJECTIVE-ADVERB

As example (3) shows, pa^{4}do^{5} is functioning as an intensity adverb.

(4) "It is extremely cold":
ta^{2}ɣo^{2} **do^{5}ma^{4}**
cold **extremely**
ADJECTIVE-ADVERB

A way to represent the linking verb in an attributive sentence is to be specific of what is being said by adding 'ʔawɛ ʔi^{6}' before the sentence where ʔi^{6} functions as a demonstrative that refers to the adjective that is modified:

(5) ʔawɛ ʔi^{6} ta^{2}ɣo^{2} **do^{5}ma^{4}**
this is cold **extremely**

(6) “It is a very cold day”:
təni^{6}ʔi^{6} ta^{2}ɣo^{2} **do^{5}ma^{4}**
day:this cold **very**

When a noun is present in the sentence and has a classifier, the syntactic order becomes noun-classifier (CLF)-adjective-intensity adverb:

(7) “the tree was really tall”:
θe^{5} tətu^{5} ʔi^{6} tɔ **do^{5}ma^{4}**
tree one-CLF this tall **really**
AGENT-CLASSIFIER-ADJECTIVE-ADVERB

(8) “that rose is really pretty”:
p^{h}ɔta^{6}ʃɨ^{5} təpɔ^{4} ʔi^{6} xwila **do^{5}ma^{4}**
rose one-CLF this pretty really-adv
AGENT-CLASSIFIER-ADJECTIVE-ADVERB

In a verbal sentence with an intensity adverb and an adjective, the syntactic order becomes subject-verb-object-adverb-adjective-adverb-classifier:

(9) “he drives a very big car”:
ʔawɛ pokwa nɔ θolɛ^{5} **p^{h}a^{4}do^{5}** tək^{h}o^{5}
2sg boy drive car very big one-CLF
AGENT-VERB-THEME-ADJECTIVE/ADVERB-CLASSIFIER

In example (9), p^{h}a^{4}do^{5} is used to express ‘very big’ and it is functioning as an intensity adverb and adjective at the same time. It appears that this adverb is always used with the adjective ‘big’.

When there are two adjectives present, the adjective/adverb p^{h}ado follows the first adjective and the order of the verbal sentence becomes subject-verb-noun-adjective 1-adjective 2/adverb-classifier:

(10) “I buy big yellow mangoes”:
jə pwe^{6} təkɔθa ʔabɔ **p^{h}a^{4}do^{5}** təplə
1sg buy mangoes yellow big one-CLF

4.1.1 Testing with negation

In a sentence containing the adverb for negation təba^{5}, the adjective is infixed between the two particles in the following order tə-adjective- ba^{5}, where tə is negation and ba^{5} is the negation particle:

(11) “I am not happy”:
jə tə θa^{4}k^{h}u ba^{5}
1sg NEG happy NEG

(12) “I am fairly happy”:

There is no word that expresses a “fairly” happy state in Karen. The state of happiness expressed is either very happy or not, so a medium intensity state cannot be expressed in this case. However, the adjective can be modified with negation combined with an intensity adverb, giving a state of “not very” happy. The adverb that is used with negation is ʔaʔa, where ʔa alone means ‘a lot’ and when combined with negation it denotes ‘not a lot’. Thus, the order of words in an attributive sentence containing negation and an intensity adverb is subject-NEG-adjective-adverb-NEG particle:

(13) “I am not **very** happy”:
jə tə θa^{4}k^{h}u **ʔaʔa** ba^{5}
1sg NEG happy not a lot NEG
AGENT-NEG-ADJECTIVE-ADVERB-NEG PARTICLE

As shown in example (13), it seems that the sentence contains double negation, since ʔaʔa by its own is the opposite of ‘a lot’, therefore ‘not a lot’ and at the same time it is combined with the negation particles.

Example (14) also demonstrates the case of “double negation”:

(14) “it is not too cold”:
ta^{2}tə ɣu^{2} **ʔaʔa** ba^{5}
NEG cold not a lot NEG

4.2 Testing with medium-intensity adverbs

When testing sentences with medium-intensity adverbs that modify adjectives such as ‘fairly’, ‘quite’, or ‘moderately’, the adverb used is təs^{h}e^{4} and it is duplicated to express the meaning ‘a little bit’. When doing one-to-one translation of medium-intensity adverbs in English, it appears that there are no words that express a “moderate” intensity state. The adverbs that are used instead express little or weak intensity which gives evidence that Karen’s adverbs of intensity can be expressed on two scales only; either “a lot” or “a little bit”:

(15) “it is moderately cold”:
ta^{2}ɣu^{2} **təs^{h}e^{4} təs^{h}e^{4}**
cold a little bit

4.2.1 Testing with two or more adjectives in attributive sentences

The syntactic order of the attributive sentences with one, two, or more adjectives present and with an intensity adverb is subject-adjective1-adjetive2-classifier-demonstrative-adjective3-adverb:

(16) “the student is quite young”:
pwa^{6} dʒo po təɣa^{6} ʔi^{6} ʔənjabo^{5} **təs^{h}e^{4} təs^{h}e^{4}**
people:school:people one-CLF this to look young a little bit
AGENT-CLASSIFIER-DEMONSTRATIVE-ADJECTIVE-ADVERB

(17) “the yellow table is quite heavy”:
ka^{4}kwɛ^{4}li^{3} ʔabɔ təbɛ^{5} ʔi^{6} xə **təs^{h}e^{4} təs^{h}e^{4}**
null:write:book yellow one-CLF this heavy a little bit
AGENT-ADJ-CLASS-DEM-ADJ2-ADVERB

(18) “the big yellow table is moderately heavy”:
ka^{4}kwɛ^{4}li^{3} ʔabɔ p^{h}a^{4}do^{5} təbɛ^{5} ʔi^{6} xə **təs^{h}e^{4} təs^{h}e^{4}**
null:write:book yellow big one-CLF this heavy a little bit
AGENT-ADJ1-ADJ2-CLASS-DEM-ADJ3-ADVERB

4.2.2 Testing with quantifiers

(19) “it was almost cold”:
ta^{2}**bu^{4}kɨ**ɣo^{2} li
QNF:cold near to end

(20) “I am almost happy”:
jə **bu^{4}kɨ**θa^{4}kɨ li
1sg QNF:happy near to end

In examples (19) and (20), the quantifier appears before the adjective it modifies and the quantifier particle ‘li’ appears after the adjective at the end of the phrase.

(21) “it was hot enough”:

The quantifier “enough” does not exist in Karen or is not really used.

5.0 Modifying verbs:

(22) “he likes you very much”:
ʔawɛ pokwa ba^{5}θa^{4} na^{6} **do^{5}ma^{4}**
2sg boy like 2sg very
AGENT-VERB-THEME-ADVERB

Other adverbs that modify verbs express “how often” the action occurs. They are contrasted with adverbs or quantifiers that express “amount”. The contrast is demonstrated in the following examples:

(23) “I went to school a lot”:
jə lɛ^{6} tɔ^{5}dʒo **tɔbo**
1sg go school always

Or

(24) jə lɛ^{6} tɔ^{5}dʒo **ʔablə**
1sg go school many times
AGENT-VERB-THEME-ADVERB

(25) "I eat a lot"
jə ʔɔ^{5} **ʔablə**
1sg eat many times
AGENT-VERB-ADVERB

(26) "I eat a lot of food":
jə ʔɔ^{5} ta^{2} **s^{h}u^{5}**
1sg eat food a lot
AGENT-VERB-THEME-QUANTIFIER

In example (25), the adverb ʔablə modifies the verb whereas in example (26), s^{h}u^{5} is a quantifier and it modifies the noun by expressing "how much".

(27) "I buy a lot":
jə pwe^{6}ta^{5} **s^{h}u^{5}/ʔablə**
1sg buy many times
AGENT-VERB-ADVERB

(28) "I buy a lot of food":
jə pwe^{6} ta^{2}ʔɔ^{5} **p^{h}a^{4}ʔa**
1sg buy unspecified food:eat a lot
AGENT-VERB-THEME-QUANTIFIER

As mentioned before, s^{h}u^{5} can be used as both a quantifier and an intensity adverb. In (27) it is used as an intensity adverb that modifies the verb. Again, its use is limited to certain adverbs such as cry, talk, walk, and eat:

(29) "I eat a lot of rice":
jə ʔɔ^{5} me^{6} **s^{h}u^{5}**
1sg eat rice a lot
AGENT-VERB-THEME-QUANTIFIER

(30) "I cry a lot":
jə hɔ^{5} **s^{h}u^{5}**
1sg cry a lot
AGENT-VERB-ADVERB

(31) "I speak a lot":
jə tɛta^{2} **s^{h}u^{5}**
1sg speak a lot
AGENT-VERB-ADVERB

5.1 Testing with weak-intensity adverbs and quantifiers

(32) “I eat very little”:
jə ʔɔ^{5} ta^{2} **təs^{h}e^{4}p^{h}o**
1sg eat unspecified food little
AGENT-VERB-ADVERB

Or:

(33) jə ʔɔ^{5} **ʃa^{6}blə**
1sg eat few times
AGENT-VERB-ADVERB

(34) “I eat very little rice”:
jə ʔɔ^{5} me **təs^{h}e^{4}p^{h}o**
1sg eat rice little

In examples (32) and (34), it appears that the adverb təs^{h}e^{4}p^{h}o expresses both “very little” or “little” in the same sense that p^{h}a^{4}do^{5} expresses “big and “very big”.

5.2 Testing with negation

Negation tə always precedes the verb it modifies, and the phrase is always ended with the negation particle ba^{5}.

(35) “I do not eat”:
jə tə ʔɔ^{5} ba^{5}
1sg NEG eat NEG
AGENT-NEG-VERB-NEG PARTICLE

(36) “I do not eat a lot”:
jə tə ʔɔ^{5} **ʔablə** ba^{5}
1sg NEG eat many times NEG
AGENT-NEG-VERB-ADVERB-NEG PARTICLE

(37) “I do not eat a lot of rice”:
jə tə ʔɔ^{5} me^{6} **p^{h}aʔa/s^{h}u^{5}** ba^{5}
1sg NEG eat rice a lot NEG
AGENT-NEG-VERB-QUANTIFIER-NEG PARTICLE

(38) “he hardly played”:
ʔawɛ po^{3}kwa lo^{3}kwɛ tə θe nɔ **təs^{h}e^{4}** ba^{5}
2sg boy play NEG can little NEG
AGENT-VERB-NEG-VERB-ADVERB-NEG

In example (38), saying “he hardly played” in Karen is the same as saying “he can play very little” or “he cannot play”.

6.0 Modifying Adverbs

When modifying adverbs, adverbs of intensity are again sentence-final.

(39) “he talks extremely fast”:
ʔawɛ po^{3}kwa kəto^{6}ta^{2} k^{h}lɛ **do^{5}ma^{4}**
2sg: boy talk fast extremely
AGENT-VERB-ADVERB-INTENSITY ADVERB

(40) “he eats quickly”:
ʔawɛ ʔɔ^{5}ta^{2} **k^{h}lɛ k^{h}lɛ**
2sg eat:unspecified food quickly
AGENT-VERB-ADVERB OF MANNER

When adding an intensity adverb to sentence (40), the ‘second’ k^{h}lɛ is removed and the intensity adverb is substituted:

(41) “he eats very quicky”:
ʔawɛ ʔɔ^{5}ta^{2} k^{h}lɛ **do^{5}ma^{4}**
2sg eat:unspecified food quick very
AGENT-VERB-ADVERB OF MANNER-INTENSITY ADVERB

Adding an object to the phrase:

(42) “he eats the apple very quickly”:
ʔawɛ ʔɔ^{5} p^{h}ɔθa^{5} k^{h}lɛ **do^{5}ma^{4}**
2sg eat apple quick very
AGENT-VERB-THEME-ADVERB OF MANNER-INTENSITY ADVERB

6.1 Testing with negation

(43) “he sings completely well”:
ʔawɛ θa^{4}wi^{5}ta^{2} ɣi^{6} **do^{5}ma^{4}**
2sg sing good completely
AGENT-VERB-ADVERB-INTENSITY ADVERB

Negation precedes the adverb of manner and the adverb of intensity:

(44) “he does not sing very well”:
ʔawɛ θa^{4}wi^{5}ta^{2} tə ɣi^{6} **ʔaʔa** ba^{5}
2sg sing NEG good very NEG
AGENT-VERB-NEG-ADVERB-INTENSITY ADVERB-NEG PARTICLE

(45) “it rained very heavily”:
ta^{2}su^{6} **do^{5}ma^{4}**
rain a lot/heavily

(46) “it rained quite heavily”:
ta^{2}su^{6} **təs^{h}e^{4} təs^{h}e^{4}**
rain a little

(47) “it did not rain very heavily”:
ta^{2} tə su^{6} **ʔaʔa** ba^{5}
NEG rain not a lot NEG

7.0 Conclusion

In conclusion, adverbs of intensity in Sgaw Karen are found to modify adjectives, verbs, and adverbs. The most frequently used adverb expressing strong intensity is do^{5}ma^{4}. Adverbs that express weak intensity include ʃa^{6}blə and təs^{h}e^{4} p^{h}o. To express negation, the adverb ʔaʔa is added after the adjective being modified accompanied by the negation particle ba^{5}.

Serial Verb Constructions in Sgaw Karen: A comparison of Karenic VP structures

Meredith Lucey-Weinhold

In many sentences in Sgaw Karen, more than one verb appears. This paper examines the different ways in which multiple verbs can interact in a sentence, such as the effects of word order and whether all the verbs are transitive. The historical development of multiple verbs is addressed: some complex modern verbs have been built from simpler structures, while some verbs have come to serve grammatical functions. These various complex verb structures in Sgaw Karen are compared to those in related languages to establish a better understanding of Karenic verbs.

1.0 Introduction

Serial verb constructions are characterized by multiple verbs appearing in a single clause and selecting nominal arguments as a unit, as in (1). While the primary focus of this paper is on serial verbs, I will also address other similar complex verb phrases.

(1) jə¹ kə¹ plɛ¹ lɔ⁶ ne² ke⁶ nə¹ ma¹
1.sg FUT let fall get return 2.sg wife
"I will release your wife."

In this paper, I investigate the structure of complex verb phrases in Sgaw Karen (Burmese dialect), drawing on other research done with the related Karenic languages Kayah Li and Pwo Karen. Earlier research in Sgaw Karen is also considered, providing a diachronic perspective. The details of structure in different types of serial verbs are examined, as well as the impact of word order on the meaning of a serial verb. Structures affecting the valency of the verbal complex are discussed. Evidence for diachronic changes in the status of verbal particles is addressed. Finally, I analyze the structure of complex serial verbs as a product of a hierarchical process.

1.2 Types of Verb Serialization

Many of the verbs in the data appear in both simple and serialized constructions. Some, such as [lɔ⁶] 'fall,' can only appear in serial constructions; other bound verbal morphemes are discussed later as resulting from grammaticalization.

Verbs of motion and action seem especially prone to serialization, as a way of refining the meaning of the verbal phrase (2-4).

(2) jə¹ ku⁵ lɔ⁶ θe⁵
1.sg cut fall tree
"I cut down a tree."

(3) jə¹ kə¹ hi⁴ kwi³ nə¹ ma¹
1.sg FUT take CAUS.move 2.sg wife
"I will take away your wife."

(4) pə¹jo⁶ ... ke⁶ twi² ʔɔ⁶ sʰu¹ ʔa¹ pu⁶
dragon return pull 3.sg to 3.sg hole
"The dragon pulled her back to his hole."

To establish categories of serial verbs, I draw on Solnit's research with Kayah Li, a related language. The above examples (2-4) are typical Resultative or Directional constructions, according to his classification (Solnit, 2006:149). Sequential constructions, where the serial verb indicates a series of actions (2006:152), appear as well (5). We also see some constructions in which the second verb mainly serves to introduce an argument to an intransitive verb (Solnit, 2006:153), where "separate" is added to the verb phrase so that it will select "table" as an argument, making the serial verb construction transitive (6). A similar construction to Solnit's 'verb + adjectival predicate' class (2006:154) also appears, in which the adjectival predicate modifies the entire clause (7). ([nɔ¹] is unable to appear without an object, and takes [θo¹le⁵] as a place-filler when no real-world referent exists.)

(5) ʔa¹wɛ¹ lɛ⁶ xɨ¹ ke⁶
3.sg go find return
"He went to find and bring back [his wife]."

(6) tʰi¹də¹ lɔ⁶ tɛ³ kʰa⁴kwe⁴li³
bottle fall separate table
"The bottle fell off the table."

(7) ʔa¹ nɔ¹ θo¹le⁵ kle¹
3.sg drive - car is-fast
"He drives (a car) very fast."

Some of the constructions Solnit lists (2006:149) have not surfaced in the current data, particularly 'Directives'. Additionally, while he analyses modal-like constructions as a type of serial verb in Kayah Li, the few modals encountered in Sgaw Karen have not required such an interpretation: the example in (8) is parallel to English modal structures.

(8) jə¹ nɔ¹ θo¹le⁵ kə¹ θe¹
1.sg drive car FUT can / be able
"I will be able to drive."

It is unknown whether these discrepancies represent a systematic difference between Kayah Li and Sgaw Karen, or whether they are due to an accidental gap in the data. Nevertheless, Solnit's terminology for different types of serial verb constructions will be used in this paper where applicable, namely 'Resultative', 'Directional', and 'Sequential'.

2.0 Word Order

Within a serial verb construction, the order of the words is fixed according to semantic factors. In Sequential serial verbs, the verbs appear in the order they are performed (9), in agreement with Solnit's findings in Kayah Li (2006:152). Changing the word order therefore changes the meaning of the sentence: (10) is ungrammatical not because of any syntactic violations, but because of the nonsensical meaning of buying a fish after having cooked it.

(9) jə¹ lɛ⁶ pwe⁶ pʰɔ¹ʔɔ⁵ ɲa⁵ tə¹ be⁵
1.sg go buy cook fish one CL
"I am going and buying and cooking a fish / I went and bought and cooked a fish."

(10) *jə¹ lɛ⁶ pʰɔ¹ʔɔ⁵ pwe⁶ ɲa⁵ tə¹ be⁵
1.sg go cook buy fish one CL

In Directional serial verbs, the word order determines which verb contributes the primary meaning to the sentence. In the following examples (11-14), the first verb carries most of the semantic weight, while the subsequent verbs serve to modify it. Inverting the order of the constituent verbs can effect a parallel change in meaning, provided the sentence remains grammatical (13-14).

(11) ʔa¹ ma¹ hɛ¹ ke⁶ tʰɔ⁵ lə¹ ta² pu⁶
3.sg wifecome return climb PREP NOM inside
"His wife came back out of the hole."

[ta²] is interpreted here as a nominalizer, turning the postposition [pu⁶] "inside" into a noun "hole", as "a place that one is inside".

(12) ʔa¹wɛ¹ lɛ⁶ jɔ² lɔ⁶ lə¹ hɔ⁵kʰo⁵ lo⁶
3.sg go step fall PREP ground surface
"She went and stepped down to the ground."

(13) jə¹ lɔ⁶ xʷe²
1.sg fall run
"I fall."

(14) jə¹ xʷe²lɔ⁶
1.sg run fall
"I run down."

Very few Resultative constructions, in which one of the actions is interpreted as causing or resulting in the other action, have surfaced. [kwi³] "throw" or "cause to move" seems fairly productive in this construction. It is therefore difficult to generalize about the word order of Resultatives, working with so little data, but in most cases the first verb is the action which results in the second verb. In (15-16), the first verbs "shout" and "slice" are actions performed in order to accomplish the second verb, "cause to move" and "get" respectively.

(15) hɔ^{1} kwi^{3}
shout CAUS.move
"[She] shooed / shouted in order to move [the snake]."

(16) ku^{4} ne^{2} wɛ^{1} ʔa^{1} θwi^{5}
slice get CONT 3.sg blood
"[He] got his blood by slicing."

While the precise significance of word order in serial verbs depends on the type of serialization, in the majority of cases it plays a role in determining the meaning of the serial verb construction.

3.0 Valency

Differing valency does not prevent verbs from appearing together in a serial verb, although variation exists in which arguments each verb selects. Above in (9), the Sequential construction combines intransitive [lɛ^{6}] "go" with the transtitives [pwe^{6}] "buy" and [pʰɔ^{1}ʔɔ^{5}] "cook". In the Directional construction at (17), "fall" is intransitive and selects "tree" as its sole argument; "cut" is transitive and selects "tree" as its object and "I" as its subject.

(17) jə^{1} klɔ^{1} lɔ^{6} θe^{5}
1.sg cut fall tree
"I cut down a tree / I cut a tree, the tree fell."

The reading of this sentence as "I cut a tree, and I fell," is not available, presumably for pragmatic reasons. Having multiple verbs in a single clause implies a relation between them, and the hypothetical reading would join two semantically unrelated concepts.

Reciprocal structures in Sgaw Karen are superficially similar to those in Pwo Karen (Kato, 2009:90), particularly in terms of phonetics: [lo^{3}] +[θa^{5}] and [lóθà], respectively. However, the reduction in valency which Kato found is not present in Sgaw Karen. In the following examples, there is no difference in transitivity between the simple and reciprocal constructions of a verb; reciprocals are always transitive, taking a pronoun as object.

(18) jə^{1} tɔ^{2} jə^{1} ɟɔ^{1}
1.sg hit 1.sg older brother
"I hit my older brother."

(19) pə^{1}wɛ^{1} tɔ^{2} lo^{3} pə^{1} θa^{5}
1.pl hit RECIP 1.pl RECIP
"We hit each other."

(20) ʔa^{1}wɛ^{1}θeʔɛ^{5} lo^{3} ʔa^{1} θa^{5}
3.pl loveRECIP 3.pl RECIP
"They love each other."

This is also true of reflexive structures: the object position is filled with a redundant pronoun (21), in sharp contrast to Pwo Karen intransitive reflexives (70).

(21) nə¹ tʰi⁵ lɔ⁶ nə¹ kə¹sa²
2.sg see REFL 2.sg REFL
"You see yourself."

(70) ʔəwê chè làɴ θà
3sg stab down REFL
'He stabbed himself.' (from Kato, 2009:87)

The verb for "do" or "make" is able to change valency in both Sgaw and Pwo Karen. Kato's example of "make-die" for "kill" (2009:76) is exactly paralleled in the current data. [θi^{1}] "die" is an intransitive verb, and [ma^{6}] "make" adds another argument to the structure (22). It also provides an *Agent* theta-role, causing the subject of the sentence to assume an active role in the death; thus the meaning of "kill" is built up.

(22) jə^{1} ma^{6} θi^{1} ɲa^{5} tə^{1} be^{5}
1.sg do / make die fish one CL
"I kill a fish."

The Sgaw Karen benefactive structure is similar, in that it adds an argument and a θ-role. Chappell's 1992 paper details the functions of the Sgaw Karen benefactive marker [ne²], which also exhibits parallels to Pwo Karen benefactive constructions (Kato, 2009:81). Introducing [ne²] into the verb phrase increases the valency by one argument, to which it assigns a *Benefactor* theta-role (Chappell, 1992:17). This analysis is perfectly borne out by the current data, in which the usually transitive [pwe^{6}] "buy" (23) becomes ditransitive with a *Benefactor* argument with the addition of [ne²] (24).

(23) jə¹ pwe^{6} ta¹kʰɔ⁴θa³ ʔa¹bɔ¹
1.sg buy mango is-yellow
"I buy yellow mangoes."

(24) jə¹ pwe^{6} ne² jə¹ mo² ta¹kʰɔ⁴θa³ tə¹ plə^{5}
1.sg buy BEN 1.sg mother mango one CL
"I buy a mango for my mother."

Based on the above results and other data, verb phrase valency increases far more commonly than decreasing in Sgaw Karen, consistent with Kato's generalization (2009:91).

4.0 Grammaticalization

Several verbal particles which have been treated thus far as constituents of serial verb constructions may in fact have undergone grammaticalization, and instead be non-lexical function words derived from verbs. Kato interprets many of the verbal particles of Pwo Karen in this manner (2009:75), and outlines diagnostic criteria for discriminating between true serial

verbs and structures which contain verbal particles (2009:76-77). The symptoms of grammaticalization observed in the current data are semantic bleaching and decategorialization (from Heine and Kuteva, 2002:2), indicated in the current data by the speaker's inability to define the meaning of a word and the use of the word in a non-lexical function, respectively.

Chappell finds the benefactive marker [ne²] has been partially grammaticalized from the word "get", though it still functions as a verb (1992:17). In the current data, [ne²] is still recognizable to a native speaker as meaning "get", and has been glossed as such in much of the data. However, it cannot function independently as a verb: [ne²] requires support from another verb [hi⁴] "take" to express the meaning of "get" (25). With [hi⁴] now dedicated to supporting [ne²], a new construction arises to express "take" (26). Finally, the meaning "take away" or "take and get" is expressed with the addition of [kwi³] (27).

(25) ʔa¹wɛ¹ hi⁴ ne² wɛ¹ lə¹ sa⁵pʰo¹kɔ²pʰo¹ tʰwi⁵ tə¹ di¹
3.sg take get CONT PREP animal dog one CL
"He got a dog."

(26) jə¹ kə¹ hi⁴ kwi³ nə¹ ma¹
1.sg FUT take CAUS.move 2.sg wife
"I will take (away) your wife."

(27) hi⁴ ne² kwi³
take get CAUS.move
"take away and get"

This apparent change in the degree of grammaticalization of [ne²] could be due to a dialectal difference in consultants. It could also be the result of a diachronic change, with the current data representing a further advancement of the grammaticalization of [ne²] to the point where it can no longer appear as an independent verb. The 1896 dictionary of Sgaw Karen provides support for [ne²] and [hi⁴] functioning as independent verbs (Cross, 1896:996, 1416). There is no evidence for a earlier construction meaning "take away", suggesting the latter is a more recent development created to fill the semantic gap left when [hi⁴] was devoted to supporting [ne²].

As indicated in previous examples, [kwi³] 'throw' has developed a non-lexical function in verb phrases, indicating the causation of movement (examples 3, 15, 26, 27). In contrast to the earlier data (Cross, 1896:333), it can no longer appear as an independent verb meaning "throw" (28), and must either be reduplicated or receive support from another word (29). (This is the only instance of [tə⁶] in the data; its precise function is unclear.)

(28) * jə¹kwi³plə¹ tə¹ plə⁵
1.sg throw ball one CL

(29) jə¹ {kwi³kwi³ / tə⁶ kwi³} plə¹ tə¹ plə⁵
1.sg throw.REDUP PRT throw ball one CL
"I throw a ball."

The extension of a word's use to metaphorical meanings is also indicative of grammaticalization (Solnit, 2006:155). [tʰɔ⁵] "climb" is particularly productive in this function, carrying a sense of creation or beginning (30).

(30) te¹ / bwa¹ / θu¹ tʰɔ⁵ sʰa⁴ tʰɔ⁵
build climb start climb
"to build" "start / start up"

ma⁶ kɛ¹ tʰɔ⁵ lo³ke¹ tʰɔ⁵
do invent climb transform climb
"invent / make up" "transform"

4.1 Lexicalization

Some serial verbs become lexicalized, as indicated by a native speaker interpreting them as single polysyllabic words; Solnit finds this only rarely in Kayah Li (2006:154). In the current data, the clearest instance of lexicalization is [kɨ²xo¹sɔ¹me⁵] "cremate" (31), in which the meanings of the constituent verbs are inaccessible to the speaker. Data from the 1896 dictionary further supports this being a development over time of a serial verb, as the contemporary word for "cremate" was [kɨ²xo¹]; [sɔ¹] "clean" and [me⁵] "fire" have since been added (Cross, 1896:235, 549, 1222).

(31) ʔa¹ ma¹ ne⁵ kɨ²xo¹sɔ¹me⁵ ke⁶ ʔa¹ wa⁶
3.sg wife TOP cremate return 3.sg husband
"His wife cremated her husband."

Two apparent lexicalizations, [ʔɔ⁵hi⁵] "steal" (32) and [pʰɔ¹ʔɔ⁵] "cook" (33), both appear to combine [ʔɔ⁵] "eat" with a second word. [hi⁵] historically appeared as an independent verb meaning "steal" (Cross, 1896:1420) but it could also be a serial verb with [ʔɔ⁵] "eat" (1896:1445), providing evidence that a serial verb once existed and later developed into the single word [ʔɔ⁵hi⁵]. The folk etymology for [ʔɔ⁵hi⁵] claims that stealing is only ever done in order to eat, suggesting that the lexicalization was recent enough that the individual meanings of the verbs are still accessible.

(32) ʔa¹wɛ¹ po³ kwa¹ ʔɔ⁵hi⁵ li³ tə¹ be⁵
3.sg child male steal book one CL
"The boy stole the book."

(33) jə¹ pʰɔ¹ʔɔ⁵ ɲa⁵ tə¹ be⁵
1.sg cook fish one CL
"I cooked a fish."

However, the word for "cook" was [pʰɔ¹ʔɔ⁵] already in 1896 (Cross, p. 1147), suggesting that is it either an even earlier lexicalization, coincidental homophony, or possibly a compound-like verb structure. Further, since the diachronic data available for Sgaw Karen is relatively

fragmentary and largely concerned with the more formal registers of the language, the claim of lexicalization for any phrase is at best a hypothesis.

5.0 Complex Serial Verbs

Thus far, the serial verb constructions under examination in this paper have primarily been only two verbs long. Serial verbs have surfaced which are up to four verbs long, built up hierarchically from the basic patterns of serial verb constructions. While not all of the permutations found in Kayah Li (Solnit, 2006:155-156) have been found in the current data, they still exhibit some similar constructions. Multiple Sequentials seem fairly productive, as in examples (9) and (34). The meaning depends on word order, indicating the order in which actions are carried out.

(34) ʔa¹wɛ¹ lɛ⁶ xɨ¹ ke⁶
3.sg go find return
"He went to find and bring back [his wife]."

Directionals are also capable of multiple serializations, as in (35). [hɛ¹] "come" supplies the basic meaning of the serial verb, with [ke⁶] and [tʰɔ⁵] refining its exact meaning.

(35) ʔa¹ ma¹ hɛ¹ ke⁶ tʰɔ⁵ lə¹ ta² pu⁶
3.sg wife come return climb PREP NOM inside
"His wife came back out of the hole."

"Release" (36) is a more complex structure, built on the Directional serialization [plɛ¹ lɔ⁶] "let fall". The benefactive marker [ne²] adds another argument to carry the theta-role of *Benefactor*, here the second-person singular [nə¹]. Finally, [ke⁶] "return" is related to the serial verb possibly as a Resultative, with the interpretation that "let fall" is the action which leads to the wife returning.

Incidentally, this is direct counter-evidence for Chappell's analysis that [ne²] is a prenominal clitic (1992:19), since [ke⁶] intervenes between [ne²] and the *Benefactor*-marked NP.

(36) jə¹ kə¹ plɛ¹ lɔ⁶ ne² ke⁶ nə¹ ma¹
1.sg FUT let fall BEN return 2.sg wife
“I will release your wife.”

The example in (37) can be divided into two dual-verb serializations: [si⁵ lɔ⁶] is a standard Directional meaning "run down", and [θi¹ kwi³] is a Resultative meaning "cause to die by throwing". These two serial verbs are then joined in a Sequential construction, with a final aggregate meaning of "run down, then cause to die by throwing".

(37) si⁵ lɔ⁶ θi¹ kwi³ ʔa¹ θa⁵ me⁵ʔu¹ pu⁶
run fall die CAUS.move 3.sg REFL fire inside
"[She] ran down and died by throwing herself in the fire."

6.0 Conclusion

Sgaw Karen serial verbs are fairly typical for the Karenic family. They exhibit evidence of a rigid underlying hierarchical structure for serial verbs in the effect of word order on the interpretation of meaning, and in the construction of complex serial verbs. In most complex verb phrases a single head verb exists, usually the first one, which carries most of the semantic weight. The language continues to change across time with the development of new grammatical particles and of new words created from serial verbs. The verb serialization process is cyclically productive, creating highly specialized verb meanings through repeated serialization. Many questions remain which the current data is incapable of answering, but the foundation of Sgaw Karen's complex verb structures is established.

References

Chappell, Hilary. (1992). The benefactive construction in Moulmein Sgaw Karen. *Linguistics of the Tibeto-Burman Area* 15 (1):11-30.

Cross, E.B. (1896). *A Dictionary of the Sgau Karen Language.* Rangoon: American Baptist Mission Press.

Heine, Bernd & Tania Kuteva. (2002). *World Lexicon of Grammaticalization.* Cambridge: Cambridge University Press.

Kato, Atsuhiko. (2009). Valence-changing particles in Pwo Karen. *Linguistics of the Tibeto-Burman Area.* 32 (2):71-102.

Solnit, David B. (2006). Verb serialization in Eastern Kayah Li. *Serial verb constructions: A Cross-linguistic typology,* eds. Alexandra Y. Aikhenvald & R. M. W. Dixon, 144-159. Oxford: Oxford University Press.

Subordination and Coordination in Karen Sgaw

Cassandra Chapman

Subordination and relative clause structure differ greatly among the languages of the world. Different strategies can be used to subordinate while some languages do not allow subordination whatsoever. Most European languages allow subordination, which can be found in common, everyday speech. These structures are actually preferred in these languages. Other languages allow, but do not prefer, subordination. Their strategies with respect to these syntactic constructions may therefore be highly restricted. Karen Sgaw, part of the Sino-Tibetan language family, is one of these languages.

There is limited previous research on subordination and relative clauses in Karen Sgaw. The few sources I found were not based on this dialect but instead on the Karen Pwo dialect. We have been told by our native speaker that these two dialects are not mutually intelligible. As such, I did not rely heavily on external sources. One source I did find was Thurgood and LaPolla's *The Sino-Tibetan Languages* (2007), which did discuss a complementizer in Karen Pwo. This complementizer was [lɤ^{1}], which is the same complementizer I will be arguing exists in Karen Sgaw. Due to the lack of sources on this topic, the majority of the analysis found in this paper will be based on my own linguistic insights.

In this paper, I will survey the use of subordination in the language. This will be introduced in Section 1. In Section 2, I will discuss complementizers in the language. Section 3 will analyze Karen Sgaw within wh-typology. In Section 4 and 5, I will survey relative clauses and their syntactic structure. In Section 6, I will discuss Keenan and Comrie's Noun Phrase Accessibility Hierarchy and argue that Karen Sgaw follows this hierarchy within relative clause structure. I will look at recursion in relative clauses in Section 7 and I will pay special attention to the topicalization mechanisms used within subordinated structures in Section 8. In Section 9, I will look at other methods of subordination in the language. Finally, in Section 10, I will conclude this paper within the Minimalist Framework. Subordination in Karen S`gaw will be seen as an option that is becoming more and more dialectal and individualized by native speakers.

1.0 Subordination in Karen Sgaw

Subordination seems to be disfavored in Karen Sgaw. Instead of subordination, conjunctions can be used to adjoin clauses together. The most commonly-used conjunction is [dɔ^{4}], which means "and." Here is an example:

(1) [nɤ^{1} kɤ^{1}ba^{5} kɤ^{1}no^{1} dɔ^{4} ja^{6} dɔ^{4} nɤ^{1} kɤ^{1}no^{1} dɔ^{4} ja^{6} me^{2} nɤ^{6} ne^{5} jɤ^{1} kɤ^{1}-
you have to wrestle with me and you wrestle with me COP win dis I will-

he^{5}- na^{6} lɤ^{1} jɤ^{1}-ta^{2}t^{h}u^{4}ta^{2}tɔ^{6} kɛ^{1} lɤ^{3}]
give-you to my-property all

"You have to wrestle with me and (if) you wrestle with me and you win, I will give

you all my property"

When we were told a Karen short story by a native speaker, subordination was not really used. Instead, simple sentences were adjoined together using conjunctions. The only subordination that used the complementizer [lɤ^{1}] was a quotative clause. This complementizer will be looked at in more detail in the next section. However, it is important to note that [dɔ^{4}] can also be seen as a discourse marker in the language. Nonetheless, it has the capacity to adjoin simple sentences together, such as the one listed above.

Contrary to our initial thoughts, subordination does exist in the language. Unfortunately, it was not used during the elicitation of the short story but our consultant did use such structures when I asked for specific sentences. I will go over complementizers in the next section then proceed to describe subordination environments and their restrictions.

2.0 Complementizers in Karen Sgaw

There is only one complementizer in Karen Sgaw. This morpheme seems to take on many functions. For instance, it can also be used as a preposition, even though it does not have a particular meaning. It is instead a more general preposition, taking on different meanings depending on the verb used. Here are some examples:

(2) [lɤ^{1} mɛ^{5}tʰi^{5}kɤ^{1}la^{1}]
inside mirror
"inside the mirror"

(3) [lɤ^{1} mɤ^{1}ha^{1}]
on yesterday
"yesterday"

(4) [lɛ^{6}-kɤ^{1}da^{1} lɤ^{1}-ʔa^{1}-kʰɨ^{4}
go-back to-his-farm
"go back to his farm"

(5) [lɤ^{1} hi^{5}pu^{6}]
in house-in
"inside the house"

(6) [lɤ^{1} ʔa^{1}tʰo^{4}]
at her-pig
"at her pig"

(7) [lɤ^{1} hi^{5}pu^{6}]
from house-inside
"from inside the house"

Its distribution as a preposition seems to vary and is highly dependent on the semantics of the verb used. Although an important distinction to make, this will not be discussed further in this paper.

We will instead discuss this morpheme's distribution as the language's complementizer. It is arguably the language's only complementizer. In any case, it is the only one that we have come across in our data. I argue that it is a complementizer due to its use with many verbs that take complementizers in most of the world's languages. Here are a few examples:

(8) [na^{3} lɤ^{1}]
believe that
"believe that"

(9) [pla^{1}lɔ ɣa (lɤ^{1})]
looks like that
"looks like that"

(10) [tʰi^{5} (lɤ^{1})]
see that
"sees that"

(11) [pla^{1} lɤ^{1}]
seems that
"seems that"

(12) [tɛ^{1} ʔa^{1} lɤ^{1}]
tell him that
"told him that"

(13) [si^{4} (lɤ^{1})]
say that
"said that"

This complementizer will be glossed as "that" for this paper. The motivation behind this will be discussed in the section on Relative Clauses (Section 4).

3.0 Wh-typology in Karen Sgaw

This section will look at wh-typology in Karen Sgaw. I decided to include this section in my analysis due to the fact that, in many languages, Complementizer Phrase (CP) structure is based on wh-questions. English and French are wh-movement languages, where the wh-phrase is said to move up to the specifier of CP. It is derived in the same syntactic position for interrogative structures as it would be for declarative structures. The movement of the wh-phrases also forces T to C movement (or do-insertion in English), which is characteristic of wh-movement languages. For instance:

(14) What$_i$ did you eat t_i?

(15) Where$_i$ did you go t_i?

(16) Où$_i$ es-tu allé t_i?

(17) Qu$_i$'avez-vous mangé t_i?

English also has wh-fronting in subordinate clauses. Here are a few examples:
(18) I wonder what$_i$ he bought t_i

(19) *I wonder he bought what

These sentences simply show the productivity of wh-fronting in languages like English. This is not the case for all languages of the world, however. Other languages, like Chinese, are considered wh-in-situ languages. The wh-phrase stays in the position where it was generated and does not move. Questions in Karen Sgaw follow this pattern. The language is therefore wh-in-situ. This can be seen by the following examples:

(20) [mɤ^{1}tɤ^{1}ɣa^6 ʔɔ^5 pʰɔ^{1}θa^5 lɛ^5]
who eat apple Q
"Who ate the apple?"

(21) [nɤ^1 tʰi^5 mɤ^{1}tɤ^{1}ɣa^6 lɛ^5]
2sg see who Q
"Who did you see?"

(22) [nɤ^1 ʔɔ^5 pʰɔ^{1}θa^5 pɛ^{1}lɛ^5 tɤ^{1}plɤ^5 lɛ^5]
2sg eat apple which one-round Q
"Which apple did you eat?"

(23) [nɤ^1 ʔɔ^5 pʰɔ^{1}θa^5 pʰɛ^1 lɛ^5]
2sg eat apple where Q
"Where did you eat the apple?"

(24) [nɤ^1 tʰi^5 di^1 lɛ^5]
2sg see how Q
"How did you see?"

Although the language has wh-words meaning *who, what, where, how,* and *which*, these words remain in-situ in interrogative sentences. A Q(uestion)-particle, [lɛ^5], is inserted to show that the sentence is interrogative. The wh-phrase does not move in these constructions but remains in the position where it was generated. Questions, therefore, do not tell us anything about the CP structure in Karen Sgaw. They do, however, indicate that CP structure may not be as rich in Karen as it is in English. We will look at this in more detail in the following section on Relative Clauses.

On a side note, it is important to note that the Q-particle for non-wh-questions is different. Instead of [lɛ^5], [ʔa^1], is used. This can be seen more clearly in the following example:

(25) [nɤ^{1} ʔɔ^{5} pʰɔ^{1}θa^{5} θe ʔa^{1}]
2sg eat apple able Q
"Can you eat the apple?"

3.1 Control/ ECM constructions

Control verbs in English are verbs that take an infinitival complement. The subject of the matrix clause is said to also be the subject of the infinitival clause. As infinitives cannot assign Case, PRO is inserted to satisfy the EPP of the lower clause. Here are some examples:

(44) I_i want PRO_i to go.

(45) I_i promise PRO_i to leave early.

ECM constructions in English are similar to control verbs in that they also take an infinitival complement. However, they also take an overt DP in the infinitival clause. As infinitives cannot assign Case, the main verb is said to exceptionally Case mark the lower subject (with accusative case). Here are some English examples:

(46) I-NOM want John-ACC to go to the store.

(47) I-NOM believe John-ACC to be a believer.

What is interesting about these verbs with respect to CP structure is that control verbs are said to take complementizer phrases as their complement while ECM verbs are said to take TPs (tense phrases) as their complements. This is to prevent Case from being assigned in Control structures. It also allows Case to be assigned in ECM constructions. Furthermore, all of these constructions can merge an overt complementizer before the subordinate clause. This forces the verb to become finite and all Case problems are solved. Here are some examples:

(48) I want that John come with me.

(49) I promise that I will go.

(50) I believe that John is a believer.

There are many control- and ECM-like verbs in Karen. They all, however, take the complementizer [lɤ^{1}]. The majority of the time, this complementizer cannot be omitted:

(51) [jɤ^{1} ʔɤ^{5}lɔ^{6} lɤ^{1} jɤ^{1} kɤ^{1} lɛ^{6}-tʰɔ^{5} ta^{2}plɔ^{1} ʔa^{1}mu^{4} ne^{5}lɔ^{6}]
1sg promise that 1sg fut go-up wedding event SFP
"I promise that I will attend the wedding"

(52) [jɤ^{1} ba^{1}θa^{1} lɤ^{1} ʔa^{1}wɛ^{1} pʰo^{1}kwa^{1} kɤ^{1} hɛ^{1} dɔ^{4} ja^{6}]
1sg want that 3sg boy will come with 1sg
"I want that he will come with me"

The subject of the embedded verb cannot be omitted. This will result in ungrammaticality. We could therefore argue that there is a strong EPP (Extended Projection Principle) feature in this

language. All specifiers of TP must be filled with an overt NP. Control and ECM verbs do not seem permitted in the language. Instead, a simple CP structure is used with a finite phrase. This satisfies both Case requirements and the EPP.

3.2 Raising

Interestingly, raising verbs in Karen are all introduced by a complementizer clause:

(53) [pla (*lɤ^{1}) John ʔa^{1}wɛ^{1} ne^{5} θa^{4}kʰɨ^{1}]
seem that John he T happy
"It seems that John is happy" OR "John seems to be happy"

(54) [plalɔɣa (lɤ^{1}) John ne^{5} θa^{4}kʰɨ^{1}]
look like that John T happy
"It looks like John is happy" OR "John seems to be happy"

The curious thing about these phrases is that an NP does not need to precede the matrix verb. As Karen does not have expletives, the spec TP position in the tree must remain empty for these kinds of verbs. In these sentences, John, cannot move higher up in the structure. He must remain in the position where he is generated. He cannot raise as in English.

4.0 Relative Clauses

Relative clauses exist in many of the world's languages. They are useful in predicting syntactic distribution within a given language. To begin, it is important that we define the notion of a relative clause (RC). The most comprehensive definition of a relative clause is given in Lehmann (1986). He states that a relative construction is "a construction consisting of a nominal … (which may be empty) and a subordinate clause interpreted as attributively modifying the nominal. The nominal is called the head and the subordinate clause the RC. The attributive relation between head and RC is such that the head is involved in what is stated in the clause" (664). It is within this definition that we will discuss relative clauses in Karen Sgaw.

This section will discuss the typology of relative clauses and where Karen Sgaw fits into this typology. Although Karen Sgaw uses wh-phrases in interrogatives, they are unable to be used in relative clauses. Hence, there are no wh-relative clauses:

(24) [*po^{1}mɨ^{5}pʰo^{1} tɤ^{1}ɣa^{6} mɤ^{1}tɤ^{1}ɣa^{6} lɤ^{1} jɤ^{1} θe^{5}ɲa^{1} ne^{5} ʔa^{6}θa^{5}
girl one-person who-one-person C 1sg know T her-heart

xa^{6} do^{5}ma^{4}]
good very

"The girl who I know has a very good heart."

(25) [*ta^{2}lɔ^{2} lɤ^{1} pʰɛ^{1}lɛ^{5} jɤ^{1} ʔɔ^{5} ta^{2} ne^{5} xa^{6} do^{5}ma^{4} lɔ^{6}]
place that where 1sg eat place that good very SFP
"The place where I eat is very good"

Instead, the language allows that-relatives:

(26) [po^{1}mɨ^{5}p^{h}o^{1} tɤ^{1}ɣa^{6} lɤ^{1} jɤ^{1} θe^{5}ɲa^{1} ne^{5} ʔa^{6}θa^{5} xa^{6} do^{5}ma^{4}]
girl one-person that 1sg know T her-heart good very
"The girl who I know has a very good heart."

(27) [ta^{2}lɔ^{2} lɤ^{1} jɤ^{1} ʔɔ^{5} ta^{2} ne^{5} xa^{6} do^{5}ma^{4} lɔ^{6}]
place that 1sg eat place that good very SFP
"The place where I eat is very good"

For this reason, we will gloss all instances of [lɤ^{1}] as "that." As [lɤ^{1}] can be interpreted as a general preposition in its other meanings, perhaps it could also be interpreted as a general complementizer. It is always a functional projection and so, perhaps we could also see it as a general FP (Functional Phrase). More data would be required to support this idea, however.

4.1 Other Methods of Subordination

Of course, relative clauses are not the only way to subordinate in Karen. As previously argued, conjunction is the preferred option to subordination. Here are a few of these conjunctions:[dɔ^{4}] "and"

[p^{h}ɛ^{1}ne^{5}] "then"
[p^{h}ɛ^{1}] "when"
[ʔa^{1}xo^{1}] "so"

The most significant of these is [p^{h}ɛ^{1}], "when," which could be considered to be a wh-word in Karen. As such, it should not be able to subordinate a phrase. The following examples show that this is not the case:

(42) [ʔa^{1}wɛ^{1} k^{h}i^{6}lɛ^{1} ʔa^{1}hi^{5} p^{h}ɛ^{1} nɤ^{1} ʔɔ^{5} ha^{1}ta^{2} ʔa^{5} wi^{6}ʔa^{1}ka^{1} lɔ^{6}]
3sg go his-home when 2sg eat dinner 3sg finish SFP
"He will go home when you have finished eating dinner"

(43) [jɤ^{1} ʔo^{5} p^{h}ɛ^{1}ne^{5} p^{h}ɛ^{1} da^{2}ke^{1}ta^{5}θa^{4} ʔa^{1}ka^{1}]
1sg stay there when happen that-period
"I was there when it happened (in the period that it happened")

Curiously, this is the only wh-phrase permitted in subordinate clauses. If you attempt to insert any other wh-phrase in these types of positions, the sentence was completely ungrammatical for our consultant. Nimrod always told me that the sentence could be remedied by inserting the complementizer. I would argue that the use of "when" implies a specific period in time for the speaker and is therefore able to have a different distribution than all other wh-phrases.

5.0 Syntax of Relative Clauses in Karen Sgaw

In Karen Sgaw, all relative clauses are post-nominal. The RC follows the noun phrase which it modifies. There are no examples of pre-nominal RCs in the language. As outlined in Lehmann (1986), RCs in the languages of the world can take on several main properties. He argues that the RC can either be adjoined to the main clause or embedded inside of it (666). Karen Sgaw is parallel to English in this respect. The RC is always embedded inside of the main clause. It modifies an NP, forming a sole constituent with that NP. It is thus part of the NP. In this way, the RC is subjected to the same syntactic properties and restrictions as the nominal to which it is adjoined. The difference is that the RC can be omitted, without changing the syntactic structure of the phrase. The RC provides an provides an optional was of adding information in the language, specifically to the NP. Like in English, relative clauses are optional in Karen Sgaw. The information presented in these types of sentences could be removed and inserted into a following sentence. The use of RCs enables a speaker to use one sentence, instead of two. The language does not use a lot of adjectives. Perhaps this could explain the lack of relative clauses used in common speech.

As far as the syntax of relative clauses go in Karen, they tend to have an overt nominal head. One will seldom find a headless relative. In fact, a lot of the time, the nominal head is repeated. In these types of constructions, the RC is both head-internal and head-external. Here are a few examples:

(28) [po^{1}mɨ^{5} ʔa^{1}wɛ^{1} ne^{5} me^{2} po^{1}mɨ^{5} lɤ^{1} nɤ^{1} θe^{5}ɲa^{1} pɛ^{1}lɛ^{5} tɤ^{1}ɣa^{6} lɛ^{5}]
girl she Topic is girl that 2sg know which one-person Q
"The girl, who is she?"

(29) [p^{h}o^{1}kwa^{1} lɤ^{1} jɤ^{1} t^{h}i^{5} lɤ^{1}mɤ^{1}ha^{1} tɤ^{1}ɣa^{6} ne^{5} ʔa^{1}wɛ^{1} xa^{6} do^{5}ma^{4}]
boy that 1sg see yesterday one-person Topic 3sg good very
"The man I saw him yesterday, he is a good person"

In both of the above phrases, the Topic marker follows the phrase that is being topicalized. The noun phrase must then be repeated within the following sentence. This may also be due to restrictions on the verb. Subject pro-drop does not seem to occur in this language. Noun phrases do not seem to be able to remain empty. The speaker must clarify their referent in a relative clause. This could also be a restriction on the topic marker, which will be discussed in a later section. If we remove the bolded NPs from the examples, the sentences are ungrammatical. This demonstrates that headless relatives are not permitted.

6.0 Noun Phrase Accessibility

The hypothesis of Noun Phrase Accessibility was first developed by Keenan and Comrie (1977). Lehmann (1986) modified this hypothesis in order to explain relative clauses in languages of the world. He argues that certain noun phrases are easier to relativize than others, according to the following hierarchy:

1) Subject/absolutive
2) Direct object/ergative

3) Indirect object
4) Other complements
5) Adjuncts

(Lehmann 1986: 668)

Noun Phrases found lower in the hierarchy may be relativizable but are harder for native speakers to parse when compared to those Noun Phrases found higher in the hierarchy. In this way, the hierarchy provides us with implicatures. If a language is able to relativize adjuncts, it is also able to relativize all other types of Noun Phrases. In English, all of these types of noun phrases can be relativized. Karen Sgaw seems to also follow this hierarchy quite closely. We will take each of these types of Noun Phrases one by one.

6.1 Subject

Subjects can be relativized in the language. This is always done using a that-clause.

(30) [pʰo^{1}kwa^{1} lɤ^{1} jɤ^{1} tʰi^{5} lɤ^{1}mɤ^{1}ha^{1} tɤ^{1}ɣa^{6} ne^{5} ʔa^{1}wɛ^{1} xa^{6} do^{5}ma^{4}]
boy that 1sg see yesterday one-person T 3sg good very
"The man that I saw yesterday is a good person"

6.2 Direct object

Direct objects are harder to relativize for native speakers of Karen Sgaw.

(31) [… be^{1} lɤ^{1} ʔa^{1} hɛ^{1} lɤ^{1} kɔ^{2}-ʔe^{1}kɤ^{1}li^{4} ...]
beans that 3sg come from country-English
"... beans that come from England ... "

It was preferable for the speaker to change this into an attributive phrase such as "English beans" or "fried beans." The relativization of subject noun phrases therefore seems to be more accessible to Karen speakers due to the difficulty the speaker had producing sentences like (31) compared to sentences like (30).

6.3 Indirect object

Relativizing indirect objects is only a little more difficult but is still possible, as in the following example:

(32) [pʰo^{1}kwa^{1} lɤ^{1} jɤ^{1} tɛ^{1} te^{2} ʔɔ^{6}]
man that 1sg say past to him
"The man to whom I spoke" literally "The man that I spoke to him"

6.4 Other complements

Karen Sgaw does not seem to have any other sorts of complements so we could argue that this position could be excluded from the hierarchy.

6.5 Adjuncts

As there are few adjuncts in Karen Sgaw, it is highly unlikely that they would be relativizable. As we have seen in the course, Karen does not like having adjuncts in the left periphery and those found in the right periphery are highly restricted. Therefore, I see no way for these types of clauses to be part of a relative clause.

Karen Sgaw, therefore, seems to follow this hierarchy quite closely. It is easiest to relativize subjects, followed by direct then indirect objects. The rest of the hierarchy is not allowed in the language. This is predicted by Keenan and Comrie's hypothesis. Due to the acceptability of the first three types of relative clauses, it is odd that these types of constructions did not come up naturally when a short story was elicited by the native speaker. Either the speaker simply prefers to not use relative clauses (personal preference, avoidance of difficult structures) or they are actually not common in habitual discourse. I will leave this topic for another time. It is beyond the scope of this paper.

7.0 Recursion in Relative Clauses

Recursion is very productive in English relative clauses. Here are a few typical examples:

(34) The man who John met who I know is a very good cook. (subject RC recursion)

(35) The food that Mary brought that we ate last night was great. (object RC recursion)

Although recursion in relative clauses is not as natural, it is still possible in Karen Sgaw. Following the Noun Accessibility Hierarchy, the easiest of these is subject recursion:

(36) [po^{1}mɨ^{5}pʰo^{1} lɤ^{1} jɤ^{1} tʰi^{5} lɤ^{1}mɤ^{1}ha^{1}tɤ^{1}ni^{6} ne^{5} lɤ^{1} (*ʔa^{1}) me^{2} bwa^{6}ʤo^{1}pʰo^{1}
girl that 1sg see yesterday T that she is student

ne^{5} ko^{1}to^{6}ta^{2} dɔ^{4} ja^{6} lɔ^{6}]
T talk to 1sg SFP

"The girl that I saw yesterday that is a student talked to me."

Recursion of the object of the relative clause is also possible, although this is not preferred. Instead of embedding the direct object with a relative clause, the speaker preferred to use an attributive structure, such as "English beans" or "fried beans." In the end, I was able to elicit a recursive structure. Here is an example:

(37) [bwa^{6}pʰo^{3}kwa^{1} lɤ^{1} ʔa^{1}-ba^{5}θa^{4} be^{1} lɤ^{1} (*ʔa^{1}) hɛ^{1} lɤ^{1} kɔ^{2}-ʔe^{1}kɤ^{1}li^{4}
man that 3sg-like beans that 3sg come from country-English

ne^{5} hɛ^{1} pʰɛ^{1}ʔi^{6} lɔ^{6}]
T come here SFP

"The man that likes beans that come from England is here"

The above examples demonstrate that recursion within RCs in Karen Sgaw is possible. The complementizer is able to embed sentences twice. In both cases, the relative clause cannot be headless. The NP must be repeated after the second complementizer. Without this referent, the sentence would be considered ungrammatical.

8.0 Topicalization in Subordinate/Relative Clauses

The morpheme [ne^5] has been a hard one to define throughout the course. Like the complementizer, it may fulfill several functions. However, I would like to argue that it is exclusively a TOPIC marker in relative clauses. It will have the feature [+topic].

This marker comes up frequently within relative clauses (RC) in Karen. However, within subordinate and RCs, there seems to be some restrictions on its distribution. Syntactically, the topic marker tends to appear before the verb (or tense/aspect marker, if applicable). In attributive sentences, it appears before the adjective. As it is typically used to topicalize a noun phrase, it frequently follows a noun phrase.

I will argue that the use of the topic marker implies that everything before it is the topic and is therefore, important. This morpheme is used to categorize what is important to the speaker. Here are a few examples:

(39) [pla^{1}lɔ^{6}ɣa^3 lɤ^1 John (ne^5) θa^{4}kʰɨ^1]
seems that John T happy
"It seems that John is happy"

John is important. It is important to the speaker that it is John who is happy.

(40) [John tɛ^1 ʔa^1 lɤ^1 ʔa^{1}wɛ^1 (ne^5) kɤ^1 hɛ^1]
John tell him that he T will come
"John told him that he would be coming"

The fact that *John told him* that he would be coming is important. The predicate is less important in this clause. The matrix clause has more value than the subordinate clause.

In both of the previous sentences, the marker appears before the predicate. The predicate is least important when the topic marker is used by the speaker. As shown in my transcriptions, the topic marker in the above sentences is optional. It denotes some sort of speaker preference. Although it is not mandatory in these types of constructions, it seems to be most natural. Most of the sentences elicited already contained the topic marker. I had to ask whether or not it was optional.

Furthermore, in recursive structures, the topic marker can appear twice:

(41) [po^{1}mɨ^{5}pʰo^1 lɤ^1 jɤ^1 tʰi^5 lɤ^{1}mɤ^{1}ha^{1}tɤ^{1}ni^6 ne^5 lɤ^1 ʔa^{1}me^2 bwa^{6}ʤɔ^{1}pʰo^1 ne^5
girl that 1sg see yesterday T that she-is student T

ko^{1}to^{6}ta^2 dɔ^4 ja^6 lɔ^6
talk to 1sg SFP

"The girl that I saw yesterday that is a student talked to me."

In both instances, the topic is the same. Perhaps the topic marker reinforces the subject of the sentence. It forces the fact that the relative clause is modifying the subject.

9.0 Conclusion

This paper has surveyed the use of subordination in Karen Sgaw. More specifically, the language was analyzed with respect to its relative clause structure. I compared the use of relative clauses in Karen Sgaw to their use in English. Keenan and Comrie's Noun Accessibility Hierarchy was used to explain the acceptability of certain RCs over others. Although subordination is not preferred in the language, it is more than acceptable. Arguably, it could be a part of everyday, common discourse. However, for some speakers, it may be easier to break up a more complex sentence into two smaller sentences adjoined by a conjunction. In this way, subordination is disfavored by native speakers. This provides us with minimal phrase structure and lack of any feature violation.

The aim of this paper was to provide some insight into the Complementizer Phrase structure in Karen Sgaw. I have tried to see the CP as complex but this does not seem probable. The CP is used as a last resort in the language. It ensures that EPP and Case Checking are not violated. It also provides us with certain economy conditions. It is more economical to merge the CP than it would be to Move DPs to a higher clause. Unlike English, Karen Sgaw embraces these notions, providing us with an elegant way of looking at CP structure in the language. Within the Minimalist Framework, the language provides us with both an economical and elegant phrase structure.

References

Keenan, E. & B. Comrie (1977). Noun phrase accessibility and universal grammar. Linguistic Inquiry. 8: 63-99.

Lehmann, C. (1986). On the typology of relative clauses. Linguistics. 24 (4): 663-680.

Thurgood, G. & R.J. LaPolla (2007). The Sino-Tibetan Languages. London: Routledge.

The [ne^5] particle and its role in a topic-prominent interpretation of Karen Sgaw

Hugo St-Amant Lamy

There are two main tendencies in how information is structured in the different languages of the world. In English (and most European languages), a typical sentence is composed of a subject and a predicate. The predicate is grammatically linked to the subject and usually modify it or describe its state. A typical subject-predicate sentence could be "I eat an apple". In many south eastern languages, we can find another construction : a combination of a topic and a comment. The comment gives information concerning the topic, but it is not grammatically linked to it. Such a sentence could be "The apple, I eat it". In this paper, I try to support the idea that Karen Sgaw, a Tibeto-Burman language, uses the topic-comment structure. To achieve this goal, I report a number of features that are caracteristic of topic-prominent languages. By doing so, I explore many aspects of the Karen syntax such as verbal voice, word order, coreference control, impersonnal verbs, etc. The main finding is that Karen Sgaw is not a pure topic-prominent, but it is closer to those languages than to subject-predicate languages.

1.0 Introduction

Even a superficial glance at Karen Sgaw discourse reveals the unusually high frequency of a particle. This frequent particle (it is found in nearly half the clauses) is pronounced [ne^5]. Its grammatical function is somewhat unclear and it is found in a number of seemingly heterogeneous environments. As if this was not confusing enough, it can also be omitted in most cases without affecting the meaning of the clause (according to the informant).

This paper has two goals. It first attempts to give a descriptive overview of the particle: the environments where it is found, those where it seems necessary, its role in such processes as topicalisation, etc. The second goal of the paper is to define the function of [ne^5] and what it can teach us about the typology of Karen Sgaw, more precisely about the relevancy of a topic-prominent approach to the language.

1.1 Warning

The following analysis is based on a very small corpus and is therefore inevitably incomplete. In addition, data was gathered during elicitation sessions (translation of short sentences by the informant and his judgments on their grammaticality) and during the recitation of a folk tale. These forms of language do not necessarily correspond to the language spoken naturally between two native speakers of Karen. Since the subject of topicalisation is relatively close to pragmatic, this could have a significant impact on the results. We must therefore consider the assumptions made in this work with a pinch of salt and interpret the findings as paths to follow as more data becomes available.

2.0 Descriptive overview of [ne^5].

The particle [ne^5] is ubiquitous in Sgaw Karen. Indeed, the short folk tale told by the speaker comprises about 60 clauses and we can find [ne^5] in about 45 of them. It is therefore apparent that the particle has an important role in the language. Before trying to define the function it performs, it is important to see where it can be found and where it seems to be necessary.

2.1 Possible environments

The following list shows environments where [ne^5] can be found. In cases studied in this section, the particle is optional and its presence or absence has no significant effect on semantics (according to the judgments of the speaker).

First, it is possible to find [ne^5] immediately after an adverbial phrase (in the data collected, it is always a time adverbial phrase located early in the proposal).

'One morning, when her husband goes to the field...'
[tə^1- ɣɔ^6 ne^5 p^{h}ɛ^1 wa^6 lɛ^6 lə^1 k^{h}ɯ^{4}pu^6]
one morning (class.) ? when husband go PREP. field

'Everyday, he works...'
[ko^{4}ni^{6}dɛ^4 ne^5 ʔa^{1}wɛ^1 ma^6 wɛ^1]
everyday ? 3P.S. work PROG.

In a similar context, (locating the temporality of action), it can also be found after a relation marker early in the proposal.

'Then, he says to his wife...'
[p^{h}ɛ^4 ne^5 ʔa^{1}wɛ^1 te^{2}xa^3 ʔa^1- ma^1]
then ? 3P.S. say 3PS POSS. wife

'And later, her comb is in her hair...'
[dɔ^4 lə^1 ki^1 ne^5 ʔa^1- θi^5 ʔo^5 lə^1 ʔa^1- k^{h}o^{5}lo^6]
and PREP. later ? 3PS POSS. comb stay/is PREP. 3PS POSS. hair

One of the environments where [ne^5] is most frequent is after the grammatical subject of the clause. However, this subjet must be nominal, pronominal subjects can't be followed by [ne^5].

'And the dragon did not say anything to him'
[dɔ^4 pə^{1}jo^6 ne^5 tə^1- te^2 ʔɔ^6 ta^{2}nɔ^{1}tə^{1}ni^6 ba^6]
and dragon ? NEG. Say 3PS nothing NEG.

'His wife stays...'
[ʔa- ma ne ʔo wɛ]
3PS POSS. wife ? stay PROG.

'I eat' *'I eat'*
*** [jə^1 ne^5 ʔɔ^5] vs. [jə^1 ʔɔ^5]**

1PS ? eat 1PS eat

The direct and indirect objects can be also followed by [ne^5]. It can be interesting to note that the pronominal direct objects (accusative case) can precede [ne^5], contrary to the pronominal subjects. However, it seems that [ne^5] cannot appear at the very end of a sentence. If the sentence continues after the object (eg.: when the first clause is follwed by another), then it seems [ne^5] can appear without any problem.

Nominal direct object
'The man eat the dog...'
[pwa^6 po^{3}kwa^1 ʔɔ^5 twi^5 ne^5 ...]
person man eat dog ?

Pronominal direct object
'I see you...]
[jə^1 t^{h}i$^?$ na^6 ne^5...]
1PS see 2PS ACC. ?

Nominal indirect object
'You give a mango to Hugo...'
[nə^1 he$^?$ tə^{1}k^{h}ɔ^{4}θa^5 tə^1- plə^5 s^{h}u^1 hu^{1}go^2 ne^5...]
2PS give mango one fruit (class.) PREP. Hugo ?

Pronominal indirect object
'You give me a mango...'
[nə^1 he$^?$ tə^{1}k^{h}ɔ^{4}θa^5 tə^1- plə^5 s^{h}u^1 ja^6 ne^5...]
2PS give mango one fruit(class.) PREP. 1PS ACC. ?

In conditionnal constructions, the condition or hypothesis clause can be followed by [ne^5]. It is the only case where this particle can be found after a verb.

'If you fight with me and don't win, I will take away your wife'
[nə^1 kə^{1}no^1 dɔ^4 ja^6 me^2 tə^1 nə^1 ne^5
pron. 2ps fight and pron. 1ps is (NEG) win ?
jə^1 kə^1 hi^4 kwi^3 nə^1- ma^1]
1ps. (FUT.) take throw away 2ps POSS. wife

Finally, [ne^5] can also appear after the possessor in genitive constructions.

'Kunole's son is dead
[kɯ^{5}nɔ^{2}le^1 ne^5 ʔa^1- p^{h}o^{1}kwa^1 θi^1]
Kunole ? 3PS POSS. son dead

There seems to be a restriction to the realisations of [ne^5] that have just been presented. All sentences ending with the particle have been judged agrammatical by the informant. [ne^5] can not be used in that position. However, as we can see in the following exemple, is the clause ending in [ne^5] is coordinated to another clause, the sentence is judged grammatical.

'I see you.'
*[**jə^{1}** **t^{h}i$^{?}$** **na^{6}** **ne^{5}**]
1PS see 2PS ACC. ?

'I see you and I am happy'
[**jə^{1}** **t^{h}i$^{?}$** **na^{6}** **ne^{5}** **dɔ^{4}** **jə^{1}** **θa^{4}kɯ^{1}**]
1PS see 2PS ACC. ? and 1PS happy

2.1.2 [ne^{5}lɔ^{6}]

In addition to the previous cases, it is possible to find [ne^{5}] in a somewhat different position. Quite frequently (in twelve of the seventy proposals of the tale), the end of a sentence is marked by the word [ne^{5}lɔ^{6}]. While the speaker could not easily express the meaning of [ne^{5}] in other environments, he is quite definitive on the fact that [ne^{5}lɔ^{6}] is a simple particle marking the end of the sentence. It does not have any effect on semantics. Moreover, there seems to be no restriction on what can precede the particle (for example, one can find a verb before [ne^{5}lɔ^{6}], which is normally impossible in front of [ne^{5}]). This specific sentence-ending realization of the particle [ne^{5}] seems to be different from the ones seen previously. It will then be treated as a different particle altogether and will not be discussed in more detail in this work.

'He starts going'
[**ʔa^{1}wɛ^{1}** **sa^{1}tɔ^{1}** **lɛ^{6}we^{1}** **ne^{5}lo^{6}**]
3PS start go fin de phrase

2.1.3 A single [ne^{5}] or many homophones?

To summarize, [ne^{5}] may be added after six distinct elements: time adverbial phrases, relation markers, subjects, objects (direct and indirect), conditional clauses and possessors in genitive constructions.

Looking at this wide array of possible environments, one might hypothesize that we are dealing with different homophones that only have their surface form in common. However, for reasons of economy, we will try to treat [ne^{5}] as a single grammatical unit as long as it remains relevant. If no explanation can take into account all the different environments, then the homophones hypothesis will be explored.

2.2 Mandatory use of [ne^{5}]

In previous cases, [ne^{5}] was optional and did not seem to have a significant impact on semantics (that the speaker was able to explain anyway). However, there are two constructions where the presence of the particle is essential.

First, karen allows topicalization of an element by moving it to the beginning of the sentence. This process is very similar to what is available in english (eg.: 'Your shoes, they look great'). In this case, the moved element has to be followed by [ne^{5}]. Also, at the original location of the item being moved, we find a trace in the form of a pronoun (except in the case of the direct object where this trace is not necessary). Several elements may be topicalized: subjects, direct

objects, indirect objects and the genitive portion of subject noun phrases. Pronominal subjects can be topicalized, but in their accusative forms.

Subjet
'The fisherman give a book to his wife'
[pwa^{6}ma^{6}ŋa^{5} ne^{5} ʔa^{1}wa^{1} he^{ʔ} li^{3} tə^{1}- be^{5} s^{h}u^{1} ʔa^{1}- ma^{6}]

fisherman ? 3PS give book 1 flat(class.) PREP 3PSposs. wife

Accusative pronominal subject
'I give a book to my wife'
[ja^{6} ne^{5} jə^{1} he^{ʔ} li^{3} tə^{1}- be^{5} s^{h}u^{1} jə^{1}- ma^{6}]
1PS ? 1PS give book 1 flat(class.) PREP 1PSposs. wife

Direct object
'The fisherman give a book to his wife'
[li^{3} tə^{1}- be^{5} ne^{5} pwa^{6}ma^{6}ŋa^{5} he^{ʔ} s^{h}u^{1} ʔa^{1}- ma^{6}]
book 1 flat(class.) ? fisherman give PREP 1PSposs. wife

indirect object
'The fisherman gives a book to his wife'
[ʔa^{1}- ma^{1} ne^{5} pwa^{6}ma^{6}ŋa^{5} he^{ʔ} li^{3} tə^{1}- be^{5} shu^{1} ʔa^{1}ʔo^{6}]
1PSposs. wife ? fisherman give book 1 flat(class) PREP 3PS

genitive
'The fisherman's son gave a book to his wife'
[pwa^{6}ma^{6}ŋa^{5} ne^{5} ʔa^{1}- p^{h}o^{1}kwa^{1} he^{ʔ} li^{3} tə^{1}- be^{5} s^{h}u^{1} ʔa^{1}- ma^{6}]
fisherman ? 3PSposs son give book 1 flat(class) PREP 3PSposs wife

The second structure where [ne^{5}] absolutely must be present is the structure of subordination. At first glance, [ne^{5}] should be placed after the grammatical subject of the subordinate clause.

'It seems that John is happy'
[pla^{1} lə^{1} John ne^{5} θa^{4}kɯ^{1}]
seem CONJ. John ? happy

However, in other subordinate clauses, we can see that the mandatory [ne^{5}] is instead located after the topicalized element.

'He knows that *his wife, the dragon pulled her inside his hole.'*
[ʔa^{1}wɛ^{1} θe^{5}ŋa^{1} wɛ^{1} lə^{1} ʔa^{1}- ma^{1} ne^{5} pə^{1}jo^{6}
pron. 3ps know (prog.) conj. sub. 3ps wife ? dragon
ke^{6}twi^{1} nɯ^{3} ʔɔ^{6} s^{h}u^{1} ta^{1}pu^{6} ne^{5} lɔ^{6}]
pull inside 3ps to hole end of sentence

3.0 [ne^5] as a topic marker

There seems to be a very close link between topicalization and the particle [ne^5] studied here. According to the Concise Oxford Dictionary of Linguistics, the topic is what the proposal is about, the rest of the proposal being a comment on this topic. The precise function of the theme is well defined by Chafe (1976):

> *What the topics appear to do is limit the applicability to a certain restricted domain. […] The topic sets a spatial, temporal, or individual framework within which the main predication holds.*

If [ne^5] must follow the topicalized element at the beginning of a sentence, would it be possible that it plays a similar role wherever it is found, a role of topic marker? Supporting this hypothesis is the fact that the six types of elements that [ne^5] may follow can hold the function of topic as defined above. The time adverbial markers and the relation markers all define the time frame in wich the proposition takes place. The arguments of the verb (subject and complements) define the personal frame. Finally, the conditional clause can define the general framework in which the clause that follows unfolds.

Since it specifies the frame of the clause, the topic has to be definite (Li & Thompson 1976: 461). This feature might explain why the nominal case pronouns (subject position) can not be topic of their clause (followed by the marker [ne^5]). They are probably not definite enough. However, their accusative counterparts, perhaps more explicitly referential, can be topics.

If [ne^5] is really a topic marker, then topic-comment constructions are much more frequent in Karen than say, French or English. A little more than half of the sentences in the speaker's tale would contain such a structure. What could this say about the typology of the karen language?

4.0 Topic and subject?

All languages where topicalization was studied showed at least some use of topic-comment structures. One can even find them in English and French, where the structure of a standard phrase is subject-predicate.

[Moi]$^{\text{thème}}$, [je considère qu'il est un génie]$^{\text{commentaire}}$
[When it comes to music]$^{\text{thème}}$, [I suck]$^{\text{commentaire}}$

While the topic-comment structure is rare in European languages, it plays a basic role in several languages of East Asia (Japanese, Mandarin, Korean, etc.). Even in the Tibeto-Burman family, which includes Sgaw Karen, we can find some languages where the structure is almost omnipresent (eg.: Lahu and Lisu of the Lolo-Burmese family). It would be interesting to see how important the topic-comment structure is in Karen.

4.1 A word on typology

In an important article on the subject, Li & Thompson (1976) classifies the world's languages into four distinct categories. First, we find subject-prominent (SP) languages , where a basic sentence consists of the grammatical relation between a subject and a predicate. The Indo-

European languages fall into this category. Then, there are topic-prominent (TP) languages. Unlike SP languages, they focus on the relationship between the topic and a commentary on this topic. Lahu and Lisu are examples of this category of languages. There are also languages where the two types of structures are used. This typology applies to, among others, Japanese. Finally, in some other languages (such as Tagalog), the subject and the topic are merged and can not be distinguished.

4.2 Caracteristics of topic-prominent languages

The high frequency of the particle [ne^{5}] leads us to believe that the topic-comment structure is very important in Karen. In the pages that follow, we'll see if it could be defined as a topic-prominent language. To do this, we will rely on a list of eight characteristics of topic-prominent languages established by Li & Thompson (1976) and see how they apply to karen.

i. Surface coding
ii. Absence of "dummy" subjects
iii. Minimal use of passive constructions
iv. Subject/topic cooccurrence (so called "double subjects")
v. Coreference control
vi. Verb-final word order
vii. Absence of constraints on topic constituents
viii. Basicness of topic-comment sentences

4.2.1 Surface coding

The first characteristic can be achieved in two ways: by the position of the subject in the sentence (eg.: through anteposition) or with a morphological marker. In Karen, movement to the beginning of the sentence plays a role in topicalization, but it is really the presence of the morpheme [ne^{5}], after the topic, that is crucial for encoding. This is probably the feature that is most explicitly in Sgaw Karen.

4.2.2 Absence of ''dummy subjects''

In subject-prominent languages, the function of the grammatical subject is very important and its position must always be occupied, even by a semantically empty element (a "dummy" subject) if necessary. This phenomenon is found in proposals like *it rains*, where the impersonal *it* refers to nothing specific. In topic-prominent languages, the subject position is only accessory and so expletive subjects are not found.

This feature does not seem to apply perfectly to Karen. For meteorological verbs, while it is possible to find a specific subject, it is also possible to find what appears to be an expletive subject, the particle [ta^{2}].

'The sky rains'		*'It rains'*	
[mɯ^{1}ko^{5}	**s^{h}u^{6}]**	**[ta^{2}**	**s^{h}u^{6}]**
sky	rain	INDEF.	rain

'Le wind blows' *'It winds'*
[kə^{1}li^{6} ʔu^{1}] **[ta^{2} ʔu^{1}]**
wind blow INDEF. blow

The fact that these verbs do not always have an empty subject is interesting. They can hardly be characterized as traditional expletive verbs. Thus, [ta^{2}] is perhaps not quite as empty semantically as we might think, it could simply be highly indefinite (like *something* in English). It must still be conceded that the sentences using [ta^{2}] as a subject can not be regarded as following the canonical topic-comment structure since this particle is too indefinite to be a topic.

There are, however, proposals without any grammatical subject. This suggests that the function is only optional and can remain empty . The fact that we do not find [ta^{2}] as the subject here supports the idea that it is a kind of undefined agent. Since the verb does not select an agent in that sentence, we can not find [ta^{2}] as a subject.

'It seems that Kunole will win'
[lɔ^{6}ɣa^{3} lə^{1} kɯ^{5}no^{2}le^{1} kə^{1} nə^{6}]
seem CONJ. Kunole FUT win

Finally, it seems that Karen generally tries to avoid the kind of expletive sentences where an indefinite pronoun is used in French or English. For example, when asked to translate the sentence t*here is a concert tonight*, the speaker produced the following sentence:

'Some people show a concert tonight'
[tə^{1}na^{6}ʔi^{6} pwa^{5} kə^{1} pja^{4} kɔ^{1}sə^{4}]
night people FUT show concert

Basically, it is false to suggest that Karen has no referent-free subjects, since we find [ta^{2}] in some cases. However, its usage is rare and proposals with empty subject position are perfectly grammatical. This feature (the tendency to avoid "dummy subjects"), applies at least partialy to Karen.

4.2.3 Minimal use of passive constructions

In topic-prominent languages, the theta roles of the verb's arguments do not have as much importance as in subject-prominent languages. For this reason, constructions based on the modification or reversal of these roles (such as passive constructions) are quite rare.

Karen does not use of this type of construction very much, but there are still clauses that have at least the appearance of a passive construction.

'I break the table'
[jə^{1} ma^{6}ka^{2}tu^{3} ka^{4}kwɛ^{4}li^{3}]
1PS break table

'The table is broken by me'
[ka^{4}kwɛ^{4}li^{3} ne^{5} ʔa^{1}wɛ^{1} ka^{2}tu^{3} lə^{1} jə^{1}ho^{1}]
table TOP. 3PS break PREP 1PS

The second sentence seems to be a typical passive sentence. However, comparison with the first sentence gives us a clue to the contrary. In the first example, the verb break is [ma^{6}ka^{2}tu^{3}] but we find only [ka^{2}tu^{3}] in the second one. According to the speaker, the morpheme [ma^{6}] has the meaning of **action** or **work** (eg.: [pwa^{6}ma^{6}ɲa^{5}] - fisherman). It is therefore likely that the *break* verb is really [ma^{6}ka^{2}tu^{3}] while [ka^{2}tu^{3}] stands for the word *broken*. In this case, the sentence *the table is broken* is no more passive than *the table is red*. The sentence merely describes the state of the table, with the addition of a complement to indicate who caused this state.

However, the difference between the verb of the "active" sentence is not always different from the one int the "passive"sentence.

'The boy steals the book'
[ʔa^{1}wɛ^{1} po^{3}kwa^{1} ʔo^{5}hu^{5} li^{1} tə^{1}- be^{5}]
3PS boy steal book one flat(class)

'The book is stolen by the boy'
[li^{1} tə^{1}- be^{5} ne^{5} ba^{ʔ} ta^{2} ʔo^{5}hu^{5} lə^{1} ʔa^{1}wɛ^{1} po^{3}kwa^{1}]
book 1 flat(class) TOP ? INDEF steal PREP 3PS boy

Here, the same verb is found in the two sentences. The passive meaning appears to be expressed by [ba^{5}ta^{2}]. The speaker had a hard time explaining the meaning of [ba^{5}], but he says the [ta^{2}] is the same that is used with meteorological verbs. If [ta^{2}] (*something*) is the grammatical subject of [ʔo^{5}hu^{5}] in the second example, then the sentence is not passive anymore. [li^{3} tə^{1}be^{5}] would be the topic, as noted by the presence of [ne^{5}], and we would find an active sentence with an indefinite subject (*the book, someone steals it, the boy*).

But this strategy of creating a passive equivalent (since in the elicitation sessions, we are somewhat "forcing"the speaker to produce passive sentences) does not work with all verbs. For example, it is impossible (according to the informant) to produce a passive version of the next sentence.

'I see you'
[jə^{5} t^{h}i^{ʔ} na^{6}]
1PS see 2PS

'You are seen by me'
***[na^{6} ne^{5} ba^{ʔ} ta^{2} t^{h}i^{ʔ} lə^{1} jə^{1}ho^{1}]**
2PS TOP ? INDEF see PREP 1PS

As one would expect from a language where topic is very important, Karen does not really use passive forms. When we asked the speaker to directly translate passive sentences, he can sometimes create something equivalent, but nothing where there is a reversal of theta roles. Therefore, this "minimal use of passive constructions" feature seems to apply well to that language.

4.2.4 Subject/topic cooccurrence (so called "double subjects")

One of the easiest ways to locate a topic-comment structure is to find a sentence where both the subject and the subject arise. The co-occurrence of these elements is a characteristic of topic-prominent languages. There is a type of sentences in which subject and topic coexist that seems exclusive to TP languages: the so-called "double subjects" clauses. In these sentences, the topic is obviously not selected by the verb (not subject or object) and its thematic function is thus highlighted. There are a few examples of this type of sentences karen.

'Kunole, the feet are cold'
[kɯ^{5}no^{2}le^{1} ne^{5} kɔ$^{?}$ kɯ$^{?}$ do^{5}ma^{4}]
Kunole TOP foot cold very

Here, we do not seem to be dealing with a simple genitive structure where the possessor would have been topicalized. Indeed, we do not find [ʔa^{1}kɔ$^{?}$] (***his*** *feet*), but only [kɔ$^{?}$]. The frame of the clause is therefore [kɯ^{5}no^{2}le^{1}], and the commentary on this subject is [kɔ$^{?}$ kɯ$^{?}$ do^{5}ma^{4}] (*the feet are very cold*).

However, while this kind of structure is almost ubiquitous in purely topic-prominent language, Karen seems to have a more moderate amount. Still, there is a large number of sentences where the topic coexists with the subject but do not correspond to an argument by the verb. This is the case of topics which are adverbial markers, relation markers or conditional proposals. We can therefore say that this "double subject" feature corresponds to Sgaw Karen.

4.2.5 Coreference control

In subject-prominent languages, the subject controls coreference. Thus, in a sentence like *Robert gave a hat to Hubert, he then left*, the pronoun *he* is understood as referring to Robert. However, in a topic-prominent language, coreference is controlled by the topic, to the detriment of the subject. This phenomenon is found in Karen as in the following sentence:

'This tree, the leaf is very big so I love it'
[θe^{5} ʔa^{1}wɛ^{1} ne^{5} la^{2}pa^{3} do^{5} do^{5}ma^{4} ʔa^{1}ho^{1} jə^{1} ba^{ʔ}θa^{ʔ} ʔɔ^{6}]
tree 3PS TOP leaf big very so 1PS love 3PS

According to the informant, this sentence would never be interpreted by a native speaker as meaning *I like the leaf*. The pronoun necessarily refers to the topic, not the subject. We can therefore conclude that the control of coreference feature applies well to Karen.

4.2.6 Verb-final word order

Topic-prominent languages tend to have a SOV structure. This is the case, among others, with Japanese and Korean, as well as Lisu and Lahu, two Burmese-Lolo languages. It is clear that Karen, with its SVO structure, does not meet this requirement.

However, as noted by Li & Thompson (1976), Mandarin Chinese, is considered at least partially TP and it has a clear SVO structure. It seems that this attribute is more often found in pure TP languages, wich Karen is apparently not.

4.2.7 Absence of constraints on topic constituents

As we have already seen, the topic-comment structure is found to some extent in all languages, even in mainly subject-prominent languages. However, what may be considered a valid topic in these languages is very limited. Usually, only the subject and its genitive can fulfill this function.

Topic-prominent languages offer a much more diverse range of topics. In the most extreme cases of TP languages, any element of the sentence can potentially be topicalized (Li & Thompson 1976: 471).

In the case of Karen Sgaw, several elements may be topicalizes. Here are some examples to illustrate this diversity.

Adverbial phrase
'Everyday, Kunole gives an apple to Nomue'
[ko^{4}ni^{6}de^{4} ne^{5} kɯ^{5}nɔ^{2}le^{1} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1}]
Everyday TOP Kunole give apple 1 round(class) PREP Nomue

Relation marker
'Then, Kunole gives an apple to Nomue'
[p^{h}ɛ^{1} ne^{5} kɯ^{5}nɔ^{2}le^{1} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1}]
Then TOP Kunole give apple 1 round(class) PREP Nomue

Sunjet
'Kunole gives an apple to Nomue'
[kɯ^{5}nɔ^{2}le^{1} ne^{5} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1}]
Kunole TOP give apple 1 round(class) PREP Nomue

Subject's genitive
'Kunole, is son gives an apple to Nomue'
[kɯ^{5}nɔ^{2}le^{1} ne^{5} ʔa^{1}- p^{h}o^{1}kwa^{1} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1}]
Kunole TOP 3PSposs son give apple 1 round(class) PREP Nomue

Direct object
'Kunole gives an apple to Nomue'
[p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} ne^{5} kɯ^{5}nɔ^{2}le^{1} he^{1} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1}]
apple 1 rond(class) TOP Kunole give PREP Nomue

Indirect object
'Kunole gives an apple to Nomue'
[nɔ^{2}mɯ^{5}ʔe^{1} ne^{5} kɯ^{5}nɔ^{2}le^{1} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} ʔa^{1}ʔo^{6}]
Nomue TOP Kunole give apple 1 round(class) PREP 3PS

Conditionnal phrase
'If Kunole gives an apple to Nomue, she is happy'
[kɯ^{5}nɔ^{2}le^{1} he^{1} p^{h}ɔ^{2}θa^{5} tə^{1}- ple^{5} s^{h}u^{1} nɔ^{2}mɯ^{5}ʔe^{1} ne^{5} ʔa^{1}wa^{1} θa^{4}kɯ^{1}]
Kunole give apple 1 round(class) PREP Nomue TOP 3PS happy

Basically, the restrictions on the subject are not very strict in Karen. Most elements of the sentence can play this role. Therefore, it is safe to say this feature applies at least to some extent here.

4.2.8 Basicness of topic-comment sentences

Obviously, one of the most striking features of topic-prominent language is the basicness of topic-comment structures. In other words, the topic-comment construction is less marked in these languages. This structure is present in the repertoire of basic sentences of these languages, which is not the case with subject prominent languages.

As we have seen, there is a topic-comment construction in more than half the clauses we found in the story told by our informant. If we consider that it is a tale, so somewhat aimed at children, we can hypothesize that this type of construction is basic in the language. This characteristic therefore seems to apply well to Sgaw Karen.

5.0 Conclusion

In light of these findings, we can conclude that Sgaw Karen shares many characteristics with the topic-prominent languages. At least six of the eight criteria established by Li & Thompson (1976) can be found in the language. Obviously, some features are only partially present (such as the lack of "dummy" subjects) and others are entirely absent (SOV order). Moreover, we have to keep in mind that most of these features are not binary, but may be realized to various degrees (eg.: basicness of topic-comment structure).

Finally, we can say with confidence that Karen Sgaw is not a TP language as "pure" as Lisu or Lahu. On the other hand, it is clearly not a subject-prominent language, such as French or English.

5.1 Subject-topic continuum

Indeed, the classification of Li & Thompson (1976) into four language categories is a bit misleading. The boundaries between these categories are not completely clearcut, so it is more appropriate to speak of a continuum. Available data suggest that the Karen language lies between the pure TP languages and pure SP languages. However, many features of the TP languages can be observed in Karen. This suggests it is closer to the TP end of the continuum. According to Hagège (1978), this continuum is the synchronic manifestation of a diachronic process: a cycle in which the topic is slowly grammaticalized until it becomes a subject, then returns to a topical function as the syntactic restrictions get looser. If this assumption is valid, it would be interesting to compare our data with future or past data to see where Karen Sgaw can be positionned in this topic-subject cycle.

5.2 Relevance of this topic in the study of field work

The issue of Karen as a topic-prominent language can teach an important lesson in the study of field work. It reminds us to be extremely careful not to project the features of our own language on the studied language. Here, Karen was instinctively investigated as a subject-predicate language because we (the "researchers") speak primarily languages that have this

typological characteristic. This bias partly explains why it took almost four months before beginning to consider that topicalisation could have an important role in language. By asking the speaker to give rulings on the subject or translate types of sentences that are not necessarily found in Karen (eg.: passive sentences), it is dangerous to force the language studied in a typological mold that would be consistent with our expectations.

It is therefore very important to be familiar with the various linguistic features of the languages of the world and not to begin research by looking for certain features at the expense of others. In conclusion, we must always keep an open mind, even to what is, according to us, the most exotic constructions.

References

Chafe, W. (1976). 'Giveness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View.' Dans C. Li (ed.) *Subject and Topic.* London, New York: Academic Press, 25-55.

Fuller, J. (1987). 'Topic Markers in Hmong.' *Linguistics of the Tibeto-Burman Area 10* (2), 113-127.

Gundel, J. & T. Fretheim. (2004). 'Topic and Focus.' Dans L. Horn & G Ward (eds.) *The Handbook of Pragmatic Theory.* Oxford: Blackwell Publishing, 174-196.

Hagège, C. (1978). 'Du thème au thème en passant par le sujet. Pour une théorie cyclique.' *La Linguistique 14* (2), 3-38.

Li, C. & S. Thompson. (1976). 'Subject and Topic: A New Typology of Language.' Dans C. Li (ed.) *Subject and Topic.* London, New York: Academic Press, 457-489.

Lindvall, A. (2005). 'Information structure in some East Asian languages or Is there a thing called subject?' *Lund Working Papers in Linguistics 51,* 125-142.

Matisoff, J. (1991). 'Areal and universal dimensions of grammatization in Lahu.' Dans E. Traugott & B.Heine (eds.) *Approaches to Grammaticalization vol. II.* John Benjamins Publishing, 383– 453.

Analysis of the particle Ne[5] in Sgaw Karen: A Minimalist Approach to Pragmatics

Nahed Mourad

This paper has for goal to identify the functions of the particle [ne[5]] in Sgaw Karen not only for the purpose of identification but rather for prediction of its overt instances. A minimalist approach to pragmatics has been adopted to include all instances of overt ne[5] as a topicalizatino marker. Data from a native Sgaw Karen speaker has been elicited over the period of 3 months for 3 hours per week.

Our data showed functions where the particle attaches to a lexical item as a specifying determiner . However, instances where [ne$_5$] is postposed to a clause were of particular interest to me. Although it may be marked for languages to consider determiner specifiers to be postposed to the clause, it is not utterly inconceivable as per the literature on Fon and Haitian Creole (Larson) where a similar distinct definiteness marker is present. Conversley, in this particular usage, it appears to be more consice and consistent for the language to employ this particle as a topicalizer, noting that definiteness markers and topicalizers are pragmatically similar.

1.0 Background on Karen

The Karen State, a State of Burma (Myanmar) bordering Thailand, has been in rebellion against the Military government of Burma since the 1950s. This political situation has sent thousands upon thousands of Karen people to refuge to the camps of Thailand to escape danger, poverty etc. Karen, a language belonging to the Sino-Tibetan family, more specifically, Tibeto-Burman languages, is spoken in parts of South East Asia. Karen has three dialects Pa'O, Pwo, and Sgaw and is "spoken along a north-south axis roughly coinciding with the Thailand-Burma border." (Thurgood & LaPolla, 2003, p. 623) Due to the political situation in Burma, and their constant contact with Thai, and Burmese, Karen dialects have much influence from these languages.

Our informant, a native speaker of Karen Sgaw has offered us, through data elicitation, a preliminary view of the phonological, morphological, syntactic, semantic and pragmatic features of this dialect.

2.0 Focus of the paper

This paper has for goal to identify the functions of the particle [ne^5] in Sgaw Karen not only for the purpose of identification but rather for prediction of its overt instances. This particle serves many functions in the language, such as terminating clauses (and therefore a final particle of some sort), modifying pronouns, nouns, adjectives and classifiers into a more definite and specific form. Interestingly enough it was omnipresent in the recounting of a folk tale but rather disfavored in previous data elicitation of simple sentences.

Firstly we will address the issue of attempting to find an umbrella function for this particle that will allow for consistency and more systematic predictions.

So far, all the functions where the particle attaches to a lexical item can be encompassed by a specifying determiner category. The only inconsistency would be when [ne5] attaches to a clause to mark its ending. Although it may be marked for languages to consider determiner specifiers to be postposed to the clause, it is not utterly inconceivable as per the literature on Fon and Haitian Creole (Larson) where a similar distinct definiteness marker is present. Here, we can draw on previous literature on the correlation between specificity and topicalization. Yon-Mee Park , in their paper on "Topicalization in Terms of Specificity'', claims that "topicalization provides a piece of evidence for the existence of the syntactic structure where NPs can best be termed as specific" (Park, 1996). Ergo from this point forward we will treat the "ne^5" particle as a topicalization marker.

3.0 Literature review

The specific particle we will be focusing on has yet to be discussed in the literature I have been reviewing; however, it has been glossed in a few articles. In 'Grammar of Sgaw Karen' a book created by the missionary of the American Baptist Mission in order to describe and notate the language use. The ne^5 construction is descriptively analyzed every instance in which it is prevalent. Here, all the cases seen in the data elicitation (demonstrative marker, definite marker, final particle (as in ne5lɔ)). Note that all the functions noted in this literature is purely prescriptive and descriptive linguistic value must be inferred. According to the grammar compilation,

a) Ne^5 seems to denote the ending of a constituent (NP, AdjP, AdvP and VP). Why it is overt in some place and not in others remains to be considered.

(1) **ʔa^1-wa^6-ne^5** [6]
Her-husband – TOP
"Her husband"

b) Ne^5 seems to also attach itself to classifiers giving the constituent (usually the NP) a more specific connotation.

[6] Note that the examples given are from the consultant data elicitation rather than from the literature on sga grammar

(2) **dɔ^{4} kɤlɤ tɤbone^{5}**
and python one-class flat – TOP
"The python" The ne here serves the purpose of a determiner (definite article)

c) Ne^{5} seems to attach to ʔawɛ to denote the meaning of a definite demonstrative adjective (That)

(3) **John$_{i}$ tɛ^{1} ʔa^{1} lɤ^{0} ʔa^{1}w ɛ^{1} $_{i}$ ne^{5} k ɤ^{0} h ɛ^{1}**
John tell him that he – TOP FUT- come
"John told him that he would be be coming"

d) It is also postposed to time markers such as [p^{h}ɛ] to make more complementizer-like lexical items. Note that, in our data elicitation, the [ne5] does not seem to be always necessary so [p^{h}ɛ] can stand alone.

(4) **p^{h}ɛ^{1}ne^{5} ʔawɛ sat^{h}ɔ lewɛ ne^{5}lɔ^{6}**
Then 1sg start go-prog final particle
"Then he started going."

Other literature, as mentioned above, does not necessarily detail on the subject however the glosses found in these pieces of work gives insight on the function of this particle. In Chappell's "Syntax and Semantics of the Benefactive Construction", Suriya's "Numeral Classifiers"'ne5' was not analysed however it appeared in the data sets. Out of interest, the different constructions were compared with our data set to see if the glosses were similar and if and how they differed. For now I will not present the data compared, but I will state that in some cases it was not glossed, in others it was classified as a demonstrative marker and in one case it was a topicalizer.

4.0 Analysis

This paper has adopted the Minimalism program as the basis for its analysis. This program is based on a feature checking theory, where each lexical item in a language is inherently equipped with a set of features. These features can be either interpretable (N, V, ...) or uninterpretable (uN, uACC...). When the uninterpretable features search for a matching interpretation, a process called feature checking, " AGREE" takes place. A set of constraints will dictate the movement in the tree as well as agreement patterns, but for the purpose of this paper only the relevant conditions will be explained when applicable.

Topicalization is a process by which the speaker places emphasis on a topic or focus. Similarily, topicalization can also be, as I believe is the case in Karen, a specificity feature, where the link of references is of importance to the understanding of the clause.

As mentioned above, the particle was seen in many different positions but always post-posed to the lexical item or the constituent (the exception ne^{5}lɔ^{6} will be explained later)"
The first example comes from the data set in Suriya's work:
***{Note that Suriya has adopted a three tone system for Karen, while we relied on a 6 tone system, so from the position and interpretation of ne^{1} in this clause, I have concluded that it is in fact ne^{5}}

(5) **na^{1}ri^{1} ɣe^{2} k^{h}i^{2} s^{h}o2 ne^{1}**
Watch good two CLAS that - - - according to Suriya's gloss
"those two good watches"

Arguably, the use of "ne^{5}" in this instance can in fact be as a demonstrative, however, it has also been noted that demonstratives (such as in German) tend to "topicalize quite liberally" (Rambow, 1993) Pragmatically, the latter view is more comprehensible and consistent with our analysis , keeping in mind that the language does have a lexical item for the demonstrative "that" [ʔa^{1}w ɛ^{1}ne^{5}], where the ne5 has probably been grammaticalized over the years.

Therefore, as per the pragmatic analysis given forth in the previous argument, our data elicitation of relative clauses has put forward the following examples:

(6) **pla^{1} (lɤ*) John ʔa^{1}w ɛ^{1}ne^{5} Ɵa^{4}k^{h}i^{1}**
Seem (that) John that happy*
"It seems that John is happy"

Compare this example with the following:

(7) **pla^{1}lɔ^{6} ɣa^{3} lɤ John-ne^{5} Ɵa^{4}k^{h}i^{1}**
Seem that John –TOP happy
"It seems that John is happy"

Complementizers are merged under C which holds a [+/- overt] feature. This feature will decide whether the complementizer is overt or not. However, a complementizer must always be governed by the verb in the preceding clause. The Empty Category Principle states that any empty category (in this case an empty complementizer category) must be governed by the verb in the main clause. It is relevant in example (6) and (7) in order to explain that the checking of the features by government can only check once, so in (7) the features are checked in [lɤ], under C; but in (6) the features are checked in [ʔa^{1}w ɛ^{1}ne^{5}], disfavouring any claim as to an empty C complementizer, or the coexistence of [lɤ] with [ʔa^{1}w ɛ^{1}ne^{5}].

As demonstrated by the preceding examples (6 & 7), if we allow for [ʔa^{1}w ɛ^{1}ne^{5}] to be present in the sentence, [lɤ] cannot coexist. [lɤ], in Karen is a discourse marker that acts as a prepostition (that, for, at....) but since in (6) [lɤ] is disfavoured, it raises the question of whether these two function words [ʔa^{1}w ɛ^{1}ne^{5}] and [lɤ] actually hold the same features and therefore on would remain uninterpreted causing the sentence to crash (be ungrammatical).

Conversely, in example (7) ne^{5} can be present in the same clause as [lɤ] because they do not carry the same features and therefore will not interfere with feature checking.
At first glance, the following example seems to contradict the analysis given:

(8) John tɛ^{1} ʔa^{1} lɤ^{0} ʔa^{1}w ɛ^{1}ne^{5} k ɤ^{0} h ɛ^{1}
"John told him that he would be coming"

Here before we gloss we must note that [lɤ^{0}] and [ʔa^{1}w ɛ^{1}ne^{5}] are co-existing in the same sentence. However we must be careful because here [ʔa^{1}w ɛ^{1}] is the pronoun "he" and is co-indexed with "John".

(8) John$_i$ tɛ^1 ʔa^1 lɤ^0 ʔa^{1}w ɛ^1 $_i$ ne^5 k ɤ^0 h ɛ^1
John tell him that he – TOP FUT- come
"John told him that he would be be coming"

And thus, the example provided in (9) just further supports the analysis of [ne^5] being a topicalization marker.

In Chappell's work, similar findings were present in the gloss:

(9) ***d*ɔ^4 dɔ^1 pɨ^2 wɛ^2 Θə^1 Ra^6 ta^2 lə^1 ne^2 wɛ^1 ne^5 hɛ^{1}ke^6 sɔ^3 wɛ^1**
*And brother 3 CLF thing REL **get** thing-TOP return carry thing*
"and the three brothers returned carrying the things which they had got."

In this example, aside from the inconsistency in comparison with our data (tə^1 ɣa^6 in our data was a classifier for person), the post-positioning of the particle to a substantive such as wɛ^1(thing) was very customary and was indicative of a specified object, in this case probably referring to a previously specified 'thing'.

However, it can be argued that the ne^5 also topicalizes the substantive within the relative clause. In English we topicalize (in relative clauses) by moving the constituent outside of the relative clause into the matrix clause, or what we call 'topic-fronting', it seems that in Karen, when you move this constituent outside of the relative clause, it is not a covert trace that is left behind but rather an overt copy along with a topicalization particle. If we take a closer look at the sentence provided, we see that ta^2 and wɛ^1 both are glossed as 'thing'.

According to minimalism, features are not only found inherently within the lexical item, but can also be found on specific nodes of the tree, such as a [uD] feature found on the spec of TP that will ensure the satisfaction of EPP at all times. Features of this nature can either be weak or strong, strong features being able to extract elements embedded in the tree (such as ones found inside a relative clause). For the purpose of topicalization in English, it has been proposed that a stong [uTop] feature is found which will front the topic of the sentence.

Analyzing our data from Karen in this minimalist program, we will see that not only do we need a strong [uTop] feature, but we need an explanation for the inserted or residual "ne^5". Here, due to the fact that we are arguing through the minimalist program, we have little leeway on inserting morphemes after movement has occurred. "The extension condition effectively prohibits setting up syntactic relations between objects which have already been constructed inside a structure.", and therefore disallowing the insertion of the ne^5 after the movement of the substantive has been made. However, if the morpheme was there already, then having it stay residually in the position of the moved substantive is not unlikely. It is in fact what happens in English Quantifier stranding. The spec in which a strong [uTop] feature would also have features only present in the substantive [ta^2].

i.e. **dɔ^1 pɨ^2 wɛ^2 Θə^1 Ra^6 *ta^2 $_i$* lə^1 ne^2 wɛ^1 $_i$ ne^5 hɛ^{1}ke^6 sɔ^3 wɛ^1**
*And brother 3 CLF thing REL **get** thing-TOP return carry thing*

Our data, seems to be in accordance with this analysis

(10) **ta^{2}$_{i}$lɔ^{2} l ɤ^{0} j ɤ^{0} ʔ ɔ^{5} ta^{2} $_{i}$- ne^{5} xa^{6} do^{5}ma^{4}lɔ^{6}**
Place LOC 1sg-eat place-TOP good very FIN
"The place where I eat is very good.

Whenever a classifier is used, it seems that the topicalization marker ne^{5} will be postposed to it inside the relative clause.

(11) **p^{h}o^{1}kwa^{1}$_{i}$ lɤ^{0} jɤ^{0} t^{h}i^{5} lɤ^{0}mɤ^{0}ha^{1} tɤ^{0}ɣa^{6}$_{i}$ - ne^{5} ʔa^{1}wɛ^{1} xa^{6} do^{5}ma^{4}**
Boy LOC 1sg-see yesterday one-person-TOP 3sg good very
"The boy that I saw yesterday is a good person."

It is important to note that topicalization in Sgaw Karen does not always induce fronting. Object topicalization is also prominent. This is not unusual for languages, especially if the object is cognate or etymologically related to the verb. Considering that Sgaw Karen is prevalent for cognate objects (even without a verb such as happy), it is quite predictable that such objects can be topicalized in situ.

(12) jɤ0 thi^{5} lɤ^{0} nɤ^{0} ʔɔ^{1} gwi^{1} nɤ^{0} ʔɔ^{1}- ne5 lɤ^{0}
1sg see that 2sg drink all your drink - TOP all
"I see that you drank all your drink."

Here the object is a cognate object and is in fact topicalized in situ. (Gimore)

5.0 Function of ne^{5} in giving directions

When giving directions in Sgaw Karen, it seems that the function of ne^{5} topicalizes the specific clause in which the route is indicated by acting as a connector between the two clauses.

(13) p^{h}ɛ^{1} nɤ^{0}lɛ^{6}tu^{6} lɤ^{0} lwi^{2} plɤ^{1}tɔ^{2} tɤ^{1}pɤ^{1}tɔ^{2} - ne^{5} lɛ^{6}s^{h}u^{1}sɨ^{1}twɛ^{1} tɤ^{1}ke^{1}
When 2sg go arrive to four level one CLASS TOP/CON go to hand right IMP
"When you get to the fourth floor, go to your right hand."

Since there are two route directions here, ne^{5} is placed between the two clauses to connect the topicalization, notice there is no need for a conjuction like dɔ^{4}.

(14) lɛ^{6} p^{h}ɛ^{1}ne^{5} dɔ^{4} kɤ^{1}na^{3} jɤ^{0}kɤ^{0}li^{2} dɔ^{4} tɤ^{0}klo^{3} pɛ^{1}tɤ^{0}rɔ^{1} tɤ^{1}ke^{1}
Go there and listen my voice and knock door IMP
"Go there and listen for my voice and knock on the door."

It is important to note that the example in (14) shows an imperative sentence that does not constitute of route directions and therefore the topicalization rules applied in (13) do not apply. However a ne^{5} is present postposed to a p^{h}e (when) which changes the meaning to "there".

6.0 Lexicalized or grammaticalized forms of ne^5

A few forms of ne5 have arisen in the data elicitation and seem to be constant regardless of topicalization. These forms include the postpositioning of ne^5 to p^{h}ɛ^1 (when) to create a new lexical entity [p^{h}ɛ^{1}ne^5] (there or then). Note that the meaning here is not topicalized, it has in fact taken up a whole new meaning.

(15) p^{h}ɛ^{1}ne^5 ʔawɛ sat^{h}ɔ lewɛ ne^{5}lɔ^6
then-particle 1sg start go-prog final particle
"Then he started going."

Other words such as ko^{4}ni^{6}de^{4}ne^5 seem to have the ne5 particle but in a lexicalized manner, where it is no longer clear to the native speaker that the ne5 is present.
Interestingly, both instances are locative of time and or place and this sheds light on the fact that there may have been diachronically a process that topicalized locatives of time and place but that in time these words were lexicalized.

7.0 Final particle or not

One form of ne^5 persists to be in non-topicalized positions, when preposed to lɔ^6 **[ne5 lɔ^6].** In both cases there is no evidence of topicalization, or at least none that favors the analysis of topicalization over others, hence, we argue that these two forms are in fact grammaticalized and that they are no longer using the particle in order to make the lexical item but rather are inherent within the lexeme.

(16) nja^5 d ɤ^{0}be^5 bu^2 ɣa^{2}ti^1 kle^1 do^{5}ma^4 ne^{5}lɔ^6
fish 1sg one swim fast very final particle [**ne^{5}lɔ^6]**
"The fish swims very fast."

It is important however to note that not all declarative final particles include the ne5 particle. Some only end in [**lɔ^6].** When there is no overt verb but rather an attribute makes the clause [**lɔ^6]** becomes the final particle used.

(17) jɤ^{0}dopu^{2}wɛ^{2}kwa^1 me^2 bwa^{6}no^1 kli^{1}po^1 dɔ^4 ko^{4}ni^{6}de^{4}ne^5 ʔa^{1}wɛ^1 le^6 l ɤ ʔa^{1}kli^1 ne^{5}lɔ^6
My brother is sailor boat person and everyday he go to his boat FIN
"My brother is a boat sailor and everyday he goes to his boat."

(18) **j ɤ pwi^6 ne^2 j ɤ k ɤsa^2 se^{1}ka^6 da^{1}be^5 dɔ^4 ʔa^{1}xi^{1}la^6 do^{5}ma^4 l ɔ^6**
I buy BEN myself clothes one CLASS and it beautiful very FIN
"I buy myself clothes and they are very beautiful."

The only difference between the two constructions is the verb construction, or the fact that the second part of the example in (18) does not contain an actual verb but rather an attributive.

8.0 Conclusion

Topicalization is a process that focuses on a specific part of a sentence or a clause. In Sgaw Karen, topicalization can be effectuated by movement or in situ. In either case, the use of a topicalization marker [ne5] is necessary. From the data collected in this short period of time, this paper is a preliminary analysis of its findings. Much more data would be necessary to allow for a more firm conclusion. In this paper, I have highlighted the main instances of appearance of the [ne5] particle and I have attempted to offer an inclusive topicalization analysis of the phenomenon. The main purpose for this argument was to explain the phenomenon pragmatically as well as syntactically in order to allow for predictions of its overtness.

References

Adger, David. Core Syntax. Oxford: Oxfprd Press, 2003. 424 p.

Burusphat, S. 1992, "Highlighting through the particle ?swe in Sgaw Karen narrative discourse", in Papers from the First Annual Meeting of the Southeast Asian Linguistics Society, ed. M. Ratliff and E. Schiller, pp. 47-63. Arizona State University, Program for Southeast Asian Studies

Culy, Christopher. "An incorporated topic marker in Takelma." CSLI Publications (2000).

Dryer, Matthew. "Word Order in Tibeto-Burman Languages1." SUNY Buffalo (April 2000).

Gimore, David. "A grammar of Sgaw Karen." American Baptist Mission Press, 1898.

Gyi, T.K. 1916, "Introduction to the study of the Sgaw Karen language", pp. 103.

Hargreaves, David. "From Interrogation to Topicalization: Proto-Tibeto-Burman *la in Kathmadu Nuwar." (California State University).

Her, One-Soon and Chen-Tien Hsieh. "On the Semantic Difference between Classifiers and Measure Words in Chinese." Language and Linguistics 11 (2010): 527-551.

Hopple, Paulette. "Topicalization in Burmese Expository Discourse." University of Texas at Arlington (n.d.).

Hosken, Martin. "A Sgaw Karen Unicode Proposal." SIL International and Payap University. n.d.

Larson, Richard. "Event Descriptions in Fon and Haitian Creole." n.d.

Laznick, Howard. Minimalist Syntax: The essential readings. USA: Oxford Press, 2007. 449 p

Park, Yoen-Mee. Topicalization in Terms of Specificity. 1996. 13 December 2010 <http://mlsk.or.kr/main/data2/mlsk2008090412414093.PDF>.

Rambow, Owen. 1993. 10 December 2010 <http://webcache.googleusercontent.com/search?q=cache:JaG_bbrpkGwJ:www.cs.columbia.edu/~rambow/papers/german-centering.ps+demonstratives+and+topicalization&cd=4&hl=en&ct=clnk&gl=ca>.

Carnie, Andrew. Syntax: A generative Introdution Second Edition, USA: Blackwell Publishing, 2008, 474 p.

Suriya, Ratanakul. "Numeral Classifiers in Sgaw Karen." Mahidol University (n.d.).

Thurgood, Graham and Randy LaPolla. The Sino-Tibetan Languages. New York: Routledge,2003.

The tale of Kinole and Nomie

Transcription by Dominike Thomas and Marc Brunelle

1. tə^{0}ni^{6}ŋa^{1}ʔi^{6} jə^{0}-kə^{0}-tɛ^{0}pja^{4}-θɨ^{1} ta^{2}χe^{6}pu^{1} tə^{0}-jɨ2 lə^{0} ʔə^{0}-me^{2} kɨ^{5}nɔ^{2}le^{1} dɔ^{4} nɔ^{2}mɨ^{5}ʔe^{1} ne^{5}lɔ^{6}
Today-DEM I-FUT–tell-2pp story one-CLASS QUOT ?- is Kinole and Nomie END.
Today I will tell you the story of Kinole and Nomie.

2. ta^{2}si^{4}wɛ^{1} lə^{0}plə^{6} tə^{0}-blɔ^{1} pwa^{6} dɔ^{1}mu^{1}wa^{6} ʔo^{5}-wɛ^{0} kʰi^{0}-ɣa^{6}
Saying past one-CLASS people wife-husband COP-PROG two-CLASS
Once upon a time, there were a husband and a wife.

3. ʔa^{0}-mi^{6} me^{2}-wɛ^{0}-da^{5} kɨ^{5}nɔ^{2}le^{1} dɔ^{4} nɔ^{2}mɨ^{5}ʔe^{1} ne^{5}lɔ^{6}
3p.-name is-PROG-? Kinole and Nomie END
Their names were Kinole and Nomie.

4. ʔa^{0}wɛ^{0}θe^{5} ʔɛ^{5}lo^{3} kwi^{0}lo^{3} ʔa^{0}-θa^{4} ɣe^{6} ma^{4}
They love love 3p.-REC good very
They loved each other very much.

5. dɔ^{4} ʔa^{0}wɛ^{0}θe^{5} ʔo^{5}-wɛ^{0}-da^{5} lə^{0}-θə^{1}wɔ^{1} tə^{0}plə^{5} ʔa^{0}-pu^{6} ne^{5}lɔ^{6}
and they live-PROG-? LOC-village one-CLASS 3p.-hole END
And they lived in a village.

6. ʔa^{0}-wa^{6}-ne^{5} me^{2}-wɛ^{0}da^{5} pwa^{6}-ma^{6}-kʰɨ^{4}-pʰo^{1} ʔa^{0}χo^{1} dɔ^{4} ko^{4}ni^{6}dɛ^{4}-ne^{6} ʔa^{0}-wa^{6} lɛ^{6} lə^{0}-kʰɨ^{4}-pu^{6} ne^{5}lɔ^{6}
3p-husband-TOP is-PROG-? pers.-do-dry.field-pers. so and every.day-TOP 3p-husband go LOC-dry.field-hole END
Her husband was a dry-farmer and so every day he went to the rice field.

7. tə^{0}-ɣɔ^{6}-ne^{5} pʰɛ^{1} ʔa^{0}-wa^{6} lɛ^{6} lə^{0}-kʰɨ^{4}-pu^{6} ʔa^{0}-kʰa^{1} tʰi^{5}lo^{3} ʔa^{0}-θa^{4} dɔ^{4} kə^{0}lɔ^{6} tə^{0}-bo^{1}
one-morning-TOP when 3p-husband go LOC-dry.field-hole ?-? meet 3p.-REC and python one-CLASS
One morning when the husband went to farm, he met with a python.

8. dɔ⁴ kə⁰lɔ⁶ tə⁰-bo¹-ne⁵ lo³kɛ⁰-tʰɔ⁵ ʔa⁰-θa⁴ lə⁰-pwa⁶-kə⁰ɲɔ¹ dɔ⁴ kə⁰lɔ⁶ tə⁰-bo¹-ne⁵ si⁴ʔɔ⁶ lə⁰
and python one-CLASS-TOP transform-? 3p-REC LOC-pers.-human and python one-CLASS-TOP say QUOT
And the python transformed himself into a human being and said:

9. nə⁰-kə⁰-ba⁵ kə⁰no¹ dɔ⁴ ja⁶ dɔ⁴ nə⁰-kə⁰-no¹ dɔ⁴ ja⁶ me² nə⁵-ne⁵ jə⁰-kə⁰-he⁵ na⁶ lə⁰ jə⁰-ta²tʰu⁴ta²tɔ⁶ kɛ¹lə³
2p-FUT-must wrestle with me and 2ps-FUT-wrestle with me COP win-TOP 1p-FUT give you ? 1p-property all
You must wrestle with me and if you win, I will give you all my property.

10. dɔ⁴ nə⁰-kə⁰no¹ dɔ⁴ ja⁶ me² tə⁰-nə⁶ tə⁰kɔ¹ jə⁰-kə⁰-hi⁴-kwi³ nə⁰-ma¹ ne⁵lɔ⁶
and 2p-wrestle with me COP not-win one-CLASS 1p-FUT-take-throw 2ps-wife END
And if you wrestle and lose, I will take away your wife.

11. dɔ⁴ ʔa⁰wɛ¹, ʔa⁰wɛ¹ kə⁰no¹ ʔa⁰wɛ¹ sa⁴tʰɔ⁵ kə⁰no¹ wɛ¹ dɔ⁴ pə⁰jo⁶ tə⁰-bo¹ ne⁵lɔ⁶
And he he wrestle he start wrestle PROG and dragon one-CLASS END
And then he started to fight the dragon.

12. ʔa⁰wɛ¹ kə⁰no¹ kə⁰no¹-wɛ¹ dɔ⁴ lə⁰-kʰi⁰-kə⁰tə²-ne⁵ ʔa⁰wɛ⁰ xa³-wɛ⁰ ʔa⁰wɛ¹ ʃu¹wɛ¹-da⁵ ne⁵lɔ⁶
he wrestle wrestle-PROG and LOC-?-end-TOP he lose-PROG he lose-? END
He wrestled and wrestled but in the end he lost.

13. dɔ⁴ pə⁰jo⁶-ne⁵ tə⁰te¹ʔɔ⁶ ta²nɔ¹tə⁰mi⁶-ba⁶ ʔa⁰wɛ⁰ lɛ⁶ ma⁶ kʰɨ⁴ wi⁶ dɔ⁴ hɛ⁰ki⁶ kə⁰da¹ lə⁰ ʔa⁰-hi⁵ ne⁵lɔ⁶
and dragon-TOP NG-say-3p nothing-NG he go work dry.field finish and return return LOC 3p-home END
And then the dragon didn't tell him anything, so he went back to his farm, finished his work and went back home.

14. lə⁰-kʰi¹ tə⁰-ɣɔ⁶ ʔa⁰wɛ¹ lɛ⁶ kə⁰da¹ lə⁰ ʔa⁰-kʰɨ⁴ dɔ⁴ ko⁴ni⁶dɛ⁴ ne⁵ ʔa⁰wɛ¹ ma⁶-wɛ⁰ di⁰-ne⁵ ʔa⁰θo⁴ ne⁵lɔ⁶
LOC-next one-morning he go return LOC 3p-dry.field and everyday TOP he work-PROG like-TOP same END
The next morning he goes back to his farm and he keeps working like that.

15. ba^{5}sa^{5}θə^{0}na^{3}ke^{0}-ne^{5} tə^{0}-ni^{6}-ne^{5} ʔa^{0}wɛ^{0} ʔo^{5} dɔ^{4} ta^{2}ɹɛ^{5}ta^{2}klɛ^{6} lə^{0} ʔa^{0}wɛ^{0} kə^{0}-lɛ^{6}-wɛ^{1} shu^{0} ta^{2}lɔ^{2} ʔa^{0}ɣa^{6} tə^{0}-tɔ^{6}
but-TOP one-day-TOP he have and plan that he FUT-go-PROG to place another one-CLASS
But one day he had a plan to go to another place.

16. dɔ^{4} pʰɛ^{0}ne^{5} ʔa^{0}wɛ^{0} tɛ^{0}χa^{3} ʔa^{0}-ma^{1} lə^{0} nə^{0}-me^{2} nə^{0}-me^{2} ʔo^{5} lə^{0}-hi^{5}-pu6-ne^{5} tə^{0}ba^{5} si^{5}lɔ^{6} lə^{0}-hi^{5}-pʰɔ^{0}la^{3} tə^{0}ɣe^{6}
and then he say 3p-wife QUOT 2ps-be 2ps-be stay LOC-house-hole-TOP NG go-fall LOC-house-ground NG.IMP
And then he told his wife: "Stay inside the house, don't go under the house.

17. ba^{5} mə^{0}ta^{6} lɛ^{5}-ne^{5} ta^{2}ʔə^{1}ta^{2}θɔ^{1} kə^{0}-hɛ^{0} ma^{6}ti^{3}ta^{3}-na^{6} tə^{0}-kʰa^{1} kʰa^{1} ne^{5}lɔ^{6}
but who WH-TOP bad.thing FUT-come destroy-2ps one thing thing END
because a bad thing will destroy you, one of these bad things."

18. pʰɛ^{1}ne^{5} ʔa^{0}wɛ^{1} sa^{4}tʰɔ^{5} lɛ^{6}-wɛ^{1} ne^{5}lɔ^{6}.
then he start go-PROG END
Then he started leaving.

19. ʔa^{0}-ma^{1}-ne^{5} ʔo^{5}-wɛ^{1} lə^{0}-hi^{5}-pu^{6} dɔ^{4} ko^{4}ni^{6}dɛ^{4}-ne^{5} ʔo^{5} lə^{0}-hi^{5}-pu^{6} dɔ^{4} tə^{0}-si^{5}-lɔ^{6} lə^{0}-hi^{5}-pʰɔ^{1}la^{3} ba^{5}
3p-wife-TOP live-PROG LOC-house-hole and everyday-TOP live LOC-house-hole and NG-go-fall LOC-house-ground NG
His wife lived in the house everyday and didn't go down under the house.

20. ba^{5}sa^{5}θə^{1}na^{3}kə^{1}-ne^{5} tə^{0}-ni^{6}ne^{5} ʔa^{0}wɛ^{1} tʰi^{5}-wɛ^{1}da^{5} lə^{0}-ʔa^{1}-tʰɔ^{4} tə^{0}-di^{1}-ne^{5} kə^{1}lɔ^{6} hɛ^{1} kə^{1}wɔ^{6}χa^{3}-ʔɔ^{6}
But-TOP one-day-TOP she see-PROG-? LOC -3p-pig one-CLASS-TOP python come constrict-3p
But one day she saw the python come surround her pig.

21. pʰɛ^{1}ne^{5} ʔa^{0}wɛ^{1} lɛ^{6} kwa^{2}-wɛ^{1} dɔ^{4} pʰɛ^{1}ne^{5} ʔa^{1}wɛ^{1} tɔ^{2}wɛ^{1} tɔ^{2}-so^{1}-wɛ^{1} lə^{0}-hi^{5}-pu^{6} dɔ^{4} hɔ^{1}-kwi^{3}wɛ^{1} dɔ^{4} ʔa^{1}ta^{2} tə^{0}-si^{5}-ba^{5}
then she go see-PRG and then she hit-PRG hit-fr.inside-PRG LOC-house-hole and shout-CAUS-PRG and ? NG-go-NG
Then she peeked and hit it from a distance and tried to shoo it, but the snake wouldn't leave.

22. dɔ^{4} lə^{0}-k^{h}i^{1}-ne^{5} ʔa^{0}-θi^{5} ʔo^{5} lə^{0}-ʔa^{1}-k^{h}o^{5}-lo^{6} tə^{0}-be^{5} ʔa^{0}-ta^{2} lɔ^{6}-tɛ^{3}-kwi^{3}-wɛ^{1} dɔ^{4} ʔa^{0}wɛ^{1} lɛ^{6}-jɔ^{2}-lɔ^{6} lə^{0}-hɔ^{5}k^{h}o^{5}-lo^{6}
and LOC-?-TOP 3p-comb stay LOC-3p-hair-surface one-CLASS 3p-thing fall-down-PROG-CAUS and she go-touch-fall LOC-soil-surf.
And later, her comb that was in her hair fell down and she stepped down on the ground [to pick it up].

23. dɔ^{4} p^{h}ɛ^{1} ʔa^{0}wɛ^{1} lɛ^{6}-lɔ^{6} lə^{0}hɔ^{5}k^{h}o^{5}-lo^{6} wi^{6} ʔa^{0}-lɔ^{2}-khi^{1}-ne^{5} pə^{1}jo^{6}-ne^{5} hɛ^{1}-p^{h}ɔ^{5}-ʔɔ^{6} dɔ^{4} ke^{6}twi^{2}-ʔɔ^{6} s^{h}u^{1} ʔa^{0}-pu^{6} ne^{5}lɔ^{6}
and when she go-fall LOC-soil-surface finish 3p-?-?-TOP python-TOP come-catch-3p and pull-3p to 3p-hole END
And when she went down on the ground, the python came, caught her and took her to his hole.

24. dɔ^{4} t^{h}o^{5} ʔo^{5}-wɛ^{1} tə^{1}be^{5} p^{h}ɛ^{1}ne^{5} lə^{0} ʔa^{1}t^{h}i^{5} klɨ^{4}ta^{2} dɔ^{4} t^{h}o^{5}-lwi^{2} tə^{0}-bɛ^{5}-ne^{5} lɛ^{6}-ju^{6}-wɛ^{1} s^{h}u^{1} ʔa^{0}-ʔo^{5} dɔ^{4} si^{4}ba^{5}-ʔɔ^{6} lə^{0}
and bird stay-PROG one-CLASS then that 3p.see everyth. and bird-dove one-CLASS-TOP go-fly-PROG to 3p-stay and tell-3p QUOT
And there was a dove which saw everything, flew to the husband and told him :

25. nə^{0}-ma^{1}ne^{5} ta^{2} hɛ^{1}-pɔ^{5}-ʔɔ^{6} li^{1} dɔ^{4} p^{h}ɛ^{1}ne^{5} ʔa^{0}wɛ^{1} kɨ^{5}nɔ^{2}le^{1} hɛ^{1}-ke^{6} kə^{0}da^{1}
2p-wife-TOP NOM? come-catch-3p COMPL and then he Kinole come-go once
"Your wife was taken away" and then Kinole went back home.

26. hɛ^{1}-ke^{6} dɔ^{4} xɨ^{1} ʔa^{0}-ma^{1} lə^{0}-hi^{5}-pu^{6} dɔ^{4} tə^{1} t^{h}i^{5}-lə^{6}-wɛ^{1} ba^{5}
come-go and look.for 3p-wife LOC-house-hole and NG see-?-PROG NG
And he went home to look for his wife and couldn't see anyone.

27. dɔ^{4} p^{h}ɛ^{1}ne^{5} ʔa^{1}wɛ^{1} θe^{5}ŋa^{1}-wɛ^{1} lə^{0} ʔa^{0}-ma^{1}-ne^{5} pə^{0}jo^{6} ke^{6}-twi^{1}-nɨ^{3}-ʔɔ^{6} s^{h}u^{1} ta^{1}-pu^{6} ne^{5}lɔ^{6}
and then he know-PROG QUOT 3p-wife-TOP dragon get-pull-?-3p to ?-hole END
And then he knew that the dragon pulled his wife into his hole.

28. dɔ^{4} p^{h}ɛ^{1}ne^{5} p^{h}ɛ^{1} ʔa^{1}wɛ^{1} lɛ^{1}-xɨ^{1}-ke^{6}-wɛ^{1} lə^{0}-pə^{1}jo^{6} ʔa^{0}-pu^{6} ʔa^{1}k^{h}a^{1} pə^{1}jo^{6} tɛ^{1}-ʔɔ^{6} lə^{0} nə^{0}-kə^{0}-ba^{5}
and then when he go-find-get-PROG LOC-dragon 3p-hole ? dragon say-3p QUOT 2ps-FUT-must

dɨ^{4}-ʔɔ^{1}-ja^{6} lə^{0} nə^{0}-θwi^{5} me^{2} wi^{6} ma^{4} jə^{0}-kə^{0}-plɛ^{1}lɔ^{6}-ne^{2}-ke^{6} nə^{0}-ma^{1} ne^{5}lɔ^{6}
give-drink-1ps LOC 2ps-blood is finish wife 1ps-FUT-release-get-get 2ps-wife END
And then when he went to get back his wife the dragon told him "You will have to feed me your blood before I let your wife go."

29. dɔ^{4} tə^{0} tə^{0}-pɔ^{1}tɔ^{2}-ne^{5} ʔa^{0}wɛ^{1} ma^{6}-kwa^{6}-wɛ^{1} ʔa^{0}wɛ^{1} hi^{4}ne^{2}wɛ^{1} lə^{0}
and one one-?-TOP he do-test-PROG he ?-get-PROG QUOT

sa^{5}p^{h}o^{1}kɔ^{2}p^{h}o^{1} t^{h}wi^{5} tə^{0}-di^{1} dɔ^{4} ku^{4}-ne^{2}-wɛ^{1} ʔa^{0}-θwi^{5}
animal dog one-CLASS and slice-get-PROG 3p-blood
The first time, he made a test : he took a dog and made an incision to get his blood.

30. dɔ^{4} pə^{1}jo^{6}-ne^{5} tɛ^{1}-ʔɔ^{6} lə^{0} ʔa^{1}wɛ^{1}-ʔi^{6} tə^{0} me^{2} nə^{0}-θwi^{5} ba^{5} tə^{0}-k^{h}i^{1}-kə-^{1}tə^{2} ʔa^{0}wɛ^{1} ku^{4}-lɔ^{6} ʔa^{0}-si^{1}-θwi^{5}
and dragon-TOP say-3 QUOT it-TOP NG is 2ps-blood NG one-?-?-? he slice-fall 3p-hand-blood

dɔ^{4} pə^{1}jo^{6}-ne^{5} tɛ^{1}-ʔɔ^{6} lə^{0} tə^{0} me^{2} nə^{0}-θwi^{5} ba^{5} ʔa^{0}xo^{1} jə^{0}-tə^{0}- ʔɔ^{1}-ba^{1}
and dragon-TOP say-3p QUOT NG is 2ps-blood NG so 1ps-NG-drink-NG
And the dragon told him: "This is not your blood" and in the end he took his hand blood and the dragon told him: "This is not your blood, I will not drink it."

31. dɔ^{4} lə^{0}-k^{h}i^{1}-kə^{1}-tɔ^{2}-ne^{5} ʔa^{0}wɛ^{1} ʔa^{0}-kə^{0}sa^{2} ba^{5} ku^{4}-lɔ^{6}-wɛ^{1} ʔa^{0}-ko^{5}bo^{1} ʔa^{0}-θwi^{5}
and LOC-?-?-?-TOP he 3p-REFL must slice-fall-PROG 3p-neck 3p-blood
And in the end he got his neck blood.

32. dɔ^{4} di^{4}-ʔɔ^{1} pə^{1}jo^{6} dɔ^{4} pə^{1}jo^{6} nə^{0}-kwa^{2} dɔ^{4} θe^{5}ɲa^{1}-wɛ^{1} lə^{0}
and serve-drink dragon and dragon smell-see and know-PROG QUOT

ʔa^{0}-kə^{0}sa^{2} ʔa^{0}-ko^{5}bo^{1} ʔa^{0}-θwi^{5} ʔa^{0}xo^{1} pə^{1}jo^{6} plɛ^{1}-lɔ^{6}-ne^{2}-ke^{6} ʔa^{0}-ma^{1} ne^{5}lɔ^{6}
3p-REFL 3p-neck 3p-blood so dragon release-fall-get-get 3p-wife END
He served the dragon, the dragon smelled it [the blood] and knew it was blood from his neck and let go of his wife.

33. dɔ^{4} p^{h}ɛ^{1} ʔa^{0}-ma^{1} hɛ^{1}-ki^{6}-t^{h}ɔ^{5} lə^{0}-ta^{2}-pu^{6} ʔa^{0}k^{h}a^{1} ʔa^{0}-ma^{1} t^{h}i^{5} lə^{0} ʔa^{0}-wa^{6}-ne^{5} θi^{1} li^{1} ʔa^{0}xo^{1} ʔa^{0}-ma^{1}ne^{5}
and then 3p-wife come-return-? LOC-?-hole ? 3p-wife see QUOT 3p-husb.-TOP dead COMPL so 3p-wife-TOP

kɨ²xo¹-sɔ¹me⁵-ke⁶ ʔa⁰-wa⁶ dɔ⁴ ʔa⁰-kə¹sa² na⁵ke¹ si⁵-lɔ⁶ θi¹-kwi³ ʔa⁰-θa⁴ lə⁰-me⁵ʔu¹-pu⁶ ne⁵lɔ⁶
cremate-RED-return 3p-husband and 3p-REFL also run-fall die-change 3p-REC LOC-fire-hole END
And then his wife came out from the hole and saw that her husband was dead, so she cremated her husband and killed herself by throwing herself into the fire.

kə²tə² li¹!
end finish
Finished!